Dynamique cérébrale en neuroimagerie fonctionnelle
Philippe Ciuciu

▶ To cite this version:

Philippe Ciuciu. Dynamique cérébrale en neuroimagerie fonctionnelle. Autre [q-bio.OT]. Université Paris Sud - Paris XI, 2008. tel-00333734

HAL Id: tel-00333734
https://theses.hal.science/tel-00333734

Submitted on 23 Oct 2008

Université de Paris-sud   Centre d'Orsay

Mémoire d'habilitation à diriger les recherches

Philippe Ciuciu

Dynamique cérébrale en neuro-imagerie fonctionnelle

soutenu le 10 juillet 2008 devant le jury composé de

M.	Jacques	Bittoun	Président
MM.	Jean-François	Mangin	Examinateur
	Andreas	Kleinschmidt	Examinateur
Mme	Line	Garnero	Rapporteure
MM.	Patrick	Flandrin	Rapporteur
	Xavier	Descombes	Rapporteur

Table des matières

III Projet de recherches 115

IV Références bibliographiques 123

V Publications et documents annexés 139

IX Imagerie fonctionnelle ultra-rapide 247

ORGANISATION DU DOCUMENT

Ce document est organisé en cinq parties :

- La première partie intitulée **Titres et travaux scientifiques**, retrace en trois chapitres les éléments d'un bilan factuel de mon activité scientifique.

  - Le premier chapitre, *Notice individuelle*, rassemble mon curriculum vitæ, mon activité d'encadrement, mes participations à des jurys de thèse, une synthèse de mes financements obtenus ainsi que mon investissement dans l'organisation de la recherche. Il se termine par le bilan de mon activité pédagogique depuis 1997 (formation initiale et continue) à une échelle nationale et internationale. Un résumé succinct de ma thématique de recherche est fourni à la fin de ce chapitre pour le lecteur qui souhaiterait faire l'économie de la lecture du chapitre III.

  - Le deuxième chapitre agrège au sein d'une *Notice bibliographique (mars 1999-mars 2007)* les travaux dont je suis auteur ou co-auteur en y distinguant les contributions relatives à ma thèse. Dans l'ensemble du document, les citations s'y rapportant sont faites exclusivement par des **sigles numérotés** entre crochets.

  - Le troisième chapitre, *Thèmes d'activité scientifique*, présente une synthèse quasi-exhaustive de mon activité de recherche depuis la fin de ma thèse en octobre 2000, découpée en thèmes, avec des indications chronologiques et la distinction nette entre des thèmes aboutis et des thèmes émergeants. Sont indiquées les collaborations nationales et internationales et le champ disciplinaire de chaque collaboration pour préciser en quelle qualité j'ai participé. À ce titre, il est également fait mention de façon synthétique des interactions au niveau du développement de logiciels scientifiques et de leur déploiement.

- La deuxième partie s'intitule **Dynamique cérébrale en neuro-imagerie**. Il s'agit d'une présentation synthétique et personnelle, en deux chapitres qui s'enchaînent, de ma principale thématique de recherche : l'extraction de la dynamique cérébrale à partir de données d'Imagerie cérébrale fonctionnelle (IRMf), son impact notamment sur les analyses de groupe, et le lien entre la dynamique de l'activité de « fond » et évoquée. Contrairement à la présentation éclatée du Chapitre III, la cohérence d'ensemble et la progressivité du propos sont privilégiées dans cette partie du document. Les applications en neuropsychologie cognitive et expérimentale sont mises en exergue afin d'illustrer l'intérêt des efforts méthodologiques consentis.

  - Le quatrième chapitre, *Impact de la modélisation non-paramétrique du signal BOLD sur les analyses de groupe en* IRM*f* propose une synthèse sur cerveau entier des approches régionales dont je suis à l'origine pour à la fois *détecter* les *variations* d'activité cérébrale en réponse à une stimulation extérieure tout en *estimant* la dynamique sous-jacente. Dans ce contexte, nous présentons les outils méthodologiques adaptés qui prennent place dans le cadre probabiliste bayésien. L'objectif *in fine* est de montrer l'impact d'une modélisation régionale et non-paramétrique du signal BOLD sur les analyses de groupe en IRMf (ici analyses à effets aléatoires uniquement). Nous explorons plusieurs perspectives à ce travail, certaines à court terme comme l'utilisation de modèles à effets mixtes pour l'analyse de groupe qui tiennent compte de la variabilité inter-individuelle potentiellement mieux estimée par nos modèles ; d'autres, à un peu plus long terme, qui

visent la recherche de modèles parcimonieux permettant d'expliquer les données, en vue notamment de capturer certains phénomènes neuro-dynamiques qui dépassent le cadre linéaire stationnaire.

– Le cinquième chapitre, *Analyse exploratoire multifractale* synthétise les travaux en cours sur l'utilisation en neuro-imagerie fonctionnelle d'une nouvelle classe de méthodes d'analyse exploratoire univariée dite *multifractale*. Ce type d'approches nécessite un grand nombre d'échantillons temporels pour fournir une estimée précise du spectre des indices de régularité recouvrés dans chaque signal. Cette approche tente de mettre en regard les données de repos et d'activation et de mieux appréhender la variation de la tortuosité du signal IRMf induite par la stimulation expérimentale. Pour fournir des résultats statistiquement fiables, ce type d'analyse nécessite des séries temporelles relativement longues. C'est pourquoi, nous avons privilégié l'usage d'une nouvelle séquence IRM 3D, dite *Echo Volumar Imaging* (EVI) développée au SHFJ puis à NEUROSPIN par Cécile RA- BRAIT au cours de sa thèse. L'imagerie EVI fournissant pour une même durée d'examen de l'ordre de dix fois plus de volumes fonctionnels, elle garantit une meilleure précision des estimateurs en ondelette mis en jeu, et autorise des conclusions neuropsychologiques pertinentes.

• La troisième partie présente mon projet de recherches à traavers un chapitre de *Conclusions et perspectives*. Je précise de façon synthétique les pistes méthodologiques sur lesquelles je souhaite travailler prochainement, pistes évoquées *in extenso* en conclusion des chapitres précédents. Je distingue ensuite les pistes de recherche que je souhaite emprunter ces prochaines années plutôt sur le versant neuro-scientifique, pour mettre à profit mes investissements méthodologiques sur des questions de neurosciences, tant cognitives que cliniques.

• La quatrième partie, **Références bibliographiques**, rassemble la bibliographie pour l'ensemble du document, hormis le contenu de la notice bibliographique du Chapitre 3. Dans tous les chapitres, les citations se rapportant à cette bibliographie principale sont faites exclusivement par nom d'auteur et année, entre crochets.

• La cinquième partie, **Publications et documents annexés**, est constituée de trois articles parus et d'un article soumis, donnant un aperçu représentatif de mes travaux.

Dans une certaine mesure, les Parties I, II, et les quatre documents de la Partie IV peuvent être lues indépendamment.

Première partie

Titres et travaux scientifiques

NOTICE INDIVIDUELLE

I.1 État civil

Philippe Ciuciu
42 rue de Verdun
91310 Longpont sur Orge
Tél. :   01 69 01 42 56   (personnel)
Tél. :   06 31 56 16 79   (mobile)
     01 69 08 77 85   (professionnel)

Né le 12 août 1973
Nationalité française
Marié, 1 enfant

Mél : philippe.ciuciu@cea.fr

Web : http://www.lnao.fr

I.2 Expérience professionnelle

2007– : **Ingénieur-chercheur** au Commissariat à l'Énergie Atomique (cea), au sein du Laboratoire de Neuro-imagerie Assistée par Ordinateur de Neurospin, dirigé par Jean–François Mangin.

2001–06 : **Ingénieur-chercheur** au cea, au sein du Service Hospitalier Frédéric Joliot (shfj), dans l'Unité de Neuro-imagerie Anatomo-Fonctionnelle (unaf) dirigée par Denis Le Bihan.

I.3 Formation initiale

2000–01 : **Post-doctorat** en **analyse statistique** de données d'irm fonctionnelle au cea/shfj/unaf, sous la responsabilité de Jean–Baptiste Poline.

2000 : **Doctorat** en **Traitement du Signal** au Laboratoire des Signaux et Systèmes (Supélec–cnrs–Université Paris-sud (ups)), mention très honorable, félicitations du jury. Directeur : Guy Demoment. Encadrant : Jérôme Idier.

1998–99 : Scientifique du contingent au Lena/cnrs upr 640 (Hôpital de la Salpétrière) : formation en **imagerie cérébrale fonctionnelle** (meg/eeg) au sein de l'équipe *Traitement des données en imagerie cérébrale* dirigée par Line Garnero.

1996 : dea d'**Automatique et de Traitement du Signal** (ats) de l'ups, mention bien. Diplômé en **traitement d'images** de l'École Polytechnique Fédérale de Lausanne.

1996 : **Ingénieur** de l'École Supérieure d'Informatique Électronique Automatique (esiea), Paris, filière Signaux et Images, mention très honorable avec les félicitations du jury.

I.4 Formation continue

07/2006 : Formation en méthodes variationnelles d'**assimilation de données**. École d'été CEA/EDF/INRIA à Saint-Lambert des Bois, dirigée par Dominique CHAPELLE (DR INRIA).

07/2004 : Formation en **analyse en ondelette et analyse multifractale**. École d'été WAMA à Cargèse, dirigée par Patrick FLANDRIN (DR CNRS, ENS de Lyon).

08/2003 : Formation en **neuro-biologie et électro-physiologie cérébrale**. École d'été *Temps & cerveau* à Concarneau, dirigée par Pierre MEYRAND, (DR CNRS, Bordeaux).

I.5 Prix et concours

08/2003 : Co-récipiendaire avec Guillaume MARRELEC du prix jeune chercheur à la conférence *Information Processing for Medical Imaging* 2003 (5 k£).

06/2001 : Classé 2e au concours N 0717 TI (section 07), CR2 au CNRS.

05/2001 : Lauréat des concours de Maître de Conférences N 0075/0448S en section 61 du Conseil National des Universités, affectés à l'ENS de Lyon et à l'ENSPM.

I.6 Expertise scientifique

2006 – : Expert international auprès du fonds anglais de financement de la recherche académique *Biotechnology and Biological Sciences Research Council* : http://www.bbsrc.ac.uk.

11/2007 : Membre invité du jury de thèse de Melle Cécile RABRAIT.

11/2006 : Membre examinateur du jury de thèse de Mme Salima MAKNI.

12/2005 : Membre examinateur du jury de thèse de Melle Anne BOTZUNG.

2002– : Membre du comité de lecture :
   – **des revues scientifiques** : *Human Brain Mapping, NeuroImage, Journal of Magnetic Resonance Imaging, IEEE Trans. on Signal Processing, IEEE Trans. on Image Processing IEEE Trans. on Medical Imaging, IEEE Trans. on Biomedical Engineering* ;
   – **des conférences internationales** *ICASSP, EUSIPCO, MICCAI, EMBC, HBM* et **nationale** *GRETSI*.

05/2006 : Organisateur d'une session spéciale à la conférence internationale ICASSP'06 intitulée : *Advanced methods for mapping brain functions from functional MRI datasets*.

09/2007 : Editeur invité de la revue *IEEE Special Topics in Signal Processing* pour le numéro spécial intitulé *fMRI Analysis for Human Brain Mapping* dont la date limite de soumission est fixée au 15/03/2008 (*cf.* http://www.et.byu.edu/groups/ece_jstsp/), en collaboration avec Tülay ADALI.

1999– : Auteur ou co-auteur de trize articles de revue internationale et d'une trentaine de communications dans des conférences nationales ou internationales, à comité de lecture. Voir http://www.lnao.fr/spip.php?rubrique3

I.7   Partenariats scientifiques

I.7.1   Au sein de l'ifr 49 « Neuroimagerie fonctionnelle », Paris

2007– : Collaboration avec Denis LE BIHAN et Toshihiko ASO à NEUROSPIN. <u>Thème :</u>« IRM de diffusion fonctionnelle : une origine neuronale ou vasculaire ? ».

2005–06 : Collaboration avec Ghislaine DEHAENE, (DR CNRS, U562), Christophe PALLIER (CR1 CNRS, U562) et Stanislas DEHAENE (Prof. Collège de France). <u>Thème :</u>« Mise en évidence du phénomène d'habituation à la répétition de phrases en IRMf ».

2001–04 : Collaboration avec Habib BENALI, (DR INSERM, U678), Guillaume MARRELEC (CR2 INSERM, U678). <u>Thème :</u>« Développement de méthodes d'estimation robuste de la réponse hémodynamique en IRMf. ».

2001–04 : Collaboration avec Christophe PALLIER et charlotte JACQUEMO. <u>Thème :</u>« Cartographie de la dynamique des réponses acoustique et phonologique chez des sujets européens et asiatiques par IRMf ».

2001–04 : Membre de l'action concertée incitative dirigée par Line GARNERO (DR CNRS, LENA–UPR 640). <u>Thème :</u>« Fusion de données multimodales en imagerie cérébrale fonctionnelle ».

I.7.2   À l'échelle nationale

2006–07 : Collaboration avec Thomas VEIT (CR2 INRETS), et Jérôme IDIER (DR CNRS, IRCCYN). <u>Thème :</u>« Échantillonnage de l'échelle dans un problème inverse bilinéaire ».

2006– : Responsable de l'équipe du SHFJ/CEA (5 personnes, 155 k€) au sein du projet ANR-OPTIMED financé par l'Agence nationale pour la recherche (380 k€), *cf.* http://optimed.univ-mlv.fr. Collaboration avec Jean-François MANGIN (Prof. Univ. Marne-la-Vallée (UMV)), Patrick-Louis COMBETTES (Prof. Univ. Paris VI) et Jérome IDIER. <u>Thème :</u>« Algorithmes d'optimisation décomposés pour les problèmes de reconstruction de grande taille en imagerie médicale ».

2005– : Collaboration avec Patrice ABRY (DR CNRS, ENS de Lyon). <u>Thème :</u>« Développement d'approches exploratoires *multifractales* pour la détection de l'activité cérébrale en IRMf ».

2005–07 : Collaboration avec Alice de FRÉMINVILLE (ingénieur d'études INSERM, U342) et Roland PEYRON, Practicien Hospitalier (CHU de Saint-Etienne). <u>Thème :</u>« Hémodynamique cérébrale de la douleur par IRMf ».

2005–06 : Collaboration avec Nicolas WOTAWA, doctorant à l'INRIA (UR de Sophia-Antopolis, équipe **Odyssée**), sous la direction d'Olivier FAUGERAS. <u>Thème :</u>« Dynamique de la sélectivité à la direction du mouvement dans les aires visuelles primaires IRMf ».

2004–05 : Collaboration avec Anne BOTZUNG, doctorante en neuropsychologie cognitive à l'Université Louis Pasteur, Strasbourg sous la direction de Lilianne MANNING. <u>Thème :</u>« Mise en évidence du flux d'information antéro-postérieur des activations cérébrales dans les processus de mémoire autobiographique ». Prix de thèse de l'ULP et du conseil général du Bas-Rhin en 2005.

2003–05 : Collaboration avec Sophie DONNET, ex-doctorante en statistiques à l'UPS sous la direction de Marc LAVIELLE. <u>Thème :</u>« Estimation et sélection de modèles en IRMf », actuellement maître de conférences à l'Université de Paris-Dauphine.

2003–05 : Collaboration avec Jean-Luc ANTON (ingénieur de recherches CNRS), responsable du centre de recherches en imagerie cérébrale à Marseille, Hôpital de La Timone, et Patricia ROMAIGUÈRE (CR1 CNRS, UMR 6562, Faculté des Sciences de St Jérôme). « Hémodynamique du contrôle du mouvement humain ».

<div align="right">20 juin 2008</div>

I.7.3 À l'échelle internationale

2007– : Collaboration avec Maria Gabriella Tanna (Institut polytechnique de Milan).<u>Thème :</u>« Analyse de la réponse hémodynamique associée à des pointes-ondes chez le patient épileptique sur des enregistrements EEG/IRMf simultanés ».

2006– : Contribution au projet NiPy (http://scipy.org/nipy) regroupant des acteurs des Univ. de Standford, Berkeley, Boulder (USA) et Cambridge (UK).<u>Object :</u> « Développement d'une architecture logicielle en Python pour la neuroimagerie ».

2005–06 : Collaboration avec Christophe Grova (Univ. de Montréal). <u>Thème :</u>« Nouveaux schémas d'interpolation du signal IRMf sur la surface corticale ».

2002–04 : Co-développement avec John Ashburner (Wellcome Department of Imaging neuroscience, London). <u>Thème :</u>« Conception du système de *batch* du logiciel SPM2/SPM5 ».

2002–04 : Co-développement avec Matthew Brett (Prof. MRC Cognition and Brain Sciences Unit, Cambridge). <u>Thème :</u>« Analyse de l'hémodynamique régionale. Interfaçage des boîtes à outils Marsbar et HRF sous SPM2 ».

I.8 Encadrement

I.8.1 Post-doctorats

05/2007- : Encadrement à 100 % du **post-doctorat** de Laurent Risser. <u>Sujet :</u> « Détection *in vivo* de l'activation neuronale par IRM de diffusion fonctionnelle : compréhension et analyse de la dynamique des phénomènes mis en jeu. ». <u>Financement :</u> allocation post-doctorale (18 mois) de la région Ile-de-France.

I.8.2 Thèses

10/2007- : Encadrement à 50 % de la **thèse** de Lotfi Chaâri en **Informatique** à l'UMV. <u>Sujet :</u> « Méthodes de reconstruction régularisée par ondelettes en IRM parallèle ». <u>Financement :</u> allocation doctorale obtenue auprès de la région Ile-de-France. <u>Directeur de thèse :</u> Jean-Christophe Pesquet.

12/2006- : Encadrement à 90 % de la **thèse** de Thomas Vincent en **imagerie médicale** à l'UPS. <u>Sujet :</u> « Fusion d'informations issues de données de neuro-imagerie fonctionnelle hétérogènes (IRMf/EEG) ». <u>Financement :</u> allocation doctorale obtenue auprès de la région Ile-de-France. <u>Directeur de thèse :</u> Jean–Baptiste Poline.

2004– 07 : Co-encadrement à 30 % de la **thèse** de Cécile Rabrait en **imagerie médicale** à l'UPS. <u>Sujet :</u> « Développement d'une séquence d'imagerie parallèle ultra-rapide (EVI) à 1.5 Tesla. ». Encadrant principal : Franck Lethimonnier. <u>Directeur de thèse :</u> Denis Le Bihan. <u>Financement :</u> contrat CFR, soutenue le 16/11/2007. <u>Situation actuelle</u> : ingénieur séquence chez Philips (Pays Bas).

2003–06 : Encadrement à 90 % de la **thèse** de Salima Makni en **imagerie médicale** à l'UPS. <u>Sujet :</u> « Détection-estimation conjointe de l'activité cérébrale mesurable en IRMf ». <u>Financement :</u> Allocation MESR, soutenue le 06/11/2006. <u>Directeur de thèse :</u> Jean–Baptiste Poline. <u>Situation actuelle :</u> Post-doctorante à l'Univ. d'Oxford (fM-RIB) sous la responsabilité de Steeve Smith.

I.8.3   Master 2 recherche et DEA

2008 : Encadrement du stage de **Master 2** en **Signaux et Images en Biologie et en Médecine** (Univ. de Bretagne Occidentale) de Marion SOUMOY. Sujet : « Comparaison et sélection de modèles en IRMf ».

2008 : Encadrement du stage de **Master 2** en **Mathématiques Vision Apprentissage** (ENS de Cachan) d'Anne-Laure FOUQUE. Sujet : « Parcellisation adaptative pour la détection-estimation conjointe de l'activité cérébrale en IRMf ».

2007 : Encadrement du stage de **Master 2** en **Imagerie médicale** (UPS) d'Ekaterina FALINA. Sujet : « Modélisation de l'habituation de l'activité cérébrale en IRMf ».

2007 : Encadrement du stage de **Master 2** en **informatique** (IMAC, UMV) d'Anne-Laure FOUQUE. Sujet : « Apport des approches intra-sujet de détection-estimation conjointe dans les analyses de groupe en IRMf ».

2006 : Encadrement du stage de **Master 2** en **Modélisation et calcul scientifique** de Thomas VINCENT (Univ. de Cergy-Pontoise). Sujet : « Apport d'un modèle de corrélation spatiale pour la caractérisation de l'activité cérébrale en IRMf ».

2003 : Encadrement du stage de DEA d'**Automatique et de traitement du signal** (UPS) de Salima MAKNI. Sujet : « Régularisation spatio-temporelle de la fonction de réponse hémodynamique en IRMf ».

2002 : Encadrement du stage de DESS de **mathématiques appliquées** (Univ. Paris VI) d'Alexis DAVID. Sujet : « Analyse de la stationnarité de la réponse BOLD en IRMf. ».

I.9   Enseignement et pédagogie

I.9.1   Enseignement universitaire

1997– : Activité (cours, TD/TP) en **Traitement du Signal et des Images**, **Problèmes inverses**, **Imagerie biomédicale**, **Informatique** et **Mathématiques appliquées** au sein des Univ. de Paris–Sud, de Marne-la-Vallée (Niv. License & Master), à l'ENS de Lyon, et en école d'ingénieur (ENSTA, ESIEA, ENSEA) : 550 heures équivalent TD. – depuis 2003 : co-responsable avec Patrice ABRY du module de M2 « processus aléatoires multifractals et problèmes inverses en imagerie » à l'ENS de Lyon.
– depuis 2005 : responsable du module d'**imagerie biomédicale** en 5$^{\text{ème}}$ année de l'ESIEA.
– 1999-2000 : co-responsable avec Christophe VIGNAT (MDC UMV) du module de **communications numériques** en 5A de l'ESIEA.

Les cours d'imagerie biomédicale dispensés à l'ESIEA sont validés par les étudiants sous forme de projets, de différente nature (analyse bibliographique, synthèse pédagogique *i.e.,* conception de site web ou fabrication de support multi-média visant un public non-averti, projet scientifique et informatique) d'un volume équivalent de 30 heures de travail environ. Ces projets font l'objet d'une soutenance orale.

I.9.2   Tableaux récapitulatifs

Cours magistraux	96-97	97-98	98-99	99-00	03-04	04-05	05-06	06-07
Problèmes inverses en imagerie, DEA Phys. non linéaire, ENS de Lyon					10h	10h		
Problèmes inverses en imagerie, M2 Phys. non linéaire, ENS de Lyon							10h	10h
Imagerie biomédicale, ESIEA 5$^{\text{e}}$ année							18h	18h
Systèmes de Communications numériques, ESIEA 5$^{\text{e}}$ année		9h	9h					
Formation MATLAB, DEA ATS, UPS		2h						

Travaux dirigés/TDAO	96-97	97-98	98-99	99-00	03-04	04-05	05-06	06-07
Problèmes inverses en imagerie, DEA Phys. non linéaire, ENS de Lyon					3h	3h	2h	
Informatique, ESIEA 1$^{\text{re}}$ année			114h	134h				
Analyse numérique matricielle, maîtrise EEA, UMV		15h						
Bases mathématiques et algorithmes pour le traitement du signal, ENSTA 2$^{\text{e}}$ année		28h						
Calcul scientifique, ESIEA 1$^{\text{re}}$ année				8h				
Projets Pluridisciplinaires, ESIEA 1$^{\text{re}}$ année			15h					
Projets Professionnels Personnalisés, DEUG B UPS		12h						

Travaux pratiques	96-97	97-98	98-99	99-00	03-04	04-05	05-06	06-07
Problèmes inverses, DEA ATS	15h							
Systèmes de Communications numériques, ESIEA 5$^{\text{e}}$ année			3h	3h				
Traitement du signal, licence et maîtrise EEA, UMV		36h						
Analyse numérique matricielle, maîtrise EEA, UMV		27h						
Imagerie biomédicale, ESIEA 5$^{\text{e}}$ année							6h	6h

| **Total équivalent TD** | 10h | 100h | 144h30 | 157h30 | 18h | 18h | 47h | 47h |

I.9.3   Formateur en imagerie cérébrale fonctionnelle

09/2006 : Intervenant à l'école d'automne organisée à Paris sur les « Méthodes avancées d'analyse des données MEG/EEG/IRMf » par Denis SCHWARTZ (CNRS UPR 640).

11/2004 : Intervenant (cours et TP) à l'école d'automne organisée à Marseille par Jean-Luc ANTON (CNRS) sur les « Méthodes avancées d'analyse des données d'IRMf ».

09/2004 : Co-responsable de la journée de formation à l'IRMf proposée à la conférence MICCAI'04.

09/2003 : Intervenant (cours et TP) aux journées de formation de la Société d'Anatomie Fonctionnelle Cérébrale, Paris.

08/2003 : Intervenant à l'école d'été *Temps & cerveau*, dirigée par Pierre MEYRAND (DR CNRS).

I.10   Diffusion logicielle et vulgarisation scientifique

– <u>Architecture logicielle</u> : concepteur et développeur d'une toolbox `HRF` de 2002 à 2006 pour `SPM99/SPM2` ainsi que d'un prototype du système de *batch* actuel de `SPM5` en Matlab.

– <u>Développement collaboratif</u> : gestion de sources (`Perforce`, `Svn`), documentation (`Doxygen`, `Epydoc`), refactorisation de code. Acteur du développement du *plugin* `neuroSpy` dédié à l'analyse de données d'IRMf sous BRAINVISA (logiciel de l'IFR 49 en langage `Python`). Développement collaboratif au sein du projet `Pyhrf` en `Python` conçu par Thomas VINCENT depuis 2006 pour l'analyse intra-sujet des données d'IRMf par des techniques d'inférence bayésienne de détection-estimation conjointe.

– <u>Support</u> : administrateur web du site http://www.madic.org de 2003 à 2007, désormais abandonné et non renouvelé pour cause de dissolution de l'ancienne équipe MADIC dirigée par Jean-Baptiste POLINE. Modérateur du forum de discussion HRF sur http://brainvisa.info

– <u>Vulgarisation scientifique</u> : Interviews données à France 2, TéléEssonne. Article de vulgarisation paru dans La Recherche (décembre 2005). Publication dans la revue interne *Talents* du CEA (février 2008).

I.11   Résumé de ma thématique scientifique

Je m'intéresse à l'analyse de la dynamique cérébrale à partir de données de neuro-imagerie fonctionnelle issues d'examens d'Imagerie par Résonance Magnétique fonctionnelle (IRMf). Mes travaux portent à la fois sur l'étude de la dynamique évoquée par un paradigme d'activation cérébrale et celle issue de l'activité spontanée ou de « fond » lorsque le sujet est au repos (*resting state*). Cela signifie que je développe à la fois des approches orientées modèles qui s'appuient sur une connaissance explicite du paradigme expérimental mis au point par les cogniticiens mais aussi des méthodes exploratoires, n'exploitant pas ces informations issues du paradigme – typiquement la connaissance des temps d'arrivée des stimuli présentés au sujet.

Ce thème de recherche embrasse à la fois des problèmes bas niveau relatifs à la reconstruction d'images en IRM mais aussi des aspects plus haut niveau qui concernent l'estimation et la sélection de modèles hémodynamiques régionaux non-paramétriques, capables de prendre en compte la variabilité inter-individuelle de la réponse impulsionnelle du système neuro-vasculaire. Les problèmes de reconstruction sont traités à l'aide de méthodes classiques de régularisation

dans l'espace image ou des méthodes plus évoluées opérant dans l'espace transformé des coefficients d'ondelette. Les aspects inférentiels haut niveau sont majoritairement abordés dans le cadre des statistiques bayésiennes.

Toute cette méthodologie a été appliquée pour tester des hypothèses neuro-scientifiques sur le fonctionnement cérébral chez le sujet sain (étude du langage) ou chez des patients atteints de maladie neurologique (épilepsie). Ces travaux ont été essentiellement conduits dans le volume cérébral mais depuis peu, nous tendons à contraindre l'analyse sur la surface corticale, substrat anatomique où siègent la majorité des macrocolonnes de neurones.

Mon activité de recherche comprend six axes qui sont résumés au Chapitre III. Je ne donne ici qu'une description extrêmement synthétique de ces différents thèmes pour fixer les idées et offrir une vision d'ensemble aux experts du conseil scientifique.

La première voie dans laquelle je me suis investi concerne l'estimation non-paramétrique de la fonction de réponse hémodynamique, c'est-à-dire la réponse impulsionnelle du système neuro-vasculaire. J'ai développé en particulier différentes approches régularisées pour l'identification de ce filtre dans différents contexte (voxel ou régional, mono- ou multi-sessions, paradigme événementiel lent ou rapide, mono- ou multi-conditions, ...). J'ai aussi mis au point des tests statistiques pour la détection des activations à partir de l'estimateur de cette fonction de réponse. J'ai par ailleurs participé à l'élaboration d'extensions de ces approches qui s'écartent de l'hypothèse de stationarité ou de reproductibilité de la réponse à travers les stimulations successives. J'ai finalement fait diffusé ces outils au travers d'une boîte à outils logicielle HRF, de 2002 à 2006, ce qui m'a permis d'établir plusieurs collaborations nationales et internationales (voir Section III.2.5).

Le second axe de mes recherches constitue une généralisation naturelle du premier dans lequel j'ai proposé un cadre statistique unifié pour opérer simultanément la détection des activations cérébrales et l'estimation de leur dynamique. Dans le cadre inférentiel bayésien, j'ai proposé différents modèles du signal BOLD (Blood Oxygen Level Dependent) mais aussi du bruit, pour opérer des analyses cerveau entier à partir d'une parcellisation fonctionnelle préexistante. La donnée d'une telle parcellisation définit des régions fonctionnellement homogènes sur lesquelles on estime des filtres hémodynamiques différents permettant de prendre en compte la variabilité régionale du système vasculaire. La validation et la comparaison des différents modèles est opérée à l'aide du calcul d'indices quantitatifs et objectifs à savoir des facteurs de BAYES. Le choix de la parcellisation optimale est un sujet d'intérêt. Le troisième axe de mes travaux n'est en fait au'un prolongement du précédent puisqu'il vise à développer des approches similaires mais cette fois sur la surface corticale. Pour ce faire, un certain nombre de problèmes spécifiques doivent être résolus, comme la segmentation des différents tissus cérébraux, ou la projection des signaux fonctionnels acquis dans le volume cérébral sur chacun des nœuds du maillage de cette surface. En collaboration avec d'autres équipes, je me suis surtout concentré sur ce dernier point.

La quatrième partie concerne la résolution d'un problème de reconstruction en Imagerie par Résonance Magnétique parallèle, c'est-à-dire lorsque le signal est recueilli sur un réseau d'antennes de surface en lieu et place d'une antenne volumique. Le problème à résoudre s'identifie à un problème de séparation de sources spatiales, chaque source correspondant à une antenne. La complexité et le caractère mal posé de ce problème sont liés à la géométrie du réseau de capteurs mais aussi au choix d'un facteur d'accélaration R. Ce facteur réalise un sous-échantillonnage de l'espace de Fourier dans au moins une des directions de l'espace et permet ainsi de réduire le temps d'examen par le biais d'acquisition d'images de plus petites dimensions sur chacune des antennes. Les profils de sensibilité des antennes étant spécifiques, l'inversion devient réalisable c'est-à-dire qu'il s'avère possible de reconstruction une image à champ de vue complet à partir d'un ensemble d'images à champ de vue réduit. A travers un contrat ANR établi en collaboration avec différents partenaires académiques, nous développons des algorithmes de reconstruc-

tion régularisée pour les images anatomiques et fonctionnelles, si possible non-supervisés ou auto-cailbrés. Nous nous intéréssons aussi à l'impact du choix de l'algorithme de reconstruction sur l'analyse statistique des données fonctionnelles, notamment la sensibilité avec laquelle nous pouvons détecter des activations cérébrales.

Cette thématique établit d'ailleurs une connexion avec le cinquième axe de recherche décrit dans ce manuscrit à savoir l'IRM fonctionnelle « ultra-rapide », qui concerne la caractérisation statistique d'une séquence d'imagerie fonctionnelle tridimensionnelle – l'*Echo Volumar Imaging* localisée parallèle – et la comparaison de la dynamique des signaux EVI à celle des signaux mesurés par une séquence EPI classique 2D. Le lien tient au fait que l'EVI n'est réellement opérationnel que lorsqu'on couple cette stratégie d'acquisition avec un schéma d'acquisition et de reconstruction parallèles, selon deux directions de l'espace (et non plus une seule). Le principal avantage de cette modalité est sa résolution temporelle, d'un ordre de grandeur plus précise que celle de l'imagerie EPI classique puisqu'elle permet d'acquérir un volume tous les 200ms. Ceci la rend particulièrement robuste au mouvement et donc adaptée à l'imagerie du nouveau né, fondamentale pour la compréhension de la mise en place des systèmes cognitifs chez le nourrisson (*e.g.,* langage).

Le dernier thème abordé dans ce manuscrit concerne le développement d'approches exploratoires pour d'une part mettre en évidence un phénomène d'invariance d'échelle sur les signaux cérébraux. D'autre part, il s'agit de comprendre les liens entre l'activité évoquée par une stimulation externe délivrée au cours d'un paradigme d'activation cérébrale et l'activité spontanée ou « de fond » mesurable chez le sujet à l'état basal communément appelé état de repos. Ces analyses exploratoires visent à extraire des indices intelligibles sur les signaux fonctionnels acquis. Pour ce faire, il est nécessaire que ces indices soient débarassés d'un certain nombre d'artéfacts physiologiques présents dans les données. C'est la raison pour laquelle les méthodes mises en œuvre procèdent dans un espace transformé (espace des coefficients d'ondelettes) et non sur les signaux eux-mêmes. Les indices de régularité du signal calculés nous permettent de mettre en évidence un phénomène d'invariance d'échelle sur ces signaux biologiques, de façon reproductible sur différents jeux de données, notamment sur des séries fonctionnelles EVI. La synthèse de ces indices au sein d'un spectre multifractal permet de montrer les fluctuations engendrées sur ce spectre par le phénomène d'activation cérébrale et d'en déduire certains comportements stationnaires ou non liés à ce phénomène.

20 juin 2008

Chapitre II

NOTICE BIBLIOGRAPHIQUE
(OCTOBRE 2000 – SEPTEMBRE 2007)

Les références annexées en Partie V de ce mémoire sont [a.6], en Annexe VII.1, page 141 ; [a.3], en Annexe VII.2, page 159 ; [a.7], en Annexe VII.3, page 177 ; [a.4], en Annexe VII.4, page 187 ; [a.8], en Annexe VIII, page 201 ; [**?**], en Annexe VIII.2, page 217 ; [a.12], en Annexe IX, page 247.

Les références [a.1, a.2, c.1, c.7, d.1, d.2, d.3, s.1, s.2, s.3, rc.1] correspondent à mes travaux de thèse et apparaissent ainsi en rouge.

II.1 Articles publiés dans des revues internationales avec comité de lecture (a.)

[a.1] **P. Ciuciu**, J. Idier et J.-F. Giovannelli, « Regularized estimation of mixed spectra using a circular Gibbs-Markov model », *IEEE Trans. Signal Processing*, vol. 49, 10, pp. 2202–2213, oct. 2001.

[a.2] **P. Ciuciu** et J. Idier, « A Half-Quadratic block-coordinate descent method for spectral estimation », *Signal Processing*, vol. 82, 7, pp. 941–959, juil. 2002.

[a.3] **P. Ciuciu**, J.-B. Poline, G. Marrelec, J. Idier, C. Pallier et H. Benali, « Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment », *IEEE Trans. Med. Imag.*, vol. 22, 10, pp. 1235–1251, oct. 2003.

[a.4] G. Dehaene-Lambertz, S. Dehaene, J.-L. Anton, A. Campagne, **P. Ciuciu**, G. P. Dehaene, I. Denghien, A. Jobert, D. Le Bihan, M. Sigman, C. Pallier et J.-B. Poline, « Functional segregation of cortical language areas by sentence repetition », *Hum. Brain Mapp.*, vol. 27, pp. 360–371, 2006.

[a.5] S. Dehaene, A. Jobert, L. Naccache, **P. Ciuciu**, J.-B. Poline, D. L. Bihan et L. Cohen, « Letter binding and invariant recognition of masked words : Behavioral and neuroimaging evidence », *Psychological Science*, vol. 15, 5, pp. 307–313, 2004.

[a.6] G. Marrelec, H. Benali, **P. Ciuciu**, M. Pélégrini-Issac et J.-B. Poline, « Robust Bayesian estimation of the hemodynamic response function in event-related BOLD MRI using basic physiological information », *Hum. Brain Mapp.*, vol. 19, 1, pp. 1–17, mai 2003.

[a.7] G. Marrelec, **P. Ciuciu**, M. Pélégrini-Issac et H. Benali, « Estimation of the hemodynamic response function in event-related functional MRI : Bayesian networks as a framework for efficient bayesian modeling and inference », *IEEE Trans. Med. Imag.*, vol. 23, 8, pp. 959–967, août 2004.

[a.8] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Joint detection-estimation of brain activity in functional MRI : a multichannel deconvolution solution », *IEEE Trans. Signal Processing*, vol. 53, 9, pp. 3488–3502, sep. 2005.

[a.9] B. Thirion, G. Flandin, P. Pinel, A. Roche, **P. Ciuciu** et J.-B. Poline, « Dealing with the shortcomings of spatial normalization : Multi-subject parcellation of fMRI datasets », *Hum. Brain Mapp.*, vol. 27, 8, pp. 678–693, aoÃ»t 2006.

[a.10] C. Grova, S. Makni, G. Flandin, **P. Ciuciu**, J. Gotman et J.-B. Poline, « Anatomically informed interpolation of fMRI data on the cortical surface », *Neuroimage*, vol. 31, pp. 1475–1486, 2006.

[a.11] B. Thirion, P. Pinel, A. Tucholka, A. Roche, **P. Ciuciu**, J.-F. Mangin et J.-B. Poline, « Structural analysis of fMRI data revisited : improving sensitivity and reliability of fMRI group studies », *IEEE Trans. Med. Imag.*, vol. 26, 9, pp. 1256–1269, sep. 2007.

[a.12] C. Rabrait, **P. Ciuciu**, A. Ribés, C. Poupon, P. Leroux, V. Lebon, G. Dehaene-Lambertz, D. L. Bihan et F. Lethimonnier, « High temporal resolution functional MRI using parallel echo volume imaging », *Journal of Magnetic Resonance Imaging*, vol. 27, 4, pp. 744–753, mars 2008.

[a.13] S. Makni, J. Idier, T. Vincent, B. Thirion, G. Dehaene-Lambertz et **P. Ciuciu**, « A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI », *Neuroimage*, vol. 41, 3, pp. 941–969, juil. 2008.

[a.14] A. Botzung, E. Denkova, **P. Ciuciu**, C. Scheiber et L. Manning, « The neural bases of the constructive nature of autobiographical memories studied with a self-paced fMRI design », *Memory*, vol. 16, 4, pp. 351–363, avr. 2008.

II.2   Revues sans comité de lecture (sc.)

[sc.1] J.-B. Poline, **P. Ciuciu**, A. Roche et B. Thirion, « Quelle confiance accorder aux images du cerveau en action ? », *Pour la Science*, vol. 338, pp. 138–142, déc. 2005.

[sc.2] L. Hertz-Pannier, **P. Ciuciu**, D. L. Bihan et V. Lebon, « Comprendre le cerveau par l'image », *Talents CEA*, vol. 99, pp. –, fév. 2008.

II.3   Chapitres de livre (bc.)

[bc.1] J.-B. Poline, **P. Ciuciu**, A. Roche et B. Thirion, « Intra and inter subject analyses of brain functional Magnetic Resonance Images (fMRI) », in *Biomedical Image Analysis : Methodologies and Applications*, N. Paragios, Ed. In press, 2008, Springer Verlag.

II.4   Articles soumis dans des revues internationales avec comité de lecture (s.)

[s.1] **P. Ciuciu**, P. Abry, C. Rabrait et H. Wendt, « Log-cumulant wavelet-leaders multifractal analysis of EVI fMRI time series : evidence of scaling phenomenae in the brain », soumis Ã journal of special topics on signal processing, numéro spécial sur le human brain mapping, CEA/I2BM/NeuroSpin, Gif-sur-Yvette, France, mars 2008.

[s.2] T. Vincent, L. Risser, J. Idier et **P. Ciuciu**, « Spatially adaptive mixture modellinf for within-suject analysis of brain activity in fMRI », soumis à IEEE Trans. Med. Imag., CEA/I2BM/NeuroSpin, Gif-sur-Yvette, France, juin 2008.

II.5 Communications dans des comptes rendus de conférences internationales avec comité de lecture (c.)

[c.1] **P. Ciuciu**, J.-F. Giovannelli et J. Idier, « Markovian high resolution spectral analysis », in *Proc. IEEE ICASSP*, Phoenix, az, mars 1999, pp. 1601–1604.

[c.2] **P. Ciuciu**, G. Marrelec, J.-B. Poline, J. Idier et H. Benali, « Robust estimation of the hemodynamic response function in asynchronous multitasks multisessions event-related fMRI paradigms », in *Proc. 1st Proc. IEEE ISBI*, Washington, DC, USA, juil. 2002, pp. 847–850.

[c.3] G. Marrelec, H. Benali, **P. Ciuciu** et J.-B. Poline, « Bayesian estimation of the hemodynamic response function in functional MRI », in *Bayesian Inference and Maximum Entropy Methods*, R. Fry, Ed., Baltimore, MD, USA, août 2001, MaxEnt Workshops.

[c.4] F. Kherif, G. Flandin, **P. Ciuciu**, O. Simon, H. Benali et J.-B. Poline, « Model based spatial and temporal similarity measures between series of functional Magnetic Resonance Images », in *Proc. 5th MICCAI*, Tokyo, Japon, sep. 2002, LNCS 2488 (Part II), pp. 509–516, Springer Verlag.

[c.5] G. Marrelec, **P. Ciuciu**, M. Pélégrini-Issac et H. Benali, « Estimation of the hemodynamic response function in event-related functional MRI : Directed acyclic graphs for a general Bayesian inference framework », in *18th IPMI*, C. Taylor et J. A. Noble, Eds., Ambleside, Royaume-Uni, 2003, LNCS-2732, pp. 635–646, Springer Verlag.

[c.6] **P. Ciuciu**, J. Idier, A. Roche et C. Pallier, « Outlier detection for robust region-based estimation of the hemodynamic response function in event-related fMRI », in *2th Proc. IEEE ISBI*, Arlington, VA, USA, avr. 2004, pp. 392–395.

[c.7] **P. Ciuciu** et J. Idier, « regularized doppler radar imaging for target identification in atmospheric clutter », in *29th Proc. IEEE ICASSP*, Montreal, Canada, mai 2004, vol. V, pp. 265–268.

[c.8] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Semi-blind deconvolution of neural impulse response in event-related fMRI using Gibbs sampler », in *2th Proc. IEEE ISBI*, Arlington, VA, USA, avr. 2004, pp. 860–863.

[c.9] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Semi-blind deconvolution of neural impulse response in fMRI using a Gibbs sampling method », in *29th Proc. IEEE ICASSP*, Montreal, Quebec, Canada, mai 2004, vol. V, pp. 601–604.

[c.10] S. Donnet, M. Lavielle, **P. Ciuciu** et J.-B. Poline, « Selection of temporal models for event-related fMRI », in *2th Proc. IEEE ISBI*, Arlington, VA, USA, avr. 2004, pp. 992–995.

[c.11] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Joint detection-estimation of brain activity in fMRI using an autoregressive noise model », in *3th Proc. IEEE ISBI*, Arlington, VA, USA, avr. 2006, pp. 1048–1051.

[c.12] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Bayesian joint detection-estimation of brain activity using MCMC with a Gamma-Gaussian mixture prior model », in *31th Proc. IEEE ICASSP*, Toulouse, France, mai 2006, vol. V, pp. 1093–1096.

[c.13] C. Rabrait, **P. Ciuciu**, C. Poupon, D. L. Bihan et F. Lethimonnier, « Temporal analysis of the BOLD response using high temporal resolution Echo Volumar Imaging », in *14th ISMRM*, Seattle, USA, mai 2006.

[c.14] B. Thirion, A. Roche, **P. Ciuciu** et J.-B. Poline, « Improving sensitivity and reliability of fMRI group studies through high level combination of individual subjects results », in *Proc. MMBIA2006*, New York, USA, juin17-18 2006.

[c.15] T. Vincent, **P. Ciuciu** et J. Idier, « Spatial mixture modelling for the joint detection-estimation of brain activity in fMRI », in *32th Proc. IEEE ICASSP*, Honolulu, Hawaii, , USA, avr. 2007, vol. I, pp. 325–328.

[c.16] **P. Ciuciu**, P. Abry, C. Rabrait, H. Wendt et A. Roche, « Leader-based multifractal analysis for EVI fMRI time series : ongoing vs. task-related brain activity », in *4th Proc. IEEE ISBI*, Arlington, VA, USA, avr. 2007, pp. 404–407.

[c.17] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Approche bayésienne de détection-estimation conjointe de l'activité cérébrale en IRMf », in *TAIMA'07*, Hammamet, Tunisie, mai 2007, pp. –.

[c.18] T. Vincent, **P. Ciuciu** et J. Idier, « Application and validation of spatial mixture modelling for the joint detection-estimation of brain activity in fMRI »,̧ in *Proc. of the 29th annual international conference of the IEEE EMBS*, Lyon, France, aoÃ»t 2007, pp. 5218–5222.

[c.19] C. Rabrait, **P. Ciuciu**, A. Ribés, C. Poupon, G. Dehaene-Lambertz, P. Leroux, D. L. Bihan et F. Lethimonnier, « Regularized localized parallel EVI : application to the study of habituation effects in fMRI », in *Proc. 16th ISMRM*, Toronto, Canada, mai 2008.

[c.20] L. Risser and **P. Ciuciu**, « Brain activation detection using diffusion weighted MRI and BOLD MRI : a comparative study », soumis à la conférence MICCAI, CEA/I2BM/NeuroSpin, Paris, France, mars 2008.

[c.21] **P. Ciuciu**, T. Vincent, A.-L. Fouque et A. Roche, « Improved fMRI group studies based on spatially varying non-parametric bold signal modeling », in *5th Proc. IEEE ISBI*, Paris, France, mai 2008, pp. 1263–1266.

[c.22] T. Vincent, **P. Ciuciu** et B. Thirion, « Sensitivity analysis of parcellation in the joint detection-estimation of brain activity in fMRI », in *5th Proc. IEEE ISBI*, Paris, France, mai 2008, pp. 568–571.

[c.23] L. Chaâri, J.-C. Pesquet, A. Benazza-Benyahia et **P. Ciuciu**, « Autocalibrated regularized parallel MRI reconstruction in the wavelet domain », in *5th Proc. IEEE ISBI*, Paris, France, mai 2008, pp. 756–759.

[c.24] L. Risser, T. Vincent et **P. Ciuciu**, « Fast estimation of partition function in 3D Potts fields. Application to fMRI image sequence analysis », soumis à NIPS'08, CEA/I2BM/NeuroSpin, Gif-sur-Yvette, France, juin 2008.

II.6 Communications dans des comptes rendus de conférences nationales avec comité de lecture (d.)

[d.1] J. Idier, J.-F. Giovannelli et **P. Ciuciu**, « Interprétation régularisée des périodogrammes et extensions non quadratiques », in *Actes du 16ᵉ colloque GRETSI*, Grenoble, France, sep. 1997, pp. 695–698.

[d.2] **P. Ciuciu**, J.-F. Giovannelli et J. Idier, « Analyse spectrale non paramétrique haute résolution », in *Actes du 17ᵉ colloque GRETSI*, Vannes, France, sep. 1999, pp. 721–724.

[d.3] **P. Ciuciu**, J. Idier et J.-F. Giovannelli, « Estimation spectrale régularisée de fouillis et de cibles en imagerie radar Doppler », in *Actes du 18ᵉ colloque GRETSI*, Toulouse, France, sep. 2001, pp. 479–482.

[d.4] **P. Ciuciu**, J. Idier, G. Marrelec et J.-B. Poline, « Estimation régularisée et non supervisée de la fonction de réponse hémodynamique en imagerie cérébrale fonctionnelle (IRMf) », in *Actes du 19ᵉ colloque GRETSI*, Paris, France, sep. 2003.

[d.5] S. Makni, **P. Ciuciu**, J. Idier et J.-B. Poline, « Détection-estimation conjointe de l'activité cérébrale en imagerie fonctionnelle (IRMf) », in *Actes du 20ᵉ colloque GRETSI*, sep. 2005, pp. 295–298.

[d.6] **P. Ciuciu**, J. Idier, T. Veit et T. Vincent, « Application du rééchantillonnage stochastique de l'échelle en détection-estimation de l'activité cérébrale par IRMf », in *Actes du 21ᵉ colloque GRETSI*, Troyes, France, sep. 2007, pp. 373–376.

[d.7] T. Vincent, **P. Ciuciu** et J. Idier, « Mélanges spatiaux pour la détection-estimation conjointe de l'activité cérébrale en imagerie fonctionnelle (IRMf) », in *Actes du 21ᵉ colloque GRETSI*, Troyes, France, sep. 2007, pp. 133–136.

II.7 Communications à des congrès sans actes (p.)

[p.1] P. Ciuciu, G. Marrelec, J. Idier, J.-B. Poline et H. Benali, « A general tool to estimate the hemodynamic response function in fMRI data », in *Proc. 8th HBM*, Sendai, Japon, juin 2002.

[p.2] P. Ciuciu, J. Idier, A. Roche, G. Flandin, G. Marrelec et J.-B. Poline, « On the spatial variability of the BOLD HRF and some regularization strategies », in *Proc. 9th HBM Neuroimage vol. 19 (2)*, New York, , USA, juin19–22 2003.

[p.3] S. Makni, P. Ciuciu, J. Idier et J.-B. Poline, « A region-based method for the estimation of the neural impulse response in event-related fMRI », in *10th HBM*, Budapest, Hongrie, juin 2004.

[p.4] S. Donnet, M. Lavielle, P. Ciuciu et J.-B. Poline, « BOLD single-trial variability and model selection », in *10th HBM*, Budapest, Hongrie, juin 2004.

[p.5] P. Ciuciu, C. Pallier, B. Thirion, S. Mériaux, G. Dehaene-Lambertz et S. Dehaene, « Hemodynamic response estimation in auditory sentence repetition », in *11th HBM*, Toronto, Canada, juin 2005.

[p.6] S. Makni, C. Grova, P. Ciuciu et J.-B. Poline, « An interpolation method for fmri data extraction on the cortical surface », in *11th HBM*, Toronto, Canada, juin 2005.

[p.7] S. Makni, P. Ciuciu, J. Idier et J.-B. Poline, « Anatomically informed joint detection-estimation of brain activity », in *12th HBM CD-Rom*, Florence, Italie, juin 11-15 2006.

[p.8] A. Botzung, P. Ciuciu, E. Denkova et L. Manning, « The neural bases of the constructive nature of autobiographical memories studied with a self-paced fMRI design », in *12th HBM CD-Rom*, Florence, Italie, juin 11-15 2006.

[p.9] P. Ciuciu, J. Idier et S. Makni, « Modeling non-linear and non-stationary effects of the BOLD response using mixture models in fMRI », in *12th HBM CD-Rom*, Florence, Italie, juin 11-15 2006.

[p.10] P. Ciuciu, C. Poupon, A. Ribés et C. Rabrait, « Vers de nouvelles méthodes de reconstruction régularisée en IRM parallèle », in *Journées Paristic'06*, Nancy, France, nov. 22-23 2006.

[p.11] P. Ciuciu, P. Abry et C. Rabrait, « Leader-based multifractal analysis of EVI fMRI time series : evidence of scaling phenomenae in a language comprehension study », in *Proc. 13th HBM*, Chicago, IL, USA, juin10–14 2007.

[p.12] T. Vincent, P. Ciuciu et J. Idier, « Whole brain validation of spatial mixture modelling for the joint detection-estimation of brain activity in fMRI », in *Proc. 13th HBM*, Chicago, IL, USA, juin10–14 2007.

[p.13] L. Chaâri, J.-C. Pesquet, A. Benazza, P. Ciuciu et C. Rabrait, « Regularized parallel MRI reconstruction in the wavelet domain », in *Grand colloque STIC-2007*, Cité des Sciences et de l'Industrie, La Villette, France, nov. 2007.

II.8 Séminaires, workshops, conférences invitées (s.)

[s.1] **P. Ciuciu**, J. Idier et J.-F. Giovannelli, « Nouveaux estimateurs du spectre de puissance », in *Colloque Jeunes Chercheurs Alain Bouissy*, Orsay, France, mars 1998.

[s.2] **P. Ciuciu**, J. Idier et J.-F. Giovannelli, « Analyse spectrale non paramétrique à haute résolution », in GDR-PRC ISIS, *GT1*, Paris, France, déc. 1999.

[s.3] **P. Ciuciu**, J. Idier et J.-F. Giovannelli, « Estimation spectrale markovienne – application en imagerie doppler », Bagneux, mai 2000, Thomson Airsys.

[s.4] **P. Ciuciu** et J.-B. Poline, « Haemodynamic response function estimation for any fmri experiment », LondonRoyaume-Uni, mai 2002, Wellcome Department of Imaging NeuroSicence.

[s.5] **P. Ciuciu** et J.-B. Poline, « Estimation de la fonction de réponse hémodynamique en IRM fonctionnelle », Paris, juin 2002, GdR ISIS.

[s.6] « Analyse intra-sujet en IRMf », Paris, France, sep. 2003, Journée de formation en neuroimagerie de la Société d'Anatomie Fonctionnelle Cérébrale.

[s.7] **P. Ciuciu**, « Modeling the BOLD response in fMRI », Saint-Malo, sep. 2004, MICCAI'04.

[s.8] **P. Ciuciu**, « Modélisation de la réponse hémodynamique en IRMf », Marseille, France, nov. 2004, Journée de formation en neuroimagerie.

[s.9] **P. Ciuciu**, « Within-subject analysis of fMRI data analysis : advanced bold signal models », Strasbourg, août 2005, Centre d'Études de Physiologie Appliquée, Unité CNRS UPS 858.

[s.10] **P. Ciuciu** et G. Marrelec, « Estimation and characterization of the hemodynamic response in fMRI », in *JIRFNI'06*, Paris, France, sep. 2006.

II.9 Rapports de recherche (rr.)

[rr.1] **P. Ciuciu**, « Modélisation régionale de la réponse hémodynamique en IRMf : prise en compte de la variabilité inter-conditions », Rapport interne, SHFJ/CEA, Orsay, France, déc. 2004.

[rr.2] **P. Ciuciu** et J. Idier, « Modèles de mélanges spatiaux adaptatifs », rapport technique, CEA/SHFJ, Orsay, France, avr. 2005.

[rr.3] **P. Ciuciu**, « Modèles hémodynamiques régionaux non-stationnaires », rapport technique, CEA/SHFJ, Orsay, France, nov. 2005.

[rr.4] **P. Ciuciu**, P. Abry et C. Pallier, « Log-cumulant wavelet-leaders multifractal analysis for fMRI time series », rejeté à Neuroimage, CEA/SHFJ, Orsay, France, nov. 2006.

[rr.5] L. Risser et **P. Ciuciu**, « Brain activation detection using diffusion weighted MRI and BOLD MRI : a comparative study », soumis à MICCAI'08, CEA/I2BM/NeuroSpin, Paris, France, mars 2008.

II.10   Thèse, dea (m.)

[m.1] **P. Ciuciu**, « Régularisation markovienne pour l'analyse spectrale non paramétrique. Application aux signaux de radars Doppler », Rapport de stage du DEA-ATS, Université de Paris-Sud, Gif-sur-Yvette, France, juin 1996.

[m.2] **P. Ciuciu**, *Méthodes markoviennes en estimation spectrale non paramétrique. Applications en imagerie radar Doppler*, Phd thesis, Université de Paris–Sud, Orsay, France, oct. 2000.

II.11   Documents pédagogiques (dp.)

[dp.1] **P. Ciuciu**, « Imagerie biomédicale : modalités, pré- et post-processing », nov. 2005.

[dp.2] **P. Ciuciu**, « Problèmes inverses en imagerie », déc. 2003–2005.

[dp.3] **P. Ciuciu**, « Tp d'analyse spectrale : méthodes classiques et avancées », mai 1999.

[dp.4] **P. Ciuciu**, « Systèmes de communications numériques », déc. 1999–2000.

II.12   Rapports de contrat de recherche (rc.)

[rc.1] **P. Ciuciu**, J.-F. Giovannelli et J. Idier, « Analyse spectrale post–moderne. application aux signaux radars », Rapport de contrat (confidentiel) CNRS–Société THOMSON, GPI–L2S, 1997.

CHAPITRE III

SYNTHÈSE DES TRAVAUX SCIENTIFIQUES

III.1   Introduction

III.1.1   Contexte général

Depuis l'an 2000, mon activité de recherche s'est inscrite au sein du groupe *Modèles et Analyse de Données en Imagerie Cérébrale* de l'Unité de Neuro-imagerie Anatomo-Fonctionnelle (UNAF) au sein du Service Hospitalier Frédéric Joliot (SHFJ/CEA), sous la responsabilité de Jean–Baptiste POLINE, d'abord pendant mon stage post-doctoral puis ensuite en tant que statutaire au sein de l'équipe. Depuis mon arrivée à NEUROSPIN début 2007, elle prend place au cœur du Laboratoire de Neuro-imagerie Assistée par Ordinateur dirigé par Jean–François MANGIN. Synthétiquement, elle se concentre sur des problèmes d'*analyse spatio-temporelle* de données d'Imagerie par Résonance Magnétique fonctionnelle (IRMf) et se positionne à l'interface de plusieurs disciplines, aux premiers rangs desquelles figurent les *statistiques*, le *traitement du signal*, l'*informatique* et les *neurosciences.*

Les études en neurosciences s'appuient en effet de plus en plus sur la neuro-imagerie anatomique et fonctionnelle, à la fois pour élucider le code neural propre à des fonctions cognitives évoluées chez le sujet sain (mémoire, calcul, langage) mais aussi pour identifier des dysfonctionnements structurels ou fonctionnels propres à certaines pathologies neurodégénératives (Alzheimer, sclérose en plaques) ou psychiatriques (autisme, schizophrénie), en vue d'établir des marqueurs du *diagnostic* voire du *pronostic* de ces maladies. Pour atteindre de tels objectifs, il est nécessaire d'extraire un ensemble d'informations pertinentes à partir du volume important de données acquises. La pertinence de ces informations dépend fortement de l'expertise acquise dans l'ensemble des disciplines scientifiques impliquées dans la démarche. Il est donc nécessaire de travailler à l'interface de ces différents domaines et de collaborer aussi bien avec des physiciens qu'avec des cogniticiens, les premiers permettant de mieux comprendre le contenu informationnel des données, les seconds garantissant la recherche de solutions cohérentes et viables en réponse à des questions de neurosciences correctement formulées. À travers ces années, j'ai pu collaborer avec des cogniticiens notamment de l'unité INSERM U562, *Neuro-imagerie cognitive* dirigée par Stanislas DEHAENE, afin de valider chez des sujets sains les approches développées. J'ai pu aussi travailler à l'amélioration de leur *robustesse* pour tenir compte de différentes sources de variabilité : inter-individuelle, inter-site, ... en vue de leur utilisation chez des patients notamment épileptiques [TANA et coll. 2007]. Ces variabilités sont en effet les causes essentielles de comportements atypiques de méthodes d'analyse classiques (*e.g.,* régression linéaire, ...). Je me suis aussi rapproché des physiciens en vue d'influencer le choix de certains paramètres d'imagerie ou de certains reconstructeurs, et d'améliorer l'analyse ultérieure des données reconstruites.

III.1.2   Pourquoi l'acquisition des données ne suffit pas en neuro-imagerie ?

Un prérequis indispensable à l'analyse pertinente des données de neuro-imagerie tient dans la maîtrise, au moins dans les grandes lignes, des techniques d'acquisition du signal de *résonance magnétique nucléaire*, et des principes de formation des images qui en résultent. La physique atomique, l'électromagnétisme et la biochimie sont donc essentielles pour parfaire ses connaissances en IRM. Depuis son invention au début des années 70[1], cette modalité d'imagerie a connu un essor considérable en permettant l'acquisition de volumes tridimensionnels de grande précision anatomique. Sous le terme IRM, on regroupe en fait diverses modalités qui ont en commun d'être toutes acquises avec le même imageur : images anatomiques (pondérées en T_1 ou T_2), images fonctionnelles (pondérées en T_2^*), images de diffusion (orientation des principaux faisceaux de fibres de matière blanche), ... Dans le domaine de l'imagerie cérébrale fonctionnelle, la formation des images s'appuie sur le contraste *Blood-Oxygen-Level Dependent* (BOLD), marqueur endogène mais indirect de l'activité neuronale. Toutefois, le polymorphisme de l'IRM l'autorise aussi à se muer en modalité d'angiographie (visualisation des vaisseaux cérébraux), de perfusion (distribution du sang) et de spectroscopie (distribution de certains métabolites) ; sauf mention contraire, nous nous focaliserons dans la suite exclusivement sur l'IRMf par effet BOLD.

Si la maîtrise des modalités d'acquisition demeure insuffisante pour proposer des interprétations directes et sans faille des images acquises, cela tient au caractère fortement bruité des données : le rapport signal-à-bruit (RSB) est faible, notamment dans les aires cérébrales attachées aux fonctions les plus évoluées (*e.g.,* cortex pariétal). Il n'est pas rare de mesurer un pourcentage de variation du signal BOLD consécutif à l'activation neuronale de l'ordre de 1 ou 2 % autour de la ligne de base. Ce constat établi, les techniques d'analyse de données de neuro-imagerie sont devenues incontournables pour produire des résultats et des interprétations neuro-scientifiques soutenables. Une fois les premiers systèmes sensori-moteurs cartographiés à l'aide de techniques de régression multiple, les cogniticiens se sont intéressés à des fonctions plus évoluées (mémoire, calcul, langage, ...) et les cliniciens, à des marqueurs physiopathologiques de certaines affections au sein de ces systèmes. C'est dans ce champ applicatif qu'ont pris leur essor des méthodes d'analyse plus ambitieuses, recherchant des phénomènes plus ténus au prix d'une modélisation accrue. C'est aussi pour étayer des interprétations neurobiologiques plus plausibles, et enfin répondre à de nouvelles questions[2], que sont apparues depuis ces cinq dernières années des méthodologies plus abouties (connectivité fonctionnelle, modèles non-linéaires, non-stationnaires, analyse sur la surface, ...), parmi lesquelles figurent mes contributions.

III.1.3   Nouvelles questions, changement de cap ?

Comme nous le verrons par la suite, la plupart des problèmes méthodologiques que j'ai abordés en neuro-imagerie fonctionnelle se formulent comme des *problèmes d'estimation* voire de *détection-estimation*. Sous un angle plus large, on peut aussi y rattacher des questions de *sélection de modèle*. Au sein de ma communauté scientifique d'origine, ces thématiques entrent dans la famille des problèmes inverses au sens où ils sont associés à un *problème direct*, exprimable et résoluble dans le cadre des équations régissant le fonctionnement des systèmes biologiques. Le regain d'intérêt qu'ont connu les problèmes inverses est dû à l'avènement de la puissance numérique des ordinateurs. On les retrouve donc dans bon nombre d'applications médicales ou autres (optique, radar, calorimétrie, spectroscopie, géophysique), bien au-delà de l'imagerie cérébrale. La

---

[1]Paul C. LAUTERBUR a reçu le prix Nobel de Médecine en 2003, pour ses travaux sur la découverte de l'IRM, qui ont révolutionné l'imagerie médicale. Il nous a quitté le 27/03/2007.

[2]en dépassant la simple cartographie fonctionnelle, qui ne répond finalement qu'à la question *Où ?*, et en s'attaquant à la problématique plus fondamentale du *Comment ?*

résolution de ces problèmes présente la caractéristique désagréable d'être très sensible aux erreurs sur les données, ce sont souvent des *problèmes mal-posés*[3]. Un moyen désormais classique de résoudre cette difficulté passe par la *régularisation* [Tikhonov et Arsenin 1977]. De façon privilégiée dans mon activité, les outils méthodologiques mis en œuvre pour parvenir à inverser les modèles sont des formulations probabilistes de type *markovienne* [Brémaud 1999], utilisées comme *a priori* dans le cadre de l'inférence bayésienne, ou encore des fonctions de pénalisation « gibbsiennes » qui en constituent le pendant dans un cadre déterministe [Idier 2001].

Compte tenu des contraintes expérimentales pour acquérir le signal d'irmf, l'objectif est souvent de construire des solutions *raisonnables* et *robustes* dans un contexte difficile : données fortement bruitées, artefacts de nature diverse, modèle d'observation approximatif, absence de modèle paramétrique du signal à reconstruire, ... La *théorie de l'estimation statistique* fournit un cadre adapté pour construire ces solutions. Mais il est alors important de remarquer que les études asymptotiques (par exemple à rapport signal-à-bruit ou à nombre de données tendant vers l'infini) ne rendent pas compte du caractère éventuellement mal-posé des problèmes en situation réelle. En traitement de signal, le choix d'un estimateur en fonction de son comportement asymptotique est classique et légitime *lorsque ce comportement correspond au régime établi dans un système alimenté par un flot de données*. Dans le contexte spécifique des problèmes inverses, cette démarche devient risquée : ignorant la nécessité de régulariser, elle est susceptible de produire des solutions inexploitables car instables.

Dans le domaine de la neuro-imagerie, il a fallu attendre le début des années 2000 pour que ce constat soit communément admis et que les méthodologistes et cogniticiens soient conscients de cette nécessité de régulariser. En effet, jusque là, la plupart des acteurs considéraient le problème d'analyse de données résolu dès lors que la question de neurosciences à laquelle il se rattachait était correctement formulée — l'exemple le plus frappant constituant la batterie de pré-traitements[4] requis pour pouvoir détecter une corrélation entre le signal fonctionnel mesuré et le paradigme expérimental[5]. L'« intrusion » de traiteurs de signaux et d'images et de statisticiens de stature internationale[6] dans la communauté de la neuro-imagerie fonctionnelle a impulsé des changements radicaux dans l'analyse de ces données depuis le début des années 2000. Ils ont réussi le tour de force de faire émerger une méthodologie innovante dans l'analyse (*e.g.,* cesser de lisser spatialement les données, résoudre le problème de la normalisation spatiale, conduire des analyses de groupe robustes aux données atypiques ou tenant compte du caractère non-gaussien de la population échantillonnée, ...) tout en démontrant un vif intérêt pour les neurosciences et une capacité à communiquer avec les acteurs de ce domaine. C'est aujourd'hui de cette communauté que je me sens le plus proche et au sein de laquelle j'essaie d'apporter ma contribution de façon originale.

Ce changement de cap dans l'analyse des données a pour but essentiel le développement d'algorithmes produisant des résultats résistant aux *méta-analyses*, c'est-à-dire dont les interprétations neuro-scientifiques sont *reproductibles* et *généralisables* à d'autres individus ou d'autres populations. C'est dans cette perspective d'une plus grande complexité que se place mon travail. Cette complexité se décline sur le plan méthodologique et algorithmique par l'usage d'une gamme d'outils relativement sophistiqués et gourmands en temps de calcul. La résolution

---

[3]Un problème est bien posé au sens de Hadamard [Idier 2001] si sa solution existe, est unique et demeure stable vis-à-vis d'une petite perturbation des données, donc en particulier si elle dépend continûment des données. Un problème est dit mal-posé si l'une au moins des trois conditions précitées n'est pas satisfaite. Voir [Tikhonov et Arsenin 1977, Kaipio et Somersalo 2005] pour des illustrations.

[4]étapes de correction mais aussi de destruction d'une partie de l'information contenue dans les données

[5]On définit le paradigme expérimental comme la suite de stimuli présentés au sujet dans le scanner, au cours duquel on acquiert les volumes fonctionnels.

[6]On pourrait citer les membres du Fmrib à Oxford, le groupe d'Ed Bullmore à Cambridge, Keith Worsley à Mc Gill, Jonathan Taylor à Standford, le fil à Londres et en particulier Will Penny, Alexis Roche ou Bertrand Thirion à Neurospin...

de problèmes calculatoires fait partie intégrante de mon activité afin de rendre crédibles et exploitables par le plus grand nombre les approches que nous développons. J'ai par exemple investi le champ des méthodes d'optimisation déterministe (*cf.* **Thèmes 1** et **4**) [Tanner 1993, Bertsekas 1995], d'analyse combinatoire (**Sous-thème 1c**) [Paschos 2005], mais aussi celui des techniques d'échantillonnage stochastique, aussi connu sous le vocable de méthodes de Monte-Carlo par Chaînes de Markov [Robert 2001] au moment d'aborder les problèmes d'estimation non-supervisée (réglage d'hyperparamètres, *cf.* **Thèmes 1** et **2**).

Outre les aspects algorithmiques, une étape fondamentale à la validation et la diffusion de nos méthodes passe par un important travail d'implantation informatique, de documentation et de support aux utilisateurs. Elle nécessite des infrastructures que peu de laboratoires de recherche possèdent et que peu de personnes valorisent comme une activité scientifique. En France, la seule plate-forme informatique *multimodale*[7] dédiée à la neuro-imagerie anatomique et fonctionnelle a vu le jour au sein de l'IFR 49 au début des années 2000. Il s'agit de Brainvisa, au sein de laquelle nous proposons une gamme d'outils d'analyse originaux. Sur le plan international, nous nous inscrivons aussi dans une politique multi-laboratoires en contribuant au projet NiPy, *i.e., Neuroimaging in Python*[8] à travers le développement d'une librairie C de calcul scientifique[9] et des *bindings* `Python` associés. En régime de croisière, cette production logicielle s'accompagne d'une activité de conseil pour des équipes de recherche extérieures. À moyen terme, cette expertise va nous permettre d'atteindre un niveau d'exigence et de qualité que sont en droit d'attendre des partenaires industriels, cliniques et pharmaceutiques, dont certains collaborent déjà avec Neurospin à travers le projet franco-allemand Iseult financé par l'Agence pour l'innovation industrielle (Aii) depuis novembre 2006.

<div align="center">*   *
*</div>

La suite de ce chapitre est un résumé de mon activité scientifique, intitulée *Neurodynamique cérébrale*, découpée en six thèmes :
- Thème 1 – Estimation non-paramétrique de la réponse hémodynamique (2000-06)
  - Sous-thème 1a – Identification non-supervisée (2000-03)
  - Sous-thème 1b – Statistique de détection et de comparaisons (2003-04)
  - Sous-thème 1c – Extensions spatiales et recherche d'homogénéité fonctionnelle (2002-06)
  - Sous-thème 1d – Extensions non-stationnaires (2003-06)
  - Sous-thème 1e – Applications en neuropsychologie cognitive et expérimentale (2003-06)
- Thème 2 – Détection-estimation conjointe de l'activité cérébrale (depuis 2004)
  - Sous-thème 2a – Modèles de mélanges indépendants (2004-06)
  - Sous-thème 2b – Modélisation du bruit (2004-06)
  - Sous-thème 2c – Modélisation de la corrélation spatiale (depuis 2006)
  - Sous-thème 2d – Modélisation de déactivations (depuis 2006)
  - Sous-thème 2e – Modèles non-linéaires et/ou non-stationnaires (depuis 2005)
  - Sous-thème 2f – Indétermination liée à la bilinéarité (depuis 2006)
  - Sous-thème 2g – Sélection et comparaison de modèles (depuis 2007)
- Thème 3 – Analyse sur la surface corticale (depuis 2005)
  - Sous-thème 3a – Interpolation du signal IRMf sur la surface (2005-06)

---

[7]c'est-à-dire capable de traiter des données issus d'examens d'imagerie par résonance magnétique (anatomique, fonctionnel, de diffusion) mais aussi de magnéto- ou électro-encephalographie (MEEG).

[8]Voir http://projects.scipy.org/neuroimaging/ni.

[9]baptisée `fff` pour *fast functions for fMRI*, téléchargeable sur le dépot SVN à l'adresse http://neuroimaging.scipy.org/svn/ni/fff.

– Sous-thème 3b – Détection-estimation conjointe sur la surface (depuis 2007)
- Thème 4 – Reconstruction régularisée en IRM parallèle (depuis 2006)
    – Sous-thème 4a – Reconstruction supervisée (depuis 2006)
    – Sous-thème 4b – Reconstruction auto-calibrée dans l'espace ondelette (depuis 2007)
- Thème 5 – IRM fonctionnelle « ultra-rapide » (depuis 2005)
    – Sous-thème 5a – Caractérisation statistique des données EVI (2005-07)
    – Sous-thème 5b – Caractérisation de la dynamique des signaux EVI (2005-07)
    – Sous-thème 5b – Applications en neurosciences : mise en évidence du phénomène d'habituation (2007)
- Thème 6 – Analyse exploratoire multifractale des données d'IRMf (depuis 2005)
    – Sous-thème 6a – Indice de régularité des données d'activité induite et spontanée
    – Sous-thème 6b – Détection d'activation par analyse statistique des attributs multifractaux

L'ordre choisi est approximativement chronologique et les périodes d'activité sont indiquées pour chaque thème et sous-thème. Le découpage adopté rend compte de la complexité croissante des modèles élaborés au cours de mes travaux, en vue notamment de l'élaboration de méthodes pertinentes pour la fusion *symétrique* de données multimodales de données de neuro-imagerie fonctionnelles (EEG-IRMf)[10]. Pour chaque thème, les applications privilégiées ayant servi à la validation des modèles et algorithmes proposés seront précisées et les collaborations mentionnées. La fin du chapitre est consacrée à un résumé des perspectives de ces travaux, portant notamment sur deux sujets majeurs :

– le développement de modèles génératifs et explicatifs pour la fusion symétrique de données multimodales de neuro-imagerie fonctionnelle ;
– l'analyse de la dynamique des données de diffusion fonctionnelle, qui nous l'espérons, reflètera l'activité neuronale plus précocément que le contraste BOLD.

III.2 Thème 1 – Estimation non-paramétrique de la réponse hémodynamique (2000-06)

Dans ce thème s'inscrit mon travail post-doctoral au SHFJ ainsi que la collaboration avec l'équipe INSERM dirigée par Habib BENALI, de l'unité U678, durant la thèse de Guillaume MARRELEC. Le problème abordé est celui de l'*identification* de la réponse impulsionnelle du système neuro-vasculaire encore appelée *fonction de réponse hémodynamique* (FRH) dans les protocoles *événementiels* d'IRMf. Dans ce type de paradigmes en effet, des stimuli de durée brève sont présentés isolément au sujet, par opposition aux protocoles en blocs. Ils induisent ainsi des variations du signal BOLD typiques, qui sous des hypothèses de linéarité et stationarité de la réponse correspondent à la réponse impulsionnelle du système neuro-vasculaire. Si les différents stimuli sont suffisamment espacés dans le temps (typiquement 20 à 25 sec.), on parle de protocole événementiel lent. Le problème d'identification n'est alors pas difficile car les réponses BOLD aux différents essais ne se chevauchent pas. En revanche, pour des intervalles inter-stimuli plus courts (*e.g.,* 3 à 4 sec.), c'est-à-dire pour des protocoles événementiels rapides, l'estimation de l'hémodynamique devient plus complexe.

Cette question est devenue un domaine de recherche très actif essentiellement pour deux raisons. La mise en évidence de l'impact d'une bonne estimation locale de la FRH sur la détection des activations en constitue la première. De façon connexe, il s'agit de sélectionner des régresseurs de façon optimale, en un sens à définir, pour modéliser la réponse attendue du sujet

---

[10] J'ai déjà eu l'occasion d'aborder cette question sous un angle *asymétrique* dans le cadre de l'action concertée incitative de Line GARNERO durant la période 2001-04, en collaborant aussi avec Pierre-Jean LAHAYE ancien doctorant de l'équipe au SHFJ.

au cours du protocole expérimental. En d'autre termes, selon la forme de la FRH considérée, des aires cérébrales impliquées par le paradigme expérimental peuvent potentiellement ne pas être détectées comme significativement activées, et nécessiter ainsi la définition d'un modèle plus souple par l'ajout de régresseurs[11]. D'autre part, les tentatives pour expliquer la forte variabilité régionale et inter-individuelle observée dans la forme de cette dynamique constitue indéniablement une seconde motivation. Les sources de cette variabilité peuvent être de nature vasculaire mais aussi neuronale. Dès lors, beaucoup d'efforts ont été consentis pour tenter de séparer les contributions de ces deux compartiments, en jouant notamment sur le paradigme expérimental (phénomène de *suppression à la répétition*) mais aussi sur la modélisation de la réponse (voir **Sous-thème 1a**).

Dès le milieu des années 90, la plupart des travaux de ce domaine ont été consacrés à la recherche du meilleur modèle paramétrique [FRISTON et coll. 1995, LANGE 1997, COHEN 1997, RAJAPAKSE et coll. 1998] et du meilleur jeu de paramètres au sens des moindres carrés une fois ce modèle posé. Un consensus s'est dégagé autour de la contribution de [GLOVER 1999], qui a proposé, étayé par des validations expérimentales, de représenter la FRH comme une différence de deux distributions gamma. Néanmoins, il est rapidement apparu que cette modélisation pouvait s'avérer trop rigide pour prendre en compte les variabilités précédemment mentionnées [AGUIRRE et coll. 1998, BUCKNER et coll. 1998, MIEZIN et coll. 2000, HANDWERKER et coll. 2004]. L'introduction d'une certaine souplesse dans le modèle a été abordée selon plusieurs points de vue. Pour certains, il s'agit d'utiliser une base de fonctions en lieu et place d'une seule forme paramétrique [FRISTON et coll. 1995, JOSEPH ET HENSON 1999], mais la question du choix de la meilleure base demeure en suspens. Cette question a en partie été résolue par les travaux successifs de [GENOVESE 2000, GÖSSL et coll. 2001] portant sur une modélisation semi-paramétrique, puis de [WOOLRICH et coll. 2004], qui ont finalement proposé une base relativement souple dont la dimension peut évoluer en fonction de l'adéquation des données aux modèles. Ce degré d'évolutivité a un prix puisque la sélection du meilleur modèle nécessite la résolution d'un problème d'optimisation dont la dimension de l'espace de recherche n'est pas fixée *a priori*.

Tout comme les travaux précurseurs de [GOUTTE et coll. 2000], j'ai poursuivi une autre voie, qui consiste à recourir à une formulation non-paramétrique pour estimer la FRH, c'est-à-dire sans imposer de forme particulière à l'avance. Mes principales contributions dans ce domaine sont synthétisées dans les paragraphes suivants.

III.2.1   Sous-thème 1a – Identification non-supervisée (2000-03)

Mes principales publications relatives à cette partie [c.2, a.6, a.3, a.7] et concernent de nombreux développements liés à l'estimation non-paramétrique et *voxel à voxel* de la FRH dans un cadre bayésien. L'aspect non-supervisé tient au fait qu'en sus des paramètres hémodynamiques à identifier les hyper-paramètres (statistique du bruit, modèle *a priori*) sont estimés automatiquement à partir des données avec d'ailleurs une stratégie d'estimation fluctuante selon les publications : maximum de vraisemblance, moyenne *a posteriori*,... Selon la nomenclature anglo-saxonne, on peut regrouper les solutions proposées en deux familles :

1. Les approches « bayésiennes empiriques » (*cf.* [a.3]) qui consistent à estimer les hyper-paramètres au sens du maximum de vraisemblance à partir des données, puis à injecter les valeurs estimées dans la distribution *a posteriori* de la FRH, sans tenir compte de l'incertitude sur ces estimées. Autrement dit, l'on suppose ici que les hyper-paramètres sont inconnus mais déterministes. Ainsi, on fait l'hypothèse que la valeur estimée coïncide avec la

---

[11]Dans cette situation, le nombre de degrés de liberté des tests statistiques sous-jacents devient plus faible et les tests moins sensibles.

vraie valeur du paramètre, ce qui évidemment n'est qu'une approximation. Les algorithmes mis en jeu sont des techniques déterministes de maximisation de la vraisemblance, tels que les algorithmes EM (*Expectation-Maximization*) [DEMPSTER et coll. 1977, WU 1983] ou leurs avatars (ECM, GEM) [MENG et RUBIN 1993].

2. Les approches « totalement bayésiennes », qui généralisent les précédentes puisqu'on probabilise les hyper-paramètres en leur assignant une distribution de probabilité. Il s'agit alors soit d'estimer la distribution *a posteriori* marginale (*cf.* [a.6, a.7]) de l'hyper-paramètre d'intérêt ou d'échantillonner sa distribution *a posteriori* conditionnelle (*cf.* [c.8]). Un tel schéma permet *in fine* d'accéder à la distribution marginale *a posteriori* du paramètre et donc de tenir compte de l'incertitude sur l'estimée.

Par rapport au travail séminal de [GOUTTE et coll. 2000], mes principales contributions concernent l'estimation de l'hémodynamique pour des paradigmes *événementiels rapides*, nécessitant donc la déconvolution des réponses aux essais successifs du stimulus. En collaboration avec Guillaume MARRELEC, j'ai notamment proposé des généralisations concernant :

– les paradigmes *multi-conditions*[12] où il s'agit d'estimer une dynamique spécifique à chaque type de stimulus ; ceci tend à rendre le problème moins *bien-posé* car le nombre d'inconnues grandit au regard du nombre de données disponibles. Toutefois, cette modélisation est capitale pour mieux comprendre au sein d'une région fonctionnelle donnée quelle part de la variabilité de l'activité neuronale est perceptible dans les signaux d'IRMf. La modélisation d'une fonction de réponse par condition expérimentale permet en effet d'isoler localement la contribution due au changement de l'activité neuronale sous réserve que le paradigme expérimental soit bien contrôlé, c'est-à-dire que les stimuli soient de durée et d'intensité comparables à travers les conditions expérimentales. Toutefois dans ce cadre, cette séparation n'est pas directement généralisable à l'ensemble du cerveau compte tenu de la variabilité du réseau vasculaire et de ses effets sur la forme de la réponse BOLD.

– les paradigmes *asynchrones* où les stimuli sont présentés à intervalles de temps variables et potentiellement décalés par rapport aux instants d'acquisition des volumes fonctionnels ; la situation devient plus complexe numériquement ici car la matrice décrivant le problème direct perd son caractère Toeplitz.

– la modélisation et l'estimation de la dérive basse fréquence affectant les signaux d'IRMf, évitant ainsi tout filtrage passe haut préalable détruisant, sans recours possible, une partie de l'information contenue dans les données. Sur le plan algorithmique, nos premières contributions ont d'abord considéré ces paramètres de nuisance comme déterministes mais inconnus pour en proposer une estimation au sens du maximum de vraisemblance au sein d'un algorithme ECM [a.3]. La qualité de l'estimation s'est améliorée significativement en adoptant une approche complètement probabiliste [d.4]. Il s'agit ici en fait d'intégrer hors du problème les coefficients de dérive basse fréquence puis de travailler sur la loi *a posteriori* jointe marginale.

– la prise en compte de plusieurs sessions d'acquisition au sein du même voxel en vue de faire croître le nombre de données disponibles au regard du nombre d'inconnues. Évidemment, cette approche fait l'hypothèse d'une certaine reproductibilité de la FRH à travers les sessions, hypothèse discutable pour certains paradigmes expérimentaux voire chez certains sujets. C'est l'une des raisons fondamentales pour lesquelles j'ai opté par la suite pour le développement de modèles spatiaux, *i.e.,* exploitant l'information de voisinage, c'est-à-dire de similarité des profils de réponse dans des voxels voisins.

La Fig. III.1 illustre les deux premiers points en montrant des différentiations de profils de FRH estimée pour différentes conditions expérimentales (silence, contraste acoustique et phonologique, ou contrôle) entre le gyrus de Heschl (*cf.* Fig. III.1(a)) et deux zones distinctes du

---

[12]qui sont les plus intéressants en pratique pour bien contrôler la tâche à réaliser.
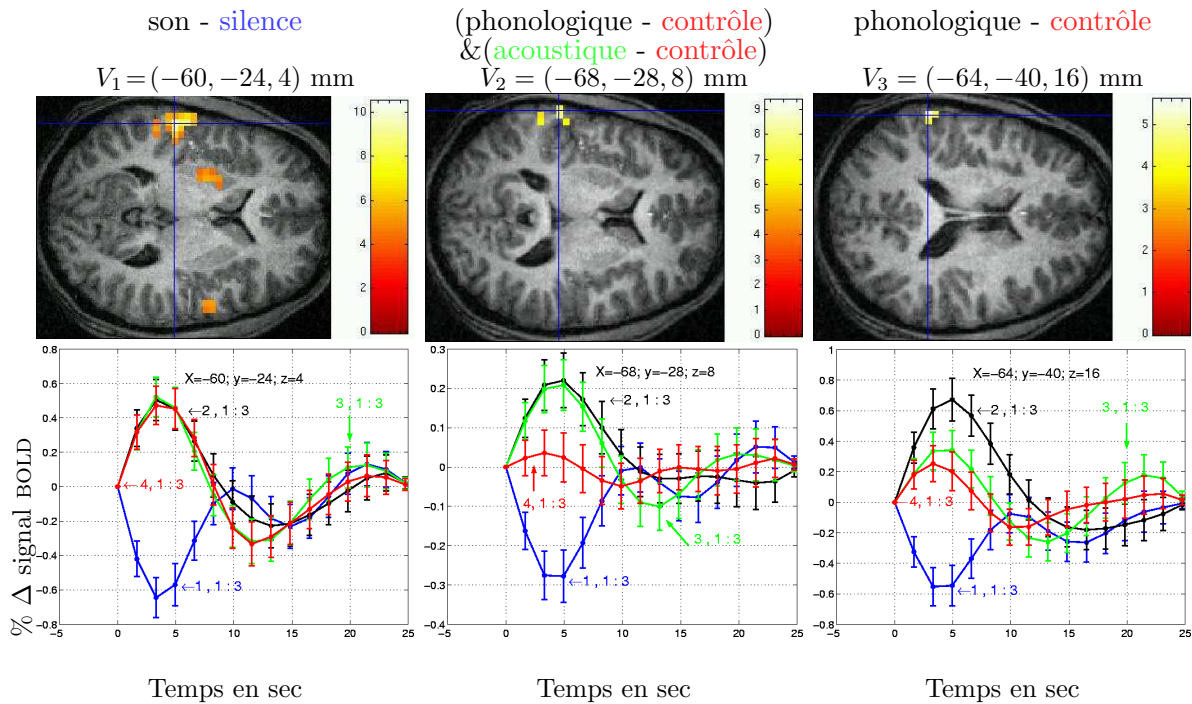
Fig. III.1 – Estimation multi-conditions et multi-sessions de la FRH dans une expérience de perception linguistique. Les conditions expérimentales **phonologique**, acoustique et contrôle sont associées aux FRHs $\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3$, respectivement et les essais silence sont représentés par $\boldsymbol{h}_4$ montrant une déactivation. En haut, les cartes statistiques de Student seuillées à $P = 0.001$ et corrigées pour les comparaisons multiples sont superposées à une coupe axiale de l'anatomie cérébrale d'un sujet. Les coordonnées des voxels sont indiquées en mm dans le référentiel de Talairach. En bas, les FRH sont estimées à partir de trois sessions dans les voxels V_1-V_3.

planum temporale impliquées respectivement dans l'encodage acoustique et phonologique (*cf.* Fig. III.1(b)-(c)). On peut également visualiser la présence de déactivations dues aux essais silence, bien prédictibles compte tenu de l'arrêt des gradients de l'IRM au cours de ces essais présentés entres les scans d'acquisition (paradigme asynchrone).

Ayant observé dans bon nombre de situations expérimentales une certaine similarité dans les formes de réponses estimées dans des voxels voisins (voir Fig. III.2), il nous a semblé pertinent de travailler à une extension spatiale de ces travaux, c'est-à-dire caractériser l'hémodynamique non plus d'un voxel mais d'une région.
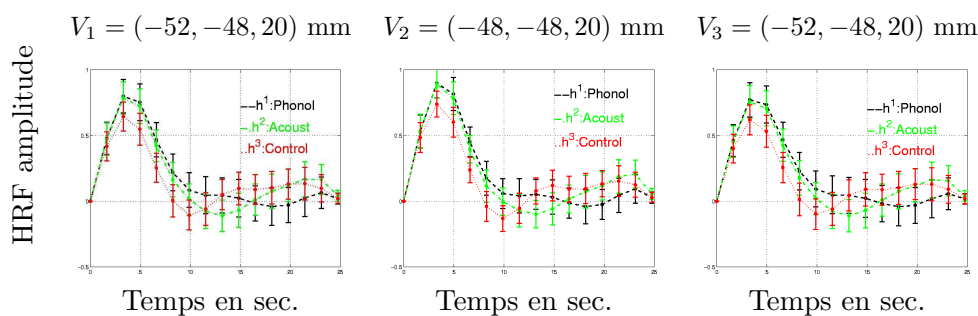
Fig. III.2 – Estimées des FRHs dans trois voxels d'une région d'intérêt illustrant une certaine reproductibilité spatiale de la dynamique recouvrée.

III.2.2 Sous-thème 1c – Extensions spatiales et recherche d'homogénéité fonctionnelle

Les extensions spatiales que j'ai proposées s'appuient sur la notion d'« homogénéité fonctionnelle régionale ». En première approximation, on pourrait définir cette notion comme la résultante de l'évaluation par une métrique appropriée, du degré de similarité de signaux IRMf issus de voxels voisins.

Mes travaux se sont alors articulés autour de deux questions :

1. Identifier au sein d'un système cognitif particulier des aires corticales dont les propriétés fonctionnelles sont discernables en termes d'amplitude et de délai d'activation (voir Fig. III.3 et [a.4, WOTAWA 2006]) ; pour ce faire, nous nous appuyons sur des algorithmes de parcellisation qui rassemblent au sein de chaque parcelle des voxels dont le profil fonctionnel est semblable tout en préservant la connexité des parcelles [a.9]. Ce profil peut être directement issu des signaux IRMf mesurés ou bien des paramètres de regression issus d'une analyse classique par modèle linéaire généralisé (MLG) comme l'illustre la Fig. III.3(c).

2. Tester l'homogénéité fonctionnelle d'une région préalablement définie, soit comme un cluster d'activation soit sur des bases anatomiques (*cf.* [c.4, a.3, c.6]). La contribution méthodologique la plus originale a consisté à adapter au contexte de l'IRMf une méthode d'estimation robuste, les moindres carrés *tamisés* [ROUSSEEUW ET LEROY 1987], dans le but d'écarter de l'analyse une proportion de voxels atypiques au sein d'une région fonctionnelle donnée. La mise en œuvre de cette approche, détaillée dans [c.6], s'appuie sur un algorithme de type SMLR (*Single Most Likely Replacement*) bien connu de la communauté des traiteurs de signaux [KORMYLO ET MENDEL 1982, MENDEL 1983, CHAMPAGNAT ET COLL. 1996] pour résoudre de façon sous-optimale un problème d'optimisation combinatoire.

Finalement, cela m'a conduit aussi à proposer un modèle explicatif régional du signal BOLD, suffisamment parcimonieux pour que son identification reste résoluble avec un coût de calcul réduit [c.6, c.9]. En substance, j'ai proposé d'introduire un modèle régional où la forme de la FRH est la même pour tous les voxels, tout en autorisant une certaine variabilité de la réponse hémodynamique d'un voxel à un autre par le biais d'une modulation d'amplitude. Ainsi, caractériser l'hémodynamique d'une région équivaut à identifier ces facteurs de modulation[13] en sus du décours temporel de la FRH. Ceci a fait l'objet du stage de DEA de Salima MAKNI au printemps 2003.

Le problème inverse linéaire d'identification d'une FRH se mue alors en un problème inverse *bilinéaire*, plus complexe, mais pas insurmontable car la linéarité conditionnelle est préservée. Nous avons contourné cette difficulté d'abord en introduisant une contrainte sur l'énergie de la FRH pour lever l'ambiguïté d'échelle sous-jacente. Nous verrons au chapitre IV qu'une contribution théorique récente permet de résoudre ce problème tout en s'affranchissant de cette contrainte (*cf.* échantillonnage de l'échelle [VEIT ET IDIER 2007, d.6]).

Il est ensuite possible d'étendre cette modélisation au cas de plusieurs conditions expérimentales en considérant deux options :
– soit la FRH reste identique indépendamment du type de stimulus présenté. Cette hypothèse n'est valide que si bien sûr aucune variabilité du délai hémodynamique n'a été constatée sur des voxels de la région pris isolément. Dans ce cas, il suffit d'introduire un facteur de modulation d'amplitude spécifique de la condition expérimentale et de chaque voxel. Cette hypothèse a été développée dans la première partie de la thèse de Salima MAKNI ; voir notamment [c.8, c.9].

---

[13]baptisés niveaux de réponse neuronale au chapitre IV.
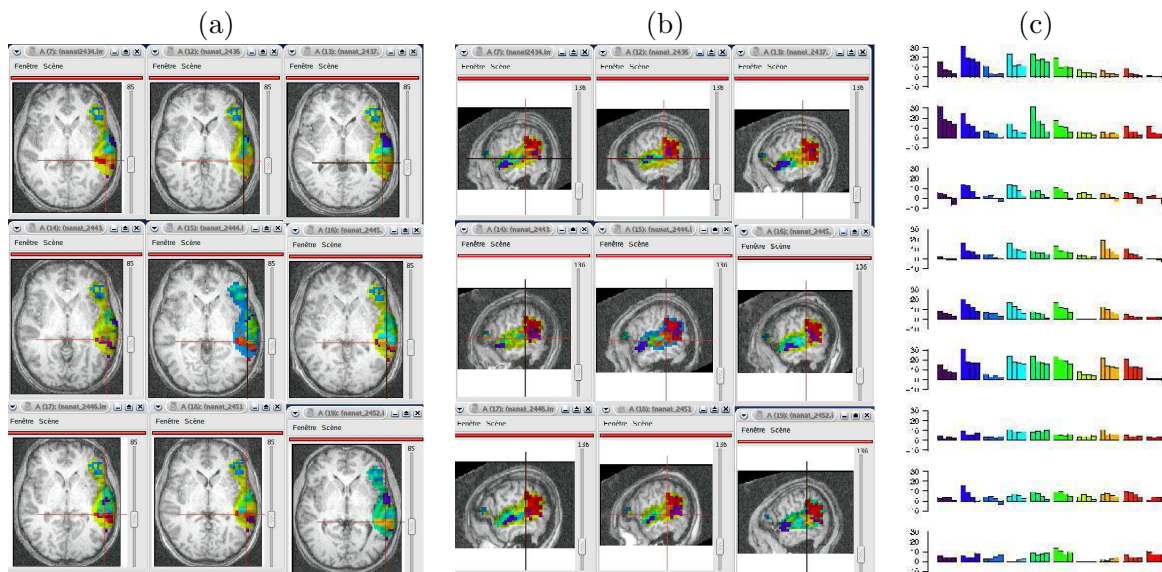
FIG. III.3 – Coupes axiales et sagitales d'une parcellisation multi-sujets (9 sujets, 3 sujets par ligne) consistant à identifier les régions qui manifestent une habituation de l'amplitude de la réponse hémodynamique c'est-à-dire un niveau de réponse plus élevé lors de la première présentation du stimulus par rapport à la seconde, la troisième ou la quatrième. Chaque couleur représente un label *i.e.*, une parcelle spécifique ; La figure (c) rapporte pour chaque *sujets* le vecteur d'effets BOLD estimés pour chacune des dix parcelles lors de l'ajustement d'un MLG standard.

– soit il apparaît plus opportun de considérer une hémodynamique régionale propre à chaque condition expérimentale. Le problème d'estimation devient plus mal posé, et ce d'autant que la variabilité spatiale est toujours prise en compte par une modulation de l'amplitude de la réponse d'un voxel à un autre. Cette modélisation n'a pas fait l'objet de publications mais uniquement d'un rapport de recherche personnel, *cf.* [rr.1].

III.2.3   Sous-thème 1b – Statistiques de détection et de comparaison

Une fois l'estimation conduite dans ce cadre non-paramétrique, il est intéressant de s'intéresser à deux questions d'ordre statistique à savoir tester l'influence de l'hémodynamique estimée sur la sensibilité de détection, et inférer l'état d'activation en fonction du profil de FRH recouvré.

La première question est habituellement abordée à l'aide de deux jeux de données d'un même sujet, le premier permettant l'estimation d'une FRH pour chaque condition expérimentale $(\widehat{h}_m)_{m=1:M}$ et le deuxième autorisant l'évaluation de la sensibilité de détection à l'aide d'un *modèle linéaire généralisé* dont les régresseurs, c'est-à-dire les colonnes de la matrice de dessin expérimental associée sont construits comme la convolution de l'hémodynamique locale estimée $(\widehat{h}_m)_{m=1:M}$ avec le vecteur binaire codant les temps d'arrivée des stimuli (*onsets*). La seconde question n'implique qu'un seul jeu de données.

La contribution [c.2] s'est intéressée à la première question, toujours dans le cadre de l'expérience de compréhension linguistique rapidement abordée au **sous-thème 1a**. Elle met clairement en évidence une sensibilité accrue lorsque le MLG est construit à partir d'une FRH estimée plutôt qu'à partir de la version canonique définie dans [GLOVER 1999]. En particulier, pour le contraste **son-silence** étudié en Fig. III.4(a), une activation bilatérale est logiquement

MLGs standards définis à partir de la FRH canonique : $f_m(t_n) \triangleq (\boldsymbol{x}^{(m)} \star \boldsymbol{h}_c)(t_n)$

$V_1 = (-60, -24, 4)$ mm $\qquad V_2 = (-68, -28, 8)$ mm $\qquad V_3 = (-64, -40, 16)$ mm

MLGs définis à partir des FRHs estimées : $f_m(t_n) \triangleq (\boldsymbol{x}^{(m)} \star \widehat{\boldsymbol{h}}_m)(t_n)$
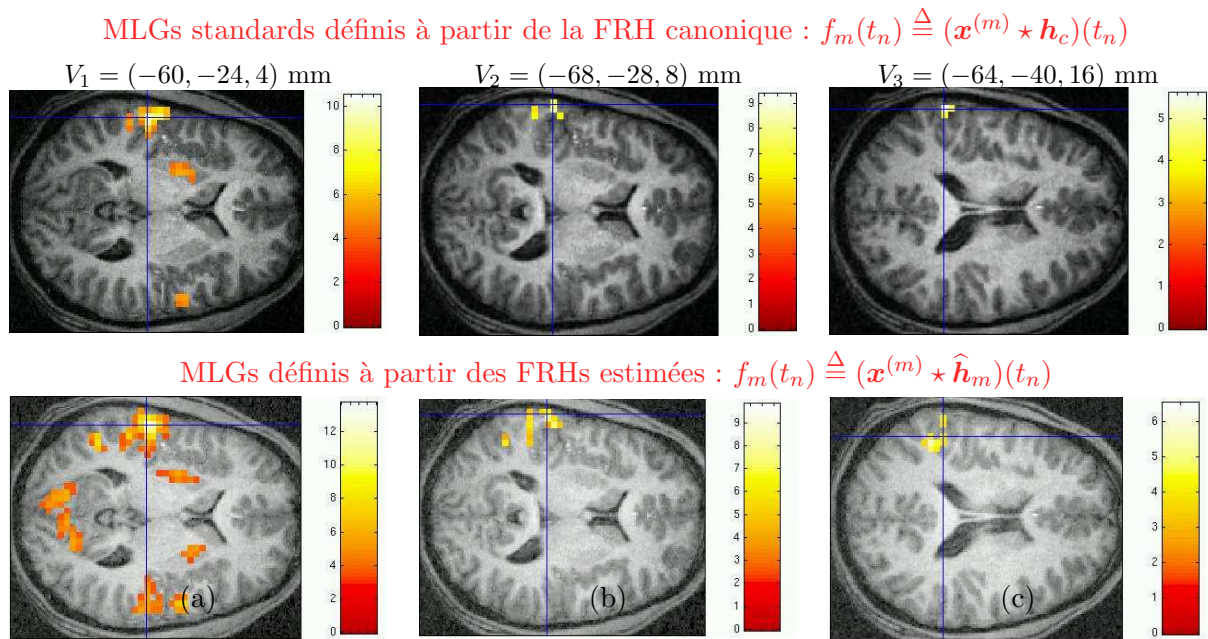
FIG. III.4 – Comparaison des cartes d'activations seuillées à $p = 0.001$ corresponsdant à des tests de Student pour les contrastes son-silence, (phonologique-contrôle)&(acoustique-contrôle), phonologique-contrôle. En haut, ces cartes sont issues de MLGs standards c'est-à-dire définis à partir de la FRH canonique $\boldsymbol{h}_c$. En bas, ces modèles s'appuient sur les FRHs estimées localement dans les voxels V_1-V_3.

restaurée avec ce dernier modèle. Précisons ici que les trois FRHs ont été estimées sur les trois premières sessions de données et que les cartes de détection issues des nouveaux modèles ont été calculées sur trois sessions d'acquisition supplémentaires. Les Figs. III.4(b)-(c) mettent aussi en évidence des activations plus étendues.

La seconde question peut être abordée comme un cas particulier du problème qui consiste à tester si une forme donnée $\boldsymbol{h}_0$ de filtre hémodynamique (par exemple issue d'une modélisation physiologique alternative ou d'une mesure empirique) représente significativement le profil hémodynamique estimé $\widehat{\boldsymbol{h}}$ en un voxel donné, voire dans une région. Dans un cadre fréquentiste, cela correspond à tester l'hypothèse nulle $(\boldsymbol{h} = \widehat{\boldsymbol{h}}_0)$ et à rejeter cette hypothèse si $\widehat{\boldsymbol{h}}$ est significativement différent de $\boldsymbol{h}_0$. En s'inspirant des travaux de [TANNER 1993], nous avons introduit dans [a.6] une *mesure de différentiation*, basée sur le calcul d'une distance de MAHALANOBIS puisque $\boldsymbol{h}$ suit une loi normale *a posteriori*, et le choix d'un seuil de significativité α_0. Ce type de tests peut être utile dans différentes situations pratiques, par exemple :

1. la détection d'activation qui consiste à poser $\boldsymbol{h}_0 = 0$.

2. la comparaison au sein du même voxel ou de la même région de deux FRHs $(\boldsymbol{h}_1, \boldsymbol{h}_2)$ modélisant la dynamique de deux conditions expérimentales, représentées par une distribution conjointe *a posteriori* gaussienne $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ où $\boldsymbol{m} = [\boldsymbol{\mu}_1, \boldsymbol{\mu}_2]$ et

$$\boldsymbol{\Sigma} = \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^{\mathrm{t}} & \Sigma_{22} \end{array} \right].$$

Dans ce cas, il faut tenir compte des termes non diagonaux dans le calcul de la distance de MAHALANOBIS.

3. la comparaison de deux FRHs estimées pour le même voxel ou la même région au cours de deux sessions d'acquisition successives. Dans ce cas, on peut supposer $\Sigma_{12} = \boldsymbol{0}$.

4. la comparaison de FRHs estimées $(\boldsymbol{h}_1, \boldsymbol{h}_2)$ dans deux voxels ou régions distinctes. Dans ce cas, il est nécessaire de contrôler les fluctuations vasculaires à l'aide d'une condition

de référence pour laquelle on estime aussi la FRH associée h_3. La comparaison $h_1 - h_2$ s'effectue alors au travers de $(h_1 - h_3) - (h_2 - h_3)$.

Le premier exemple revient donc à comparer la FRH estimée à une fonction plate reflétant que le stimulus n'a aucune influence sur le décours temporel d'intérêt, qui ne se compose alors que d'une ligne de base (tendance et bruit). Ainsi, dans ce cadre, il devient possible d'opérer une détection d'activation sur le signal fonctionnel, moyenné ou non spatialement, ayant servi à la caractérisation temporelle de la FRH.

Tous ces aspects ont été mis à la disposition des utilisateurs au sein de la boîte à outils HRF que j'ai développée depuis 2002[14].

III.2.4   Sous-thème 1d – Extension non-stationnaire (2003-06)

Les développements méthodologiques réalisés jusqu'ici font l'hypothèse que la réponse hémodynamique obéit au principe de superposition et donc demeure stationnaire. En d'autres termes, les réponses à deux stimulations identiques mais décalées dans le temps coïncident. Si cette hypothèse paraît en première approximation pertinente, il existe néanmoins des situations dans lesquelles la variabilité inter-essais de la réponse BOLD a été constatée, soit parce que des mécanismes d'apprentissage ou d'habituation du sujet aux stimuli se mettent en place (on parle de facilitation du sujet à la tâche), soit en raison d'un changement de stratégie du sujet en cours d'expérience, soit enfin à cause d'une certaine fatigue ou lassitude du sujet due à l'examen. À travers le stage de DESS d'Alexis DAVID en 2002, j'ai proposé de relâcher cette hypothèse en modélisant la variabilité inter-essais de l'*amplitude* de la réponse hémodynamique, la forme restant constante à travers les essais.

Ce travail a été repris ensuite par Sophie DONNET et développé davantage au cours de sa thèse[15]. La première contribution liée à ces travaux [c.10] a permis de mettre en évidence une telle variabilité notamment dans le cortex moteur pour la condition clic-droit (voir Fig. III.5). Le signal prédit par le modèle non-stationnaire permet de mieux s'ajuster aux données que le modèle stationnaire comme l'illustre la Fig. III.6. Les algorithmes mis en œuvre étaient du type SAEM (*Stochastic Approximation of the EM algorithm*, cf. [DELYON et coll. 1999]). Ces travaux ont ensuite donné lieu à une publication dans *NeuroImage*, cf. [DONNET et coll. 2006], à laquelle je n'ai pas été associé par volonté délibérée de certains auteurs : un remerciement chaleureux m'est adressé en fin d'article !

De façon colatérale, j'ai participé sans en être l'élément moteur, au développement de techniques de sélection de modèles pour élucider dans quelles circonstances et chez quels sujets ces modèles non-stationnaires sont pertinents pour expliquer la variabilité observée dans les données d'IRMf. Techniquement, cela consiste à comparer des vraisemblances de modèles différents (stationnaire ou non) et à mettre en œuvre des tests statistiques pour évaluer si les différences observées entre ces vraisemblances sont statistiquement significatives ou non. Le lecteur intéressé pourra consulter [DONNET et coll. 2006].

Cette expérience m'a permis par la suite d'amorcer un ensemble de réflexions sur un sujet connexe, l'inférence de modèles hémodynamiques non-linéaires et/ou non-stationnaires, qui sera évoqué dans le **Thème 2**. J'ai d'abord développé ces travaux tout seul avant de diffuser les idées qui les sous-tendent au travers du stage d'Ekaterina FALINA et du post-doctorat de Laurent RISSER.

---

[14]voir http://www.madic.org/download/HRFTBx/htmldir/HRFTBx_main.html et le tutoriel en http://www.madic.org/download/HRFTBx/tutorial/.

[15]thèse soutenue en probabilités et statistiques à l'université de Paris-Sud, centre d'Orsay, le 30/11/2006 sous la direction de Marc LAVIELLE.
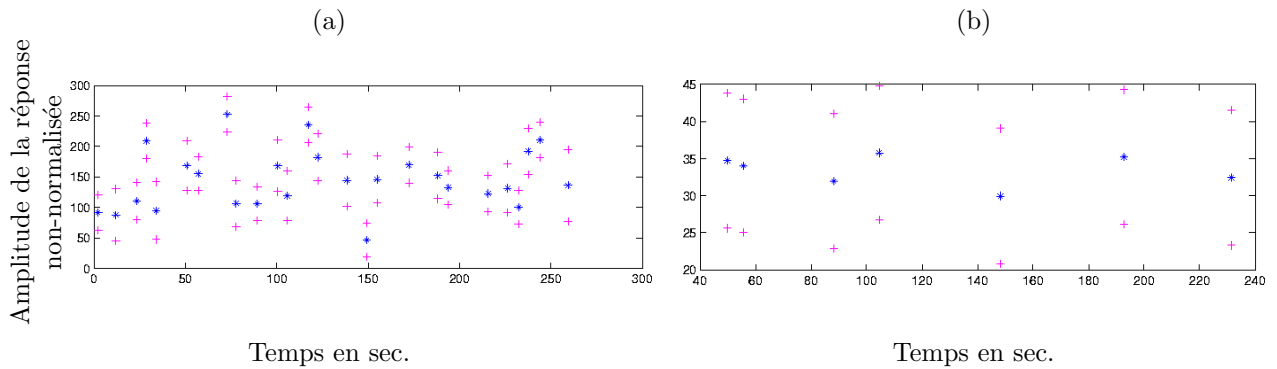
FIG. III.5 – Amplitude de la réponse hémodynamique estimée essai par essai dans les conditions (a) clic-droit et (b) visuelle à partir des signaux IRMf moyennés dans une région motrice.
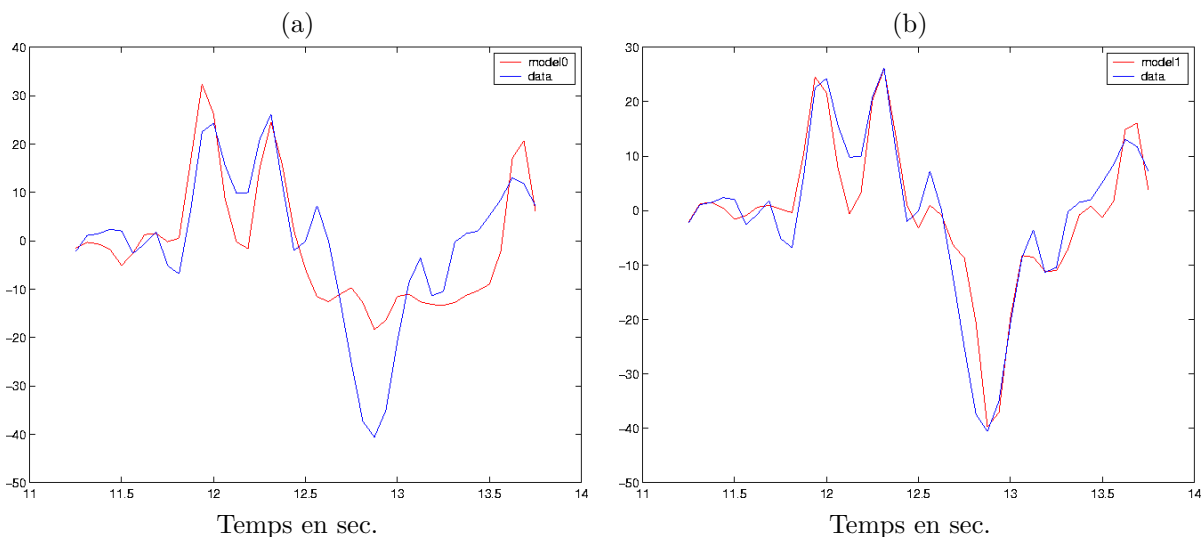
FIG. III.6 – Signal IRMf mesuré en bleu et signal ajusté en rouge c'est-à-dire prédit par les modèles (a) stationnaire et (b) non-stationnaire, ce dernier s'ajustant mieux aux données.

III.2.5 Sous-thème 1e – Applications en neuropsychologie cognitive et expérimentale

Ces développements divers ont donné lieu à l'écriture d'une boîte à outils, baptisée HRF[16], pour le logiciel SPM (versions SPM99 et SPM2). Récemment, elle a même été portée sous SPM5 dans le cadre de ma collaboration avec l'équipe de Anna Maria BIANCHI et Maria Gabriella TANNA du département de **Génie Biomédical** de l'institut Polytechnique de Milan [TANA et coll. 2007]. Par ailleurs, cette boîte à outils a fait l'objet de nombreuses présentations, tant au travers de congrès (poster [p.1, p.2], tutoriel [c.7]) que de conférences invitées [s.4] ou de formations destinées à un large public telles que les journées de formation inter-régionales en neuroimagerie organisées depuis 2003 par l'INSERM. Sur le plan des applications, j'ai pu faire fructifier plusieurs collaborations :

1. dans le cadre de la thèse de Charlotte JACQUEMOT[17], nous avons pu mettre en évidence

---

[16]voir http://www.madic.org/download/HRFTBx/htmldir/HRFTBx_main.html
[17]menée sous la direction d'Emmanuel DUPOUX en collaboration avec Christophe PALLIER

une dissociation de la dynamique des aires temporales supérieures impliquées dans la compréhension linguistique [a.3, Jacquemot, C .and Pallier et coll. 2003] tant chez des européens que chez des asiatiques (contrastes acoustique *vs* phonologique).

2. Suite aux journées de formation inter-régionales en neuroimagerie en 2004, j'ai également collaboré avec Alice de Fréminville de l'équipe inserm U342 du CHU de Saint Étienne dirigée par le Dr. Roland Peyron, qui s'intéresse à la perception de la douleur. Nous avons notamment pu mettre en évidence une plus grande précocité de la réponse hémodynamique insulaire par rapport à la réponse cingulaire.

3. dans le cadre de la thèse d'Anne Botzung, effectuée sous la direction de Lilianne Manning (Univ. Louis Pasteur, Strasbourg et Centre d'Études de Physiologie Appliquée du cnrs), nous avons étudié la dynamique cérébrale des processus de mémoire autobiographique, tant chez le sujet sain que chez le sujet atteint de sclérose en plaque [p.8, ?]. Nous avons pu mettre en évidence, comme le prédisait le modèle de Conway et Pledell-Pierce, l'implication plus précoce du cortex préfrontal dorso-latéral gauche dans le processus de récupération des souvenirs autobiographiques, en comparaison avec des structures plus postérieures. Nous avons aussi exhibé un flux temporel antéro-postérieur des activations sous-tendant les différentes phases de la récupération de souvenirs.

4. Dans le cadre de la thèse de Nicolas Wotawa, nous avons procédé à l'estimation de la dynamique de la réponse dans chaque aire visuelle primaire afin d'évaluer la sélectivité à la direction du mouvement (trois directions privilégiées dans la stimulation : 0, 45 et 180 ; pour les détails, voir [Wotawa 2006, chap. VI]). Grâce à cette caractérisation, Nicolas Wotawa a pu déduire la proportion respective de deux sous-populations de neurones sensibles à cet attribut visuel. Pour ce faire, il a évalué un indice d'adaptation directement à partir des réponses hémodynamiques estimées. Il est alors apparu que le complexe hMT+, directement suivi par l'aire V3A, étaient les plus sélectifs à la direction[18]. Une quantification de la proportion de neurones large bande et bande étroite a pu être réalisée conduisant notamment à des taux de 35% et 65% respectivement dans hMT+ et de 45% et 55% dans V3A. Il a aussi été mis en évidence que la région V2 n'était pas sélective à la direction du mouvement.

5. Dans un travail avec Ghislaine Dehaene-Lambertz et Christophe Pallier, nous avons pu mettre en évidence au sein du sillon temporal supérieur un gradient antéro-postérieur d'habituation de l'amplitude de la réponse hémodynamique à la répétition de phrases [p.5, a.4], tandis qu'aucune habituation n'a été constatée dans d'autres régions primaires (*e.g.,* le gyrus de Heschl). De façon plus ténue, nous avons aussi observé une accélération de la réponse dans les zones qui s'habituent (aire de Broca, sillon temporal supérieur dans sa partie postérieure), *i.e.,* une habituation du délai.

---

[18]Pour la définition de cet indice, voir [Wotawa 2006, p. 177]. Les valeurs de cet indice étaient les plus élevées pour hMT+ et V3A.

III.3   Thème 2 – Détection-estimation conjointe (depuis 2004)

Partant du constat que d'une part, les méthodes d'estimation de l'hémodynamique locale ou régionale ne s'appliquent en général qu'à des territoires cérébraux impliqués par le paradigme expérimental, et que d'autre part, les techniques classiques de détection d'activations supposent une forme *canonique* de fonction de réponse hémodynamique, il m'a semblé fondamental de développer des approches combinées, dites de *détection-estimation conjointe* visant à répondre aux deux questions simultanément. En d'autres termes, il s'agit d'estimer la dynamique cérébrale tout en localisant les foyers d'activation. Pour tenir compte de la variabilité spatiale de l'hémodynamique, il est nécessaire d'opérer région par région. Le « pavage » du cerveau en régions fonctionnellement homogènes nécessite le recours à des techniques de parcellisation anatomo-fonctionnelles telles que [a.9], dont un exemple est représenté à la Fig. III.7. Dans la suite, nous discutons de différents aspects de modélisation *intra-parcelle*, l'ensemble des parcelles étant fixé une fois pour toutes. Nous évoquerons au chapitre IV une piste visant à s'affranchir ou à remettre en cause une configuration donnée de parcelles.
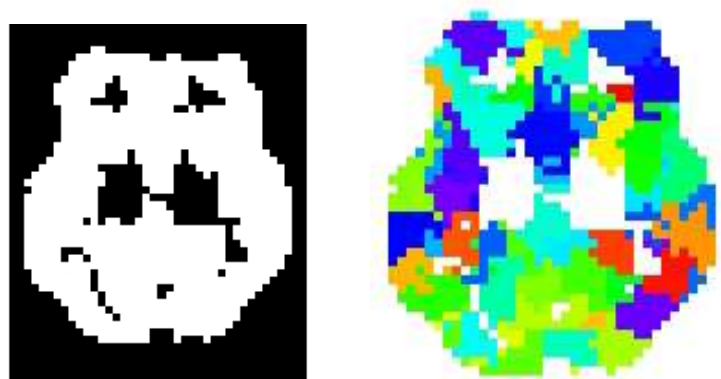
FIG. III.7 – À gauche, version dilatée d'une coupe du masque anatomique du cerveau sur la base duquel est calculée la parcellisation. À droite, le résultat de la parcellisation pour la même coupe en $z = -4mm$. Chaque couleur décrit une parcelle différente.

Ce travail, précurseur et original au sein de la communauté de la neuro-imagerie fonctionnelle, a constitué le cœur de la thèse de Salima MAKNI. Il a par ailleurs occupé Thomas VINCENT durant son stage de Master recherches en 2006 et sa première année de thèse. Il a fédéré en 2007 le travail de deux étudiantes de master, Ekaterina FALINA et Anne-Laure FOUQUE, et constitue un cadre théorique intéressant pour de plus amples développements dans le futur notamment en sélection de modèles. D'inspiration bayésienne, ces travaux représentent un volet important de mon activité de recherche et contribuent significativement à ma reconnaissance internationale (*cf.* [a.8, c.11, c.12, c.15, c.18, **?**, **?**]). Du point de vue algorithmique, ces travaux s'insèrent dans un cadre *totalement bayésien*, c'est-à-dire qu'ils exploitent des méthodes de MONTE-CARLO par Chaînes de MARKOV pour simuler des réalisations de la distribution *a posteriori* jointe.

Les sous-thèmes suivants résument à la fois les points clés qui font l'originalité de ces travaux et les problèmes théoriques résolus. Le chapitre IV est consacré à une présentation aboutie de ces travaux. Dans tous les cas, le modélisation du signal BOLD retenue illustrée par la Fig. III.8 est non-paramétrique.
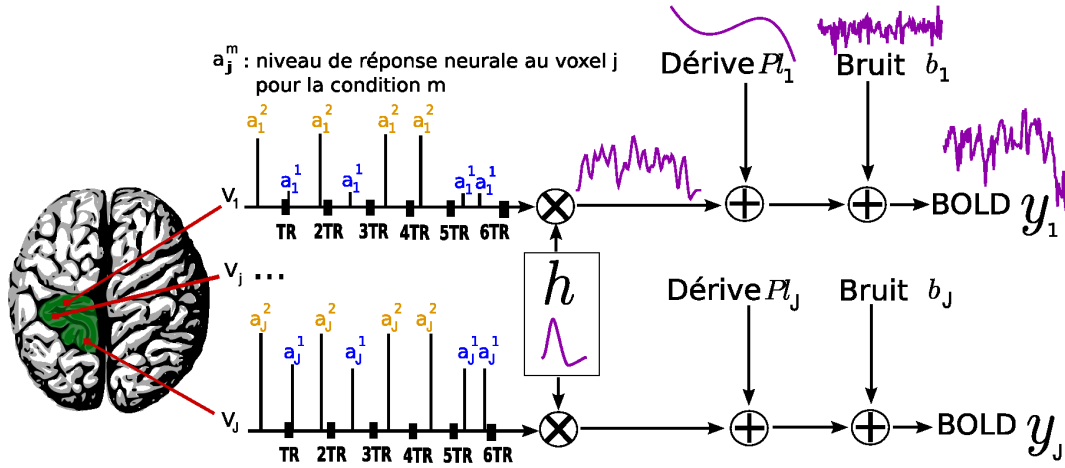
FIG. III.8 – Modèle d'observation non-paramétrique retenu dans chaque parcelle $\mathcal{P}$ (en vert ici).

De façon équivalente, le modèle s'écrit sous forme vectorielle dans chaque voxel V_j :

$$\boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}^m \boldsymbol{h} + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j, \tag{III.1}$$

où $\boldsymbol{y}_j = (y_{j,t_n})_{n=1:N}$ désigne la série temporelle IRMf mesurée en V_j aux instants $(t_n)_{n=1:N}$ (N étant le nombre de scans) avec $t_n = nTR$ et TR définit le temps de répétition. La matrice $\boldsymbol{X}^m$ de taille $N \times (D+1)$ est binaire et encode les temps d'arrivée des stimuli pour la $m^{\text{ième}}$ condition expérimentale. Le vecteur $\boldsymbol{h} \in \mathbb{R}^{D+1}$ représente la FRH inconnue dans la parcelle $\mathcal{P}$, qui est considérée ici de même forme pour l'ensemble des stimuli. Le scalaire a_j^m définit le *niveau de réponse neurale* (NRN) dans le voxel V_j pour la condition m (M spécifie le nombre de conditions expérimentales présentes dans le paradigme). Alors, le profil d'activation associé au $m^{\text{ième}}$ type de stimulus en V_j s'obtient comme le produit $\boldsymbol{h} \times a_j^m$. La matrice $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_Q]$ définit une base orthogonale de taille $N \times Q$ à basses fréquences. Les fonctions qui la définissent $\boldsymbol{p}_q = (p_q(t_n))_{n=1:N}$ restent au choix de l'utilisateur et demeurent un degré de liberté du modèle. Le vecteur $\boldsymbol{\ell}_j$ définit la représentation canonique dans cette base de la dérive basse fréquence en V_j. Enfin, $\boldsymbol{b}_j \in \mathbb{R}^N$ modélise le bruit et les erreurs de modèle en V_j. Dans [WOOLRICH et coll. 2001, WORSLEY et coll. 2002], un modèle autorégressif (AR) a été introduit pour prendre en compte l'autocorrélation temporelle des données d'IRMf dans la phase de détection des activations. Lorsque cette corrélation est correctement estimée, le nombre de faux positifs décroit produisant ainsi des résultats plus conservatifs.

III.3.1   Sous-thème 2a – Modèles de mélanges indépendants (2004-06)

À partir du travail mené au **Sous-thème 1b** (recherche de l'homogénéité fonctionnelle), le choix du modèle *a priori* portant sur les amplitudes de la réponse hémodynamique a été « optimisé ». Plus précisément, au lieu de supposer ces amplitudes spatialement distribuées selon une loi normale de moyenne et variance inconnue, nous avons considéré un modèle de mélange à deux classes, l'une pour les voxels inactivés (classe 0) et l'autre pour les voxels activés (classe 1) [a.8]. Ce modèle s'applique bien sûr pour chaque condition expérimentale prise séparément. Plus tard, il a d'ailleurs été étendu afin de prendre en compte des désactivations potentielles au sein d'une troisième classe [c.17]. Des travaux antérieurs [EVERITT ET BULLMORE 1999, VAEVER HARTVIG ET JENSEN 2000] ont déjà démontré le double intérêt de recourir à une

modélisation par mélange probabiliste. D'une part, elle permet de mieux prendre en compte des hétérogénéités locales tout en conservant une formulation paramétrique relativement simple. Ici, il s'agit de s'affranchir en partie de l'hypothèse d'homogénéité fonctionnelle, en autorisant un comportement différent de certains voxels au sein du mélange. D'autre part, cette modélisation permet d'opérer une *classification* des voxels en deux catégories et ainsi de détecter lesquels sont significativement activés c'est-à-dire induisent une activité évoquée par le paradigme expérimental.

Ainsi, une loi gaussienne *centrée* de variance inconnue modélise la distribution des amplitudes des voxels inactivés. Une deuxième distribution de paramètres statistiques inconnus modélise les amplitudes hémodynamiques des voxels activés. Une première contribution d'un mélange gaussien à deux classes [a.8] a montré qu'une telle modélisation pouvait s'avérer insuffisante dans des situations où le mélange dégénère. Nous ne parlons pas ici des cas extrêmes où l'une des deux classes est vide[19], mais plutôt d'une situation où les deux distributions gaussiennes deviennent peu discernables (voir Fig. III.9). Le contrôle des faux positifs, *i.e.,* de la *spécificité*, devient alors mauvais.
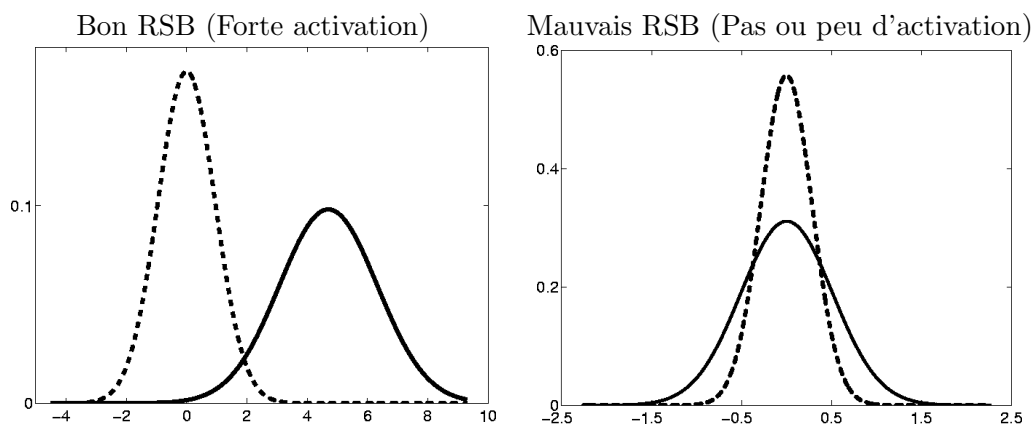
FIG. III.9 – Illustration des problèmes rencontrés avec un mélange gaussien à deux classes où les distributions gaussiennes $\mathcal{N}(0, \sigma_0^2)$ et $\mathcal{N}(\mu_1, \sigma_1^2)$ modélisent les voxels inactivés et activés respectivement. Ici aucune contrainte n'est fixée sur les variances et sur μ_1.

Pour résoudre ce problème, nous avons proposé d'introduire une contrainte de positivité sur le support de la distribution des activations. Dès lors, plusieurs choix de distribution s'ouvraient à nous. Pour des raisons techniques qui tiennent à la faisabilité de l'échantillonnage des hyper-paramètres, nous avons opté pour une loi gamma de paramètres d'échelle et de forme inconnus. L'algorithme qui en découle a été présenté dans [c.12]. Les gains obtenus en termes de compromis sensibilité-spécificité[20] sont significatifs comme illustrés à la Fig III.10, au prix d'un accroissement important du coût de calcul.

III.3.2   Sous-thème 2b – Modélisation du bruit (2005-06)

Dans un travail initial [c.3], il a été démontré que le choix du modèle de bruit, tenant compte à la fois d'imprécisions de mesure (inhomogénéités du champ magnétique, pertes de signal,

---

[19]Ce cas sera traité dans le chapitre IV.

[20]La sensibilité se mesure par le taux de *vrais positifs* c'est-à-dire le nombre de vrais positifs sur le nombre total d'événements positifs, qui correspond à la somme des faux négatifs et des vrais positifs, tandis que la spécificité correspond au taux de *vrais négatifs*, et s'obtient comme le nombre de vrais négatifs divisé par la somme des faux positifs et des vrais négatifs.
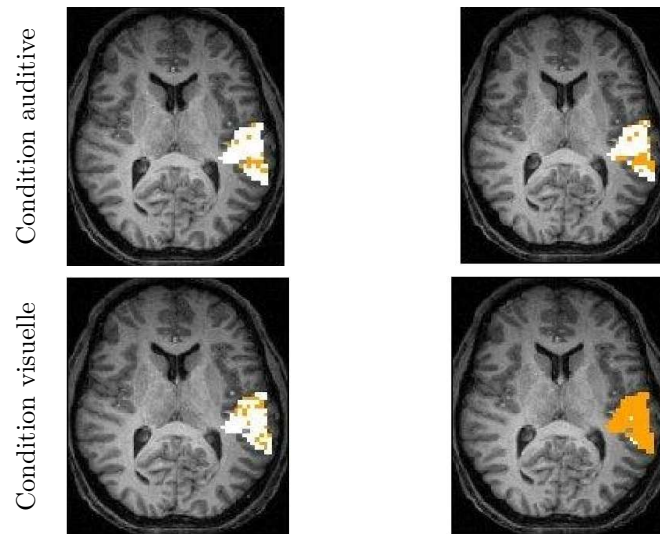
FIG. III.10 – Comparaison des résultats de détection entre deux modélisations par mélanges indépendants différents pour les conditions auditive et visuelle du protocole de localisation fonctionnelle. À gauche, mélange gaussien à deux classes, à droite mélange gamma-gaussien qui illustre notamment le gain en spécificité. Les voxels activés apparaissent en blanc, ceux inactivés en orange.

...), d'artefacts spécifiques au sujet (mouvement de la tête, battements cardiaques et respiration), mais aussi d'imperfections dues à la représentation de l'activité évoquée par un modèle linéaire, n'avait que peu de conséquences sur l'estimation de la forme de l'hémodynamique locale. À l'inverse, en *détection*, il est désormais bien connu que la sous-estimation de l'autocorrélation temporelle des signaux IRMf conduit à une élévation du taux de faux positifs dans les cartes d'activation [WOOLRICH et coll. 2001, WORSLEY et coll. 2002] et donc à une dégradation du compromis sensibilité-spécificité. Deux classes de solutions ont été envisagées dans la communauté pour tenir compte de cette autocorrélation dans un contexte de détection (voir par exemple [FRISTON et coll. 1995, FRISTON et coll. 2002]). La première, dite *pre-colouring*, tenant plus de la « recette de cuisine » que d'une véritable démarche scientifique, noie la structure de dépendance temporelle des signaux mesurés dans une forme d'autocorrélation spécifiée par l'utilisateur, donc connue, permettant ainsi de corriger le nombre de degrés de liberté de la statistique de décision et du test associé. La seconde approche, baptisée *pre-whitening*, consiste d'abord à estimer la structure d'autocorrélation en ajustant par exemple un modèle autorégressif [WOOLRICH et coll. 2001, WORSLEY et coll. 2002, PENNY et coll. 2003] aux résidus d'un MLG puis à « blanchir » les données afin de recouvrer une situation de décorrélation statistique.

Il semblait donc opportun d'étudier l'influence du modèle de bruit dans le cadre conjoint de détection-estimation. La contribution [c.11] a mis en évidence que l'utilisation d'un modèle auto-régressif du premier ordre (modèle dit AR(1)), permettait en effet de réduire le taux de faux positifs tout en n'ayant que peu d'influence sur la forme de l'hémodynamique estimée, confirmant ainsi les conclusions tirées séparément en détection et en estimation. L'introduction de ce modèle de bruit complexifie sensiblement la procédure d'échantillonnage de certains paramètres et hyperparamètres du modèle, notamment les paramètres AR. Une version étendue de ces travaux se trouve maintenant dans l'article [**?**] en cours de publication dans la revue *NeuroImage*. La Fig. III.11 illustre le type de résultats mis en avant.

Enfin, il est à noter que le compromis sensibilité-spécificité s'améliore, particulièrement à
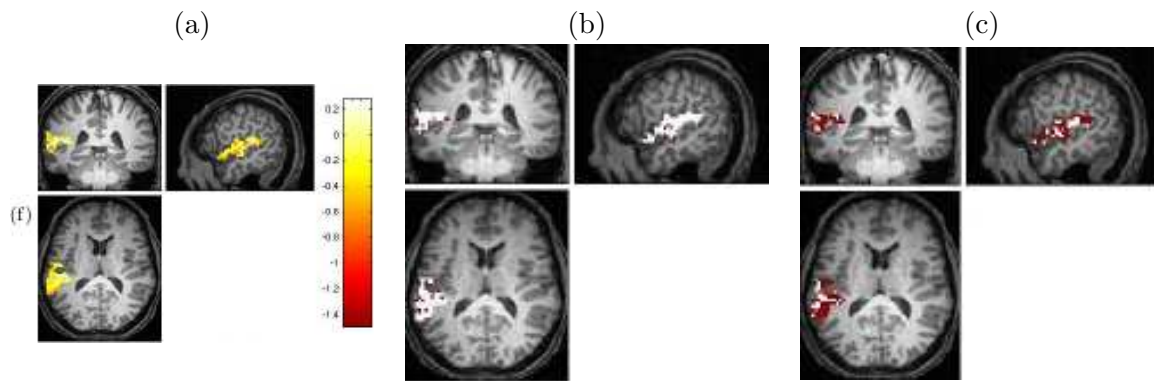
Fig. III.11 – Analyse des activations dans le lobe temporel (cortex auditif). (a) : différences des amplitudes de la réponse hémodynamique pour une condition visuelle dans un protocole de calibration fonctionnelle, pour laquelle on n'attend pas d'activations *a priori*. (b)-(c) : comparaison des résultats de détection pour cette même condition entre les modèles de bruit blanc (b) et auto-régressif du premier ordre (c). Le taux de faux positifs décroît avec la prise en compte de la corrélation temporelle. Les deux modèles sont adaptatifs dans l'espace au sens où ils estiment des paramètres statistiques de bruit dans chaque voxel. Les voxels en blanc sont déclarés activés, ceux en brun sont inactivés.

faible RSB, lorsque l'autocorrélation temporelle est prise en compte, et ce quel que soit le type de modèle de mélange exploité sur les amplitudes de la réponse hémodynamique. Toutefois, cette amélioration est plus significative lorsqu'il s'agit d'un mélange gaussien (voir [**?**]).

III.3.3 Sous-thème 2c – Modélisation de la corrélation spatiale (depuis 2006)

Les travaux présentés aux paragraphes précédents ont porté jusqu'à maintenant sur des mélanges *indépendants*, au sens où la classification d'un voxel n'influence pas celle de ses voisins. Il est pourtant bien connu que les activations cérébrales, même en dehors de tout lissage spatial des données, présentent une certaine extension spatiale. En toute rigueur, c'est au niveau du ruban cortical, siège des activations cérébrales, que cette corrélation spatiale se manifeste. Il serait donc plus pertinent de modéliser l'extension spatiale des activations cérébrales sur la surface corticale plutôt que dans le volume dont la géométrie est prescrite à l'acquisition. Toutefois, l'analyse sur la surface [Andrade 2002] nécessite d'avoir résolu certaines questions, réputées difficiles (extraction et maillage de la surface, projection des données, ...), sur lesquelles nous reviendrons au Thème 3. C'est pourquoi dans un premier temps, nous nous sommes concentrés sur l'introduction de modèles de mélanges spatiaux dans le volume cérébral, extension naturelle des mélanges indépendants, afin de favoriser les configurations où les activations apparaissent sous forme de *clusters* plutôt que sous forme de points isolés [Woolrich et coll. 2005, Ou et Golland 2005]. Cette thématique a occupé le stage de master de Thomas Vincent puis le début de sa thèse.

Nous avons privilégié l'encodage de cette corrélation spatiale sur les « états » des voxels adjacents au sens d'un système de voisinage 3D[21], plutôt que sur les amplitudes des activations. Ainsi, nous avons introduit un champ de Markov *caché* sur les états inobservés des voxels constituant le mélange au sein de chaque parcelle. Dans le cas d'un mélange à deux classes, le modèle intra-parcelle introduit correspond à un champ binaire de Ising.

---

[21]On peut considérer un système de voisinage induit par la 6 ou la 26-connexité.

Jusqu'ici, nous avons travaillé sur des mélanges spatiaux gaussiens pour des raisons de coût de calcul, mais les premiers résultats obtenus en simulation [c.15], puis sur données réelles [c.18, d.7], ont démontré une nette amélioration du compromis sensibilité-spécificité par rapport aux mélanges indépendants. En particulier, à un niveau de sensibilité donné, le taux de faux positifs obtenu est bien plus faible.

Dans ces premiers travaux, le paramètre fixant la régularité spatiale du champ a été fixée empiriquement en fonction de la résolution spatiale des volumes fonctionnels acquis, des habitudes des neuropsychologues en termes de lissage spatial des données, et de tests sur des données réelles. Néanmoins, il serait intéressant de régler automatiquement ce paramètre de régularité intrinsèque à partir des données, et rendre la régularisation spatiale *adaptative* comme dans [Woolrich et coll. 2005, Woolrich et Behrens 2006]. Bien que coûteuse, cette stratégie reste néanmoins envisageable, d'autant que l'architecture informatique que nous avons déployée permet maintenant de répartir les calculs sur une ferme de calculateurs en réseau. Ces aspects sont discutés en guise de perspective au chapitre IV.

III.3.4 Sous-thème 2d – Modélisation de déactivations (depuis 2006)

L'idée d'introduire davantage de souplesse dans le modèle *a priori* de mélange pour modéliser des déactivations remonte déjà à plusieurs années [Woolrich et coll. 2005]. Nous l'avons simplement adaptée au contexte de la détection-estimation conjointe. Dans notre cas, l'ajout d'une contrainte de négativité sur les niveaux de réponse neurale liées aux déactivations potentielles s'est fait naturellement au travers d'une loi anti-gamma, comme illustré à la Fig. III.12. Dès
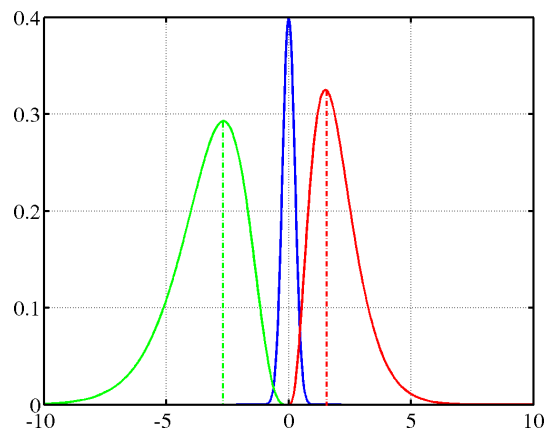
Fig. III.12 – Mélange bi-gamma gaussien à trois classes pour tenir compte de la positivité des activations et de la négativité des déactivations. Une distribution gaussienne centrée modélise toujours les inactivations.

lors, pour ce qui concerne les mélanges indépendants, l'extension est relativement directe et à été développée *in extenso* dans [**?**]. Le modèle devenant plus complexe à identifier en raison du plus grand nombre de paramètres, il nous a été difficile jusqu'à maintenant de démontrer de façon très claire l'apport de ce modèle sur des données réelles. En revanche, la situation pourrait être plus favorable dans le cas de mélanges spatiaux, où cette fois l'extension à trois classes intégrant les déactivations potentielles induit l'usage d'un champ de Potts à « trois états ». En particulier, il existe deux situations classiques générant des déactivations :

1. la première a déjà été mentionnée au **sous-thème 1b** : elle consiste par exemple au sein d'un paradigme événementiel auditif à présenter isolément des essais *silence*, au cours

desquels on arrête les gradients. En réponse à cet effet, le niveau de signal décroît par rapport à la ligne de base où le bruit d'acquisition se superpose aux stimuli auditifs ;

2. la seconde concerne l'épilepsie, pathologie au cours de laquelle durant des phases inter-critiques des pointes ondes détectables sur l'eeg génèrent des déactivations mesurables en irmf par une inversion de la réponse hémodynamique.

Pour cette seconde situation, la collaboration avec l'équipe de Milan, et peut être bientôt celle de Marseille (Christian Benar, inserm U751, Univ. de la Méditerranée), est fructueuse car elle nous permet d'accéder à des données présentant de façon certaine de telles déactivations focales. Ces données vont nous aider à valider notre modèle à trois classes intégrant la corrélation spatiale. Il s'agit déjà d'une fusion multimodale eeg/irmf *asymétrique* puisque les données eeg sont exploitées ici pour définir les *onsets* ou les temps d'arrivées des événements électriques pertinents générant une réponse hémodynamique perceptible.

III.3.5 Sous-thème 2e – Modèles non-linéaires et/ou non-stationnaires (depuis 2005)

L'objectif est ici d'appréhender des phénomènes hémodynamiques et neuronaux plus complexes tels que l'habituation ou l'adaptation de la réponse, la variabilité inter-essais, dans un contexte de détection-estimation. L'investigation de ces paramètres est particulièrement intéressante parce qu'elle devrait nous renseigner davantage sur des corrélats de l'activité neuronale tels que des paramètres d'amplitude ou de délai. Elle pourrait aussi à terme refléter chez le sujet épileptique le désordre introduit par des pointes-ondes au sein de l'activité neuronale des foyers épileptogènes.

Le choix opéré ici consiste à quitter le cadre bilinéaire sans toutefois recourir à des modèles physiologiques bien plus complexes à identifier [Buxton et coll. 2004, Riera et coll. 2004, Deneux et Faugeras 2006]. J'ai proposé une démarche intermédiaire s'appuyant sur la généralisation des modèles développés jusqu'ici, pour capturer des *non-linéarités* ou des *non-stationnarités* de la réponse hémodynamique. En particulier, je me suis focalisé sur deux sources potentielles de la *variabilité inter-essais* :

1. modélisation de l'*habituation* inter-essais ;
2. modélisation de la fluctuation de l'activité spontanée.

Le premier thème a pour objet de relâcher la contrainte de stationnarité de l'amplitude hémodynamique évoquée par le stimulus tout en introduisant une dépendance temporelle, *i.e.,* inter-essais, entre les différentes réponses par le biais d'un paramètre supplémentaire, en l'occurrence une *vitesse moyenne d'habituation*, normalisée entre 0 et 1. Ce type d'approche est particulièrement prometteur parce qu'il permet de modéliser des comportements de la réponse plus complexes au sein de modèles parcimonieux, ne nécessitant en réalité que l'ajout d'un paramètre par voxel et par condition expérimentale. Les amplitudes des réponses aux différents essais se déduisent en effet de l'amplitude de la réponse au premier essai et de la vitesse moyenne d'habituation. Il est bien évident que ce type de modèle englobe la configuration stationnaire dès lors que la vitesse d'habituation est nulle. De plus, ce dernier paramètre n'est pertinent à modéliser que pour des voxels présentant une activation. En conséquence, il semble opportun d'introduire un modèle *a priori* composite de type Bernoulli-uniforme sur ces paramètres de vitesse pour imposer la nullité de l'habituation dans les voxels inactivés. Par ailleurs, le modèle *a priori* global englobant les paramètres d'amplitude et de vitesse d'habituation est *séparable* vis-à-vis des conditions expérimentales et préserve l'indépendance conditionnelle de ces deux types de paramètres une fois la classification du voxel effectuée au sein du mélange.

Une première contribution dans cette veine a fait l'objet d'une communication [p.9]. Des améliorations concernant la prise en compte d'un modèle spatial ont occupé Ekaterina Falina

au cours de son stage de Master d'imagerie médicale de l'UPS mais ces travaux sont encore en cours de validation.

Les développements concernant le second thème sont synthétisés au sein d'un rapport de recherches [rr.3]. Ils autorisent l'inférence de modèles plus souples donc plus complexes, de dimension variable, où les fluctuations de l'activité spontanée conduisent à définir une amplitude de la réponse par essai, *indépendante*, mais pas nécessairement *identiquement distribuée*. Les mécanismes d'inférence adaptés sont du type méthodes MCMC à sauts réversibles [GREEN 1995, RICHARDSON ET GREEN 1997] et nous permettent d'évoluer au sein de familles de modèles emboîtés en vue d'une étape de sélection de modèles. Ces approches relativement ardues n'ont pas fait l'objet pour le moment d'une recherche aboutie et concrétisée en raison des compétences pointues requises. Un sujet de thèse autour de ces questions a été rédigé tout récemment dans le cadre du dossier de bourse AMN déposé par Marion SOUMOY pour la campagne 2008.

III.3.6 Sous-thème 2f – Indétermination liée à la bilinéarité (depuis 2006)

Ce travail collaboratif a été mené avec Thomas VEIT, post-doctorant à l'IRCCyN durant l'année 2006, et Jérôme IDIER, son encadrant. Il s'est concrétisé par deux communications à la conférence GRETSI 2007, une traitant des aspects théoriques [VEIT ET IDIER 2007], l'autre des aspects spécifiques rencontrés en IRMf [**?**].

Ce travail s'intéresse aux algorithmes MCMC impliqués dans la résolution de problèmes inverses bilinéaires dans un cadre bayésien. Les problèmes inverses bilinéaires incluent des applications telles que la séparation de sources [SNOUSSI ET IDIER 2006] ou la déconvolution aveugle [CHENG et coll. 1996]. Les quantités inconnues d'un problème inverse bilinéaire sont reliées aux observations par une fonction bilinéaire. Autrement dit, la fonction reliant les observations aux deux inconnues est linéaire par rapport à chacune des deux variables lorsque l'autre est constante. La résolution de ce type de problèmes dans un cadre bayésien lève l'ambiguïté d'échelle mais les échantillons issus de l'algorithme MCMC sont fortement corrélés. L'introduction d'une étape d'échantillonnage d'un paramètre d'échelle scalaire, noté s, améliore radicalement le comportement de l'algorithme MCMC. Par rapport à d'autres solutions existantes (*e.g.,* renormalisation déterministe du filtre à chaque itération, par exemple unitaire), cette approche est mathématiquement rigoureuse sans augmentation significative du coût de calcul. Elle correspond en quelque sorte à une *normalisation stochastique* des quantités mises en jeu.

Dans le contexte de la détection-estimation conjointe de l'activité cérébrale en IRMf, la modélisation retenue, décrite à la Fig. III.8, conduit à résoudre un problème de déconvolution semi-aveugle bilinéaire. Cette propriété introduit une ambiguïté d'échelle qu'on lève habituellement de façon non rigoureuse en imposant une contrainte sur la norme du filtre (*e.g.,* $\|\boldsymbol{h}\| = 1$). Si cette stratégie demeure valable en optimisation déterministe, où les problèmes de minimisation sous contrainte sont résolubles exactement, elle peut s'avérer désastreuse dans un schéma MCMC puisqu'elle peut déformer la loi *a posteriori* et modifier le point de convergence, comme illustré sur la Fig. III.13. Dans [**?**], j'ai donc proposé de lever cette ambiguïté en appliquant les résultats du travail présenté dans [VEIT ET IDIER 2007] pour garantir et accélérer la convergence de l'algorithme d'échantillonnage stochastique. Les courbes apparaissant en vert sur la Fig. III.13 valident cette approche puisque la forme de la FRH estimée converge vers celle de référence. Toutefois, il est à noter que ces résultats ne sont valables que pour un modèle de mélange *a priori* bi-gaussien sur les amplitudes neurales. L'extension au cas de mélanges *a priori* inhomogènes est rendue difficile par le fait que la loi du paramètre d'échelle n'a alors plus une forme connue et demeure donc plus coûteuse à simuler. Dans ce cas, on pourrait remplacer le rééchantillonnage exact de l'échelle par une étape de METROPOLIS à marche aléatoire.
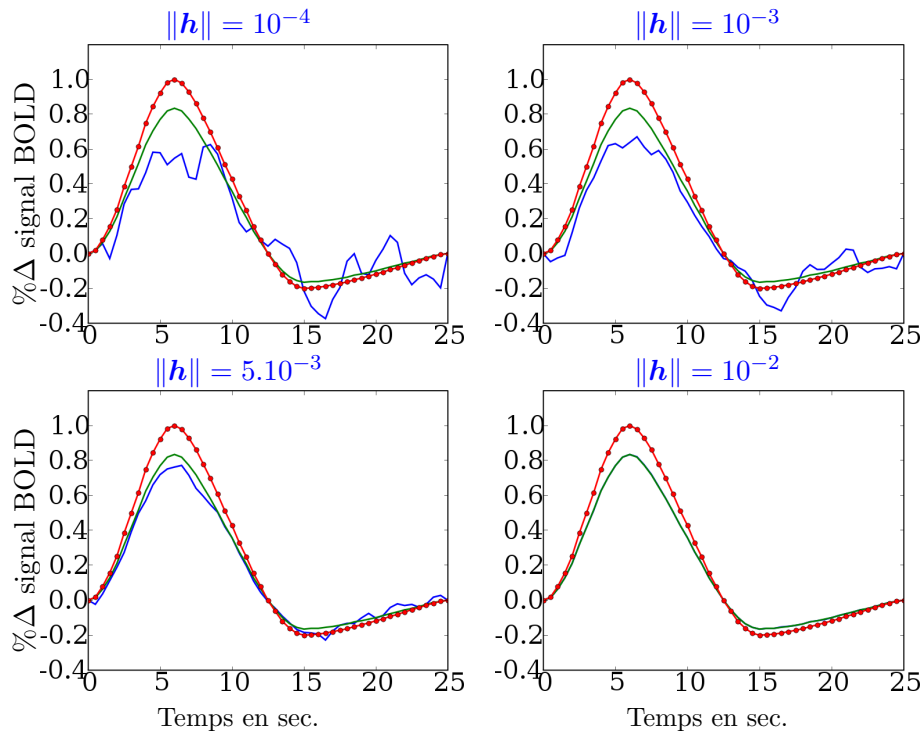
FIG. III.13 – $\boldsymbol{h}$ fonction de réponse hémodynamique de référence utilisée pour générer les données ; $\widehat{\boldsymbol{h}}$ estimées par renormalisation déterministe en fixant $\|h\|$; $\widehat{\boldsymbol{h}}$ estimées par rééchantillonnage de l'échelle. Il a été délibérément choisi de remettre à l'échelle $\widehat{\boldsymbol{h}}$ de façon discriminante vis-à-vis de $\boldsymbol{h}$ pour mieux observer les différences potentielles de forme.

III.3.7   Sous-thème 2g – Sélection et comparaison de modèles (depuis 2007)

Nous avons vu dans les sous-thèmes précédents qu'un certain nombre de choix s'offraient à nous pour décrire la grande variété de paysages d'activation cérébrale, due en partie aux variabilités individuelles, développementales, à des pathologies ... Il nous a paru opportun de concevoir des méthodes permettant de clarifier le message à délivrer au neuropsychologue et de le synthétiser. Pour ce faire, il est d'usage de recourir à des critères informationnels mesurant le compromis entre la capacité d'un modèle à s'ajuster aux données et sa complexité, mesurable notamment en termes de nombre de paramètres à identifier. J'ai récemment développé une application des travaux de [RAFTERY et coll. 2007] au domaine de la neuro-imagerie, et particulièrement au contexte de la détection-estimation conjointe. Toutefois, ces travaux sont suffisamment génériques pour s'appliquer à d'autres contextes. Les éléments clés sont synthétisés dans les paragraphes qui suivent sans démonstration.

Dans la communauté bayésienne, l'outil de référence s'appelle la vraisemblance *a posteriori* intégrée d'un modèle $\mathcal{M}$ et se note $p(\mathcal{M} \mid \mathbb{y})$ où $\mathbb{y}$ désigne ici l'ensemble des données disponibles, par exemple la suite des séries temporelles en IRMf. Dès lors, la comparaison de modèles peut se faire au travers d'une quantité remarquable, le *facteur de Bayes* [KASS ET RAFTERY 1995, CHIPMAN et coll. 2001] :

$$BF_{mn} \triangleq \frac{p(\mathcal{M}_m \mid \mathbb{y})}{p(\mathcal{M}_n \mid \mathbb{y})} = \frac{p(\mathbb{y} \mid \mathcal{M}_m)}{p(\mathbb{y} \mid \mathcal{M}_n)} \times \frac{p(\mathcal{M}_m)}{p(\mathcal{M}_n)},$$

où (m, n) référencent deux modèles disjoints. Dans le cas de modèles *a priori* équivalents, le second rapport dans le membre de droite disparaît et le facteur de Bayes se réduit au rapport

des *évidences* des modèles $\mathcal{M}_m$ et $\mathcal{M}_n$ définies par :

$$p(\mathrm{y}\,|\,\mathcal{M}_\star) = \int p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star, \mathcal{M}_\star)\,p(\boldsymbol{\theta}_\star\,|\,\mathcal{M}_\star)\,\mathrm{d}\boldsymbol{\theta}_\star, \quad \text{avec } \star = m, n.$$

Il s'agit donc d'un problème d'inférence hiérarchique, où le premier niveau concerne l'estimation des paramètres $\boldsymbol{\theta}_\star$ conditionnellement à la connaissance du modèle $\mathcal{M}_\star$, tandis que le second vise à répondre à la question du choix du meilleur modèle disponible au sein d'une famille disponible. Dans le cadre d'un schéma MCMC, il semble opportun de chercher à approcher l'évidence d'un modèle donné à partir des échantillons simulés de la loi *a posteriori* $p(\boldsymbol{\theta}_\star\,|\,\mathrm{y}, \mathcal{M}_\star)$. Il nous a donc semblé intéressant d'investir la piste de recherche qui consiste à déterminer un bon estimateur de l'évidence d'un modèle, en un certain sens à définir, à partir de ces échantillons issus de la loi *a posteriori*. De prime abord, ceci ne semble pas évident dans la mesure où :

$$p(\mathrm{y}\,|\,\mathcal{M}_\star) = \int p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star\,|\,\mathcal{M}_\star)\,p(\boldsymbol{\theta}_\star\,|\,\mathcal{M}_\star)\,\mathrm{d}\boldsymbol{\theta}_\star = \mathrm{E}_{\boldsymbol{\theta}_\star}\big[p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star, \mathcal{M}_\star)\big] \tag{III.2}$$

Autrement dit, un calcul direct montre clairement la difficulté : l'évidence du modèle peut être approchée par une moyenne empirique de L vraisemblances $p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star^{(l)}, \mathcal{M}_\star)$ à partir d'échantillons issus de la loi *a priori* ! Or si asymptotiquement cet estimateur s'avère satisfaisant, il peut être extrêmement inefficace à nombre fini d'échantillons L, tout simplement parce que sa variance peut être très grande. C'est la problématique bien connue de l'échantillonnage d'*importance*[22] où le choix d'une fonction d'importance accordant plus de poids à des points supports de la queue de la distribution peut permettre de réduire la variance de l'estimateur. Ainsi, hors du cadre asymptotique, il semble inopérant de chercher à approcher l'évidence à partir de (III.2). Fort heureusement, une alternative initialement proposée par [NEWTON ET RAFTERY 1994] consiste à exprimer l'évidence comme une espérance vis-à-vis de la loi *a posteriori* :

$$\frac{1}{p(\mathrm{y}\,|\,\mathcal{M}_\star)} = \int \frac{p(\boldsymbol{\theta}_\star\,|\,\mathrm{y}, \mathcal{M}_\star)}{p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star, \mathcal{M}_\star)}\,\mathrm{d}\theta_\star.$$
$$= \mathrm{E}\left[\frac{1}{p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star, \mathcal{M}_*)}\,\Big|\,\mathrm{y}, \mathcal{M}_*\right] \tag{III.3}$$

Ainsi, d'après (III.3) l'évidence s'écrit comme la *moyenne harmonique a posteriori* de la vraisemblance des paramètres $p(\mathrm{y}\,|\,\boldsymbol{\theta}_\star, \mathcal{M}_\star)$. Cela suggère évidemment d'approcher $p(\mathrm{y}\,|\,\mathcal{M}_\star)$ par une version échantillonnée de la moyenne harmonique des vraisemblances $p(\mathrm{y}\,|\,\boldsymbol{\theta}^{(l)}, \mathcal{M}_\star)$ calculées à partir des L échantillons $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(L)}$ issus de la loi *a posteriori* $p(\boldsymbol{\theta}\,|\,\mathrm{y}, \mathcal{M}_\star)$:

$$\widehat{p(\mathrm{y}\,|\,\mathcal{M}_\star)} = \left[\frac{1}{L}\sum_{l=L_0}^{L_1} \frac{1}{p(\mathrm{y}\,|\,\boldsymbol{\theta}^{(l)}, \mathcal{M}_\star)}\right]^{-1} \tag{III.4}$$

où $L = L_1 - L_0 + 1$. Cette fois, il apparaît clairement que les échantillons requis peuvent correspondre à ceux disponibles en sortie d'un schéma MCMC. Bien que $\widehat{p(\mathrm{y}\,|\,\mathcal{M}_\star)}$ est un estimateur consistent quand la taille L de l'échantillon augmente, sa précision n'est pas garantie. Contrairement à d'autres techniques plus spécialisées [CHIB 1995, GREEN 1995, CHIB ET JELIAZKOV 2001], l'intérêt principal de l'estimateur de la moyenne harmonique réside dans sa simplicité : en effet, il n'exploite que des échantillons *a posteriori* d'un modèle donné et

---

[22]Pour approcher $\mathrm{E}\big[\phi(X)\big]$ par une moyenne de Monte Carlo, la meilleure solution pour obtenir un estimateur à variance minimale n'est pas d'échantillonner X suivant sa loi pour en déduire $1/L\sum_{l=1}^{L}\phi(x^l)$ mais consiste le plus souvent à simuler X selon une densité g de support contenu dans celui de p, pour en déduire $\sum_{l=1}^{L}\omega(x^l)\phi(x^l)/\sum_{l=1}^{L}\omega(x^l)$ où $\omega(x^l) = p(x^l)/g(x^l)$.

des évaluations de vraisemblances qui sont la plupart du temps disponibles comme sous-produit de l'échantillonneur de la loi *a posteriori*. Toutefois, son inconvénient majeur réside dans son instabilité numérique. Bien que consistent, cet estimateur peut être à variance infinie[23] à travers les simulations, même pour des modèles simples. C'est la raison pour laquelle des versions *stabilisées* de cet estimateur ont été proposées dans [RAFTERY et coll. 2007]. Parmi les deux voies explorées dans [RAFTERY et coll. 2007], nous avons retenu celle opérant par réduction de dimension pour stabiliser l'estimateur, car plus simple à mettre en œuvre dans le contexte de détection-estimation conjointe régi par le modèle (III.1).

En quelque mot, cela consiste à remplacer $p(y \mid \boldsymbol{\theta}^{(l)}, \mathcal{M}_\star)$ par $p(y \mid f(\boldsymbol{\theta}^{(l)}), \mathcal{M}_\star)$ dans (III.4) de telle façon que f est une fonction mesurable de $\boldsymbol{\theta}$ qui définit une transformation de *réduction de dimension*. Dans notre contexte, puisque les voxels sont indépendants, nous pouvons procéder séparément pour chaque voxel : $p(y \mid f(\boldsymbol{\theta}^{(l)}), \mathcal{M}_\star) = \prod_j p(\boldsymbol{y}_j \mid f(\boldsymbol{\theta}_j^{(l)}), \mathcal{M}_\star)$. Plus précisément, cela signifie que f peut être explicitée en intégrant hors du problème analytiquement des paramètres voxelliques du modèle d'observation (III.1). Parmi les paramètres candidats, nous avons retenu pour chaque voxel V_j les NRNs $\boldsymbol{a}_j$ car de distribution conditionnelle gaussienne à $\boldsymbol{h}$ fixé, ainsi que les variances de bruit $\sigma_{\varepsilon_j}^2$. Cela semble suffisant pour garantir que $\mathrm{var}\left[(p(y \mid f(\boldsymbol{\theta}), \mathcal{M}_\star))^{-1} \mid y\right] < \infty$. Ainsi, en pratique nous avons considéré l'estimateur suivant :

$$p(\widehat{y \mid \mathcal{M}_\star}) = \left[\frac{1}{L} \sum_{l=L_0}^{L_1} \frac{1}{\prod_j p(\boldsymbol{y}_j \mid f(\boldsymbol{\theta}_j^{(l)}), \mathcal{M}_\star)} \right]^{-1}.$$

En procédant ainsi, nous avons pu calculer la log-évidence $\log p(y \mid \mathcal{M}_m)$ de plusieurs modèles se différenciant par le choix du modèle de bruit ou du modèle de mélange, avec ou sans corrélation spatiale. Le logarithme du facteur de Bayes s'obtient alors directement

$$\log BF_{mn} = \log p(y \mid \mathcal{M}_m) - \log p(y \mid \mathcal{M}_n), \quad \forall (m, n).$$

Nous avons ainsi pu démontrer de façon relativement systématique sur plusieurs jeux de données que :
- le modèle prenant en compte la corrélation spatiale des activations fournit de meilleures performances en termes de compromis sensibilité/spécificité que le modèle de mélange indépendant ; *cf.* [c.15].
- la modélisation de l'autocorrélation temporelle des données par le biais d'un processus AR(1) fournit un taux de faux positifs réduit (meilleure spécificité) notamment à faible SNR (protocoles événementiels lents par exemple) ; *cf.* [?].
- les modèles de mélanges inhomogènes indépendants surpassent leurs homologues homogènes et gaussiens, mais ce résultat n'a pas encore été démontré sur des mélanges spatiaux ; *cf.* [?].

---

[23]mesurable par $\mathrm{V}_{\boldsymbol{\theta}_\star \mid y}\left[p(y \mid \boldsymbol{\theta}_\star, \mathcal{M}_\star)\right]$.

III.4 Thème 3 – Analyse sur la surface corticale (depuis 2005)

La surface corticale, puisque constituée de macro-colonnes de neurones, définit le substrat anatomique sous-tendant les activations cérébrales. Pourtant la plupart des analyses intra-sujet réalisées en IRMf se fondent sur le volume tridimensionnel d'acquisition. Cet état de fait est la conséquence de plusieurs difficultés, aux premiers rangs desquels figurent d'une part, l'extraction robuste d'un maillage de cette surface et d'autre part, la projection des données fonctionnelles sur cette surface. Grâce aux compétences développées depuis dix ans par l'équipe de Jean–François MANGIN au SHFJ, et à la qualité des nouveaux imageurs Siemens disponibles à NEUROSPIN, la première étape est désormais bien maîtrisée. La seconde étape a fait l'objet d'intenses recherches depuis plusieurs années [ANDRADE 2002, a.10, OPERTO et coll. 2006]. J'ai participé au développement de certaines idées dans le cadre d'une collaboration entre Christophe GROVA[24] et notre équipe au SHFJ, en particulier avec Salima MAKNI [a.10]. Comme indiqué ci-après, le but était d'améliorer l'interpolation des données acquises dans le volume aux nœuds du maillage extrait. Une fois les données projetées sur la surface, il devient pertinent de développer des approches robustes pour la détection des activations sur ce support anatomique et la caractérisation de l'hémodynamique associée. En effet, des premiers travaux [ANDRADE 2002] ont montré une amélioration de la sensibilité de détection dès lors que le lissage spatial des données est opéré géodésiquement à cette surface, plutôt qu'isotropiquement dans le volume du cerveau. Toutefois, pour dépasser le paradigme classique d'analyse des données, il convient de renoncer à ce lissage et chercher à introduire un modèle génératif de l'extension spatiale de la distribution des activations. Dans les paragraphes suivants, je précise quel type d'approches j'ai commencé à promouvoir dans ce domaine.

III.4.1 Sous-thème 3a – Interpolation du signal IRMf sur la surface corticale (2005-06)

L'objectif d'une telle interpolation est d'associer une série temporelle IRMf à chaque nœud du maillage de la surface corticale. Concevoir un schéma optimal d'interpolation des données d'IRMf repose sur un compromis entre d'une part, le choix d'un nombre suffisamment grand de noyaux d'interpolation, en raison de la nature étendue et distribuée de l'activité hémodynamique, et d'autre part éviter de mélanger des données issues de différentes structures anatomiques. De plus, le choix de ces noyaux d'interpolation se doit d'être robuste vis-à-vis des distorsions présentes dans les images fonctionnelles, et d'erreur de recalage entre les données anatomiques et fonctionnelles.

Forte de ces pré-requis, la méthode proposée dans [p.6, a.10] ajuste automatiquement le niveau de ce compromis en définissant des noyaux d'interpolation autour de chaque nœud de la surface corticale à l'aide d'un diagramme de VORONOÏ géodésique à cette surface. En substance, une distance 3D géodésique est exploitée pour tenir compte de la morphologie circonvoluée du cortex, et les noyaux d'interpolation sont ensuite définis autour de chaque nœud à l'aide du diagramme de VORONOÏ ce qui garantit par construction la contrainte de proximité autour de chaque nœud et la robustesse vis-à-vis des erreurs de recalage. Pour intégrer des données fonctionnelles sur des aires suffisamment grandes, le support spatial utilisé pour générer le diagramme de VORONOÏ était constitué du masque de la matière grise segmenté à partir de l'IRM anatomique, après une étape de dilatation morphologique d'environ 3 mm, pour prendre en compte la résolution spatiale plus grossière des données d'IRMf. L'algorithme se décompose alors en trois étapes une fois la segmentation de la surface corticale disponible :

---

[24]Post-doctorant à l'Université de Mc Gill à Montréal.

1. définir les masques anatomiques et fonctionnels utilisés comme supports spatiaux pour l'interpolation ;
2. construire les noyaux d'interpolation attachés à chaque nœud de la surface à partir du diagramme de Voronoï ;
3. déduire le signal fonctionnel attaché à chaque nœud.

Un exemple d'activité 3D simulée à l'intérieur de la matière grise est présenté en rouge à la Fig. III.14. Cette activité est ensuite interpolée sur la surface corticale selon deux approches, une technique classique de moyennage local où les valeurs contenues dans une sphère centrée en chaque nœud et de rayon $R = 5mm$ sont moyennées (voir Fig. III.15(a)), et celle que nous avons développée (voir Fig. III.15(b)) qui produit un foyer d'activation plus focal autour du gyrus, reproduisant ainsi plus fidèlement l'activité simulée.
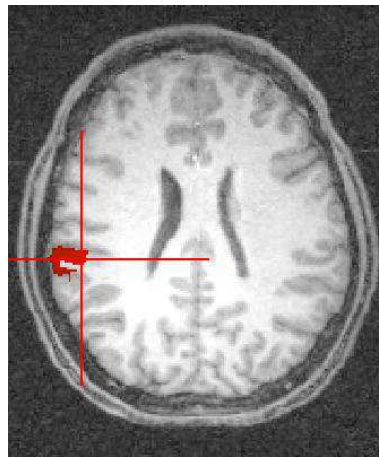
Fig. III.14 – Carte d'activation simulée.
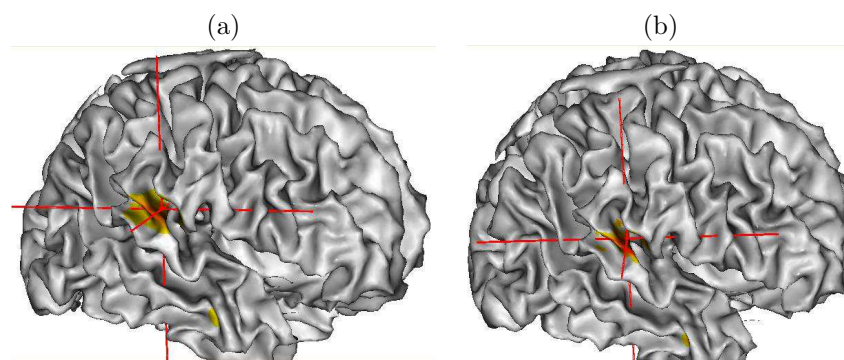
Fig. III.15 – Comparaison de schémas d'interpolation du signal fonctionnel sur la surface corticale. (a) version moyennée des valeurs fonctionelles à l'intérieur d'une sphère de rayon $R = 5mm$. (b) version développée à l'aide d'un diagramme de Voronoï geodésique à la surface.

III.4.2 Sous-thème 3b – Détection-estimation conjointe sur la surface (depuis 2007)

Il s'agit ici d'adapter sur la surface corticale les méthodes évoquées au **Thème 2**. Les nouveautés concernent essentiellement deux aspects : la définition d'une parcellisation de la surface

corticale et la régularisation spatiale le long de cette surface. Nos contributions directes visent plutôt le second point, mais néanmoins, le premier est brièvement discuté.

La *parcellisation* anatomo-fonctionnelle des données en régions homogènes le long de cette surface peut être menée selon différents critères. Récemment, l'équipe d'Olivier COULON (*i.e.,* à travers la thèse de Cédric CLOUCHOUX), faisant écho aux travaux d'Arnaud CACHIA au sein du laboratoire[25], a proposé des techniques de parcellisation anatomique de cette surface en s'appuyant sur un système de coordonnées (latitude, longitude) cohérent pour cette topologie [CLOUCHOUX et coll. 2006]. Ce système est défini à partir d'un certain nombre de contraintes structurales telles que la position ou les relations de voisinage de certains sillons ou de leurs racines sulcales (marqueur ascendant des sillons). Cette approche s'appuie donc sur la sulcogénèse, théorie développée par Jean RÉGIS, à laquelle le laboratoire participe depuis plus de dix ans au travers des travaux de Jean-François MANGIN. Les parcellisations ainsi produites bien qu'issues d'informations dites structurales mettent en évidence une bonne corrélation anatomo-fonctionnelle.

De façon alternative mais non-exclusive, Bertrand THIRION et Guillaume FLANDIN ont développé des parcellisations plutôt guidées par des informations fonctionnelles [FLANDIN 2004, a.9], bien que respectant certaines contraintes géométriques comme la *connexité* des parcelles. Leurs approches ont le mérite d'être multi-sujets mais procèdent dans le volume cérébral et nécessitent ainsi une étape préalable de normalisation de tous les cerveaux dans un espace de référence commun. Leurs travaux font l'objet actuellement d'évolutions dans le cadre de la thèse d'Alan TUCHOLKA sous la direction de Bertrand THIRION, pour opérer sur la surface.

Sur la base d'une parcellisation anatomique illustrée à la Fig. III.16, nous avons déjà commencé à tester les approches de détection-estimation conjointe sans régularisation spatiale. Un résultat illustrant des activations dans le cortex pariétal gauche durant une tâche de calcul est présenté à la Fig. III.16(b). Certaines activations parasites assimilables à des faux positifs pourraient disparaître à l'aide d'un modèle de régularisation spatiale géodésique à cette surface. Nous travaillons dans cette direction actuellement dans le cadre de la thèse de Thomas VINCENT car cette extension présente un réel intérêt pour la perspective de la fusion symétrique multimodale EEG/IRMf.

---

[25]voir par exemple [CACHIA et coll. 2003] pour des méthodes de parcellisation anatomique en gyri de la surface corticale
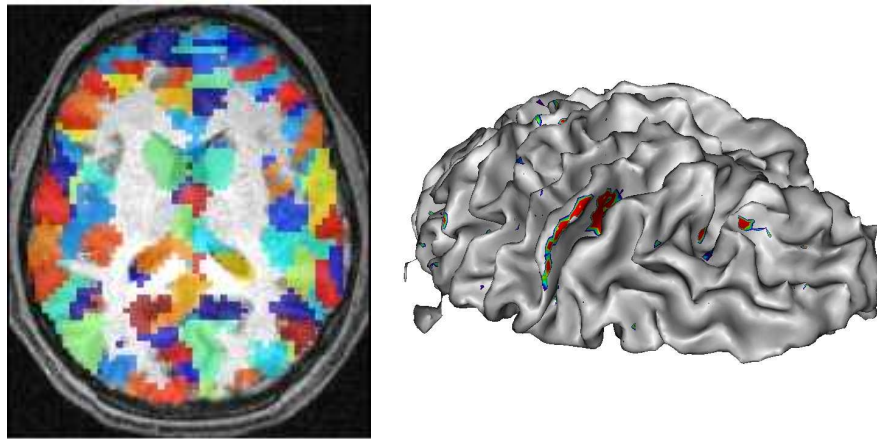
FIG. III.16 – À gauche, parcellisation anatomique du ruban cortical, où chaque hémisphère est parcellisé séparament. Ici, chaque couleur représente une parcelle différente, et la connexité des parcelles est garantie par construction. À droite, détection des foyers d'activation dans le cortex pariétal gauche en réponse à une tâche de calcul.

III.5 Thème 4 – Reconstruction régularisée en IRM parallèle (depuis 2005)

Ce thème n'étant pas détaillé dans la partie suivante, cette section débute par quelques rappels sur l'imagerie parallèle (IRMp) dont un des buts est l'amélioration de la résolution temporelle des volumes acquis. Dans un deuxième temps, la position du problème de reconstruction parallèle et les solutions déjà existantes sont précisées. Enfin, mes contributions sur la reconstruction régularisée en IRMp sont exhibées : elles imputeront directement les résultats d'analyse statistique ultérieure (*cf.* **Thème 5**).

III.5.1 Rappels sur l'IRM parallèle

Grâce à l'apparition de techniques d'IRM parallèle (IRMp), une réduction substantielle du *temps d'acquisition* est obtenue sans nécessairement accroître les performances techniques des gradients. L'idée sous-tendant ce principe est la suivante.

En IRM classique, le contraste de l'objet est encodé dans le spectre de résonance par un gradient de champ magnétique. Ce concept d'encodage de gradient pour la localisation du signal impose une restriction de taille : une seule position de l'espace-k (*i.e.,* de Fourier) peut être échantillonnée en un seul instant, faisant de la vitesse de balayage de l'espace-k le facteur déterminant du temps d'acquisition. L'IRM parallèle fonctionne sur un tout autre principe en tirant partie du fait que la contribution d'une source de signal *i.e.,* d'une antenne à la tension induite à la sortie de la bobine varie avec la position relative. En d'autres termes, la connaissance du profil de sensibilité spatiale de l'antenne nous renseigne sur l'origine du signal de résonance détecté, information qui peut être utilisée pour la formation de l'image. Contrairement à la position dans l'espace-k, le profil de sensibilité est une caractéristique de l'antenne et ne fait pas référence à l'**état de l'objet** en cours d'examen. Par conséquent, les échantillons situés à différentes positions dans l'espace-k peuvent être obtenus en *un instant* en utilisant simplement différentes antennes en parallèle, impliquant un temps d'acquisition réduit sans avoir à parcourir l'espace de Fourier plus rapidement. Ainsi, le profil de sensibilité des antennes complète l'encodage par harmoniques pures du domaine de Fourier, que l'on peut dès lors restreindre.
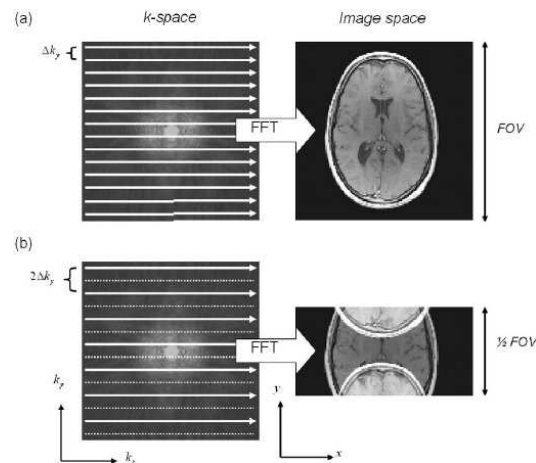
FIG. III.17 – (a) : Acquisition conventionnelle de l'espace de Fourier échantillonné complet (aux fréquences horizontales et verticales $n\Delta k_x$ et $n\Delta k_y$), donnant une image à champ de vue (FOV) complet après transformation de Fourier inverse. (b) : Acquisition sous-échantillonnée d'un facteur de réduction $R = 2$ dans la direction de phase Oy donnant une image de champ de vue réduit (FOV/2) artefactée. En trait plein, les lignes acquises de l'espace-k, en trait tireté, les lignes manquantes dans la situation (b).

Dans la pratique, il faut en fait établir un compromis entre le coût (nombre d'antennes dont on dispose, complexité de la géométrie du réseau d'antennes, chute du RSB et le temps d'acquisition, si bien qu'on n'acquiert qu'une certaine proportion de l'espace de Fourier. On « saute » ainsi des lignes de l'espace-k en omettant une fraction de l'encodage de phase, tout en gardant la même résolution spatiale et le même niveau de contraste dans l'image. Alternativement, le temps économisé en IRMp peut aussi servir à améliorer la résolution spatiale (*i.e.,* réduire la taille de voxels), mais cette fois le temps d'acquisition reste le même que pour un examen standard non parallèle. Cette amélioration de la résolution spatio-temporelle en IRMp se fait au prix d'une dégradation du RSB dans l'image. Une partie de cette dégradation provient de la nature spatialement corrélée des informations acquises sur les différentes antennes, ce qui rend mal posé le problème d'inversion à résoudre pour recouvrer une image à champ de vue complet.

Ces dix dernières années, de véritables progrès dans le développement des techniques d'imagerie parallèle ont été accomplis, produisant ainsi une multitude de techniques de reconstruction parallèles différentes. Toutes ont vocation à reconstruire les données parallèles et à supprimer l'artefact de repliement. Actuellement, les plus connues sont SMASH [SODICKSON 2000, MCKENZIE et coll. 2001], SENSE [PRUESSMANN et coll. 1999, PRUESSMANN et coll. 2001] et GRAPPA [GRISWOLD et coll. 2005]. Plus récemment, des extensions « autocalibrées » ont été proposées à travers AUTO-SMASH [JAKOB et coll. 1998, HEIDEMANN et coll. 2001], GENERALIZED-SMASH [BYDDER et coll. 2002], PILS [GRISWOLD et coll. 2000] et SPACE-RIP [KYRIAKOS et coll. 2000]. Toutes ces méthodes requièrent des informations supplémentaires sur la sensibilité des antennes afin d'éliminer les effets de sous-échantillonnage de l'espace de Fourier (illustré à la figure III.17). Cette information de sensibilité peut être obtenue en une seule fois au début de l'examen au moyen d'une *carte de sensibilité* ou bien pour chaque session d'IRMf au moyen de l'acquisition de quelques lignes supplémentaires de l'espace-k (phase d'auto-calibration).

Les méthodes actuelles de reconstruction en IRMp peuvent être ordonnées en deux grandes catégories : celles qui opèrent la reconstruction dans l'**espace image** : SENSE, PILS, consistent en une procédure inverse de dépliement, alors que celles opérant dans l'**espace de Fourier** (*e.g.,*
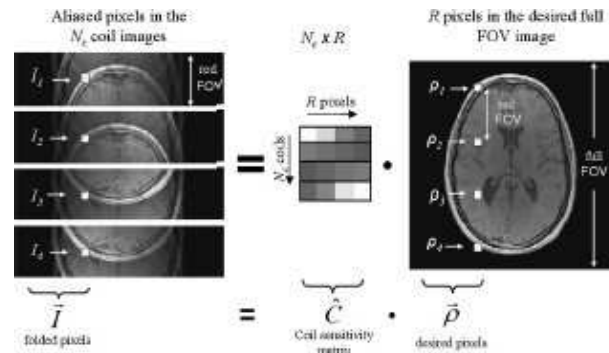
FIG. III.18 – Illustration de la relation de base de l'imagerie SENSE utilisant un facteur de réduction $R = 4$ avec $L = 4$ antennes. $\vec{I}$ contient les pixels repliés en une position donnée dans les images à champ de vue réduit (FOV/4) acquises par les antennes. La matrice de sensibilité, ici notée $\widehat{C}$ rassemble les valeurs de sensibilité correspondante des antennes aux positions des $R = 4$ pixels impliqués dans l'image à champ de vue complet $\vec{\rho}$.

SMASH, GRAPPA) se résument à un problème d'interpolation dans l'espace de Fourier (calcul des lignes manquantes). Des techniques hybrides combinant les deux approches sont aussi possibles (*cf.* SPACE-RIP).

L'imagerie SENSE sur laquelle nous nous concentrons appartient aux méthodes de reconstruction dans l'espace image ; elle procède en deux temps :

1. la première étape consiste à reconstruire pour chaque antenne l une image intermédiaire de taille réduite (FOV/R si R est le facteur de réduction ou d'« accélération ») à l'aide de la transformée de Fourier inverse discrète. Pour un échantillonnage régulier de l'espace-k, cette reconstruction est donc assez rapide car implantable par FFT. Pour un échantillonnage plus complexe, de type spirale ou autre, cette étape est beaucoup plus coûteuse (voir [PRUESSMANN et coll. 2001, WEIGER et coll. 2002] pour plus de détails).

2. la seconde étape consiste à créer une image à champ de vue complet à partir de l'ensemble des images intermédiaires. Pour réaliser cette opération, on doit inverser le processus de superposition des images. Pour chaque pixel dans le champ de vue réduit, les contributions du signal d'un certain nombre de positions du champ de vue complet doivent être séparées. Comme illustré à la Figure III.18, ces positions forment une grille cartésienne calculable à partir du facteur de réduction R.

Sous cette forme la reconstruction en imagerie SENSE s'identifie donc à un problème de séparation de sources (spatiales) linéaire dans lequel la matrice de mélange est complexe et varie dans l'espace mais peut être relativement bien estimée à partir de l'acquisition supplémentaire d'une carte de sensibilité en écho de gradient. La clé de la séparation du signal tient au fait que dans chaque image mono-canal, la superposition du signal apparaît avec des poids spécifiques qui dépendent des sensibilités locales des antennes.

Rappelons enfin que l'objet recherché est une densité de protons donc *positif*, mais les observations disponibles dans l'espace image sont complexes, car obtenues après transformée de Fourier inverse bidimensionnelle des données de l'espace-k contaminées par les imperfections du système d'imagerie.

Pour illustrer la complexité du problème de reconstruction en fonction du facteur de réduction, nous présentons à la Fig. III.19 deux résultats obtenus à l'aide de l'algorithme SENSE classique sur des images anatomiques acquises à 1.5 T. Alors que l'image reconstruite à champ de vue complet pour $R = 2$ est d'une qualité satisfaisante, des artefacts de repliement apparaissent

au centre de l'image de droite obtenue avec $R = 4$, c'est-à-dire là où le problème d'inversion s'avère le plus sévère et le plus mal posé.
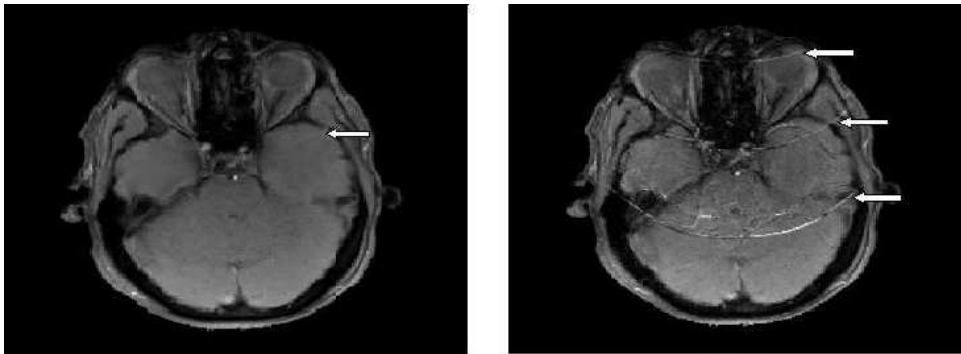
FIG. III.19 – Artefacts de repliement dans les images à champ de vue complet reconstruites à l'aide de l'algorithme SENSE classique avec $R = 2$ (à gauche) et $R = 4$ (à droite).

Dans le but de limiter la dégradation du RSB inhérent aux acquisitions parallèles, il est d'usage d'introduire une régularisation sur l'image à reconstuire. Dans la suite, je précise mes contributions en reconstruction régularisée SENSE. Ces travaux sont encore en cours. Certains aspects ont fait l'objet d'une présentation sous forme de poster aux journées PARISTIC en novembre 2006 [p.10]. D'autres ont été établis en 2007 par Lotfi CHAÂRI dans le cadre de son master et du projet OPTIMED avec Jean-Christophe PESQUET et Amel BENAZZA. Enfin, dans le cadre de la thèse de Cécile RABRAIT, nous avons testé l'apport de modèles régularisant frustes *i.e.,* quadratiques sur des données *Echo Volumar Imaging*.

III.5.2   Sous-thème 4a – Reconstruction supervisée (depuis 2006)

Fort de mon expérience en thèse sur des modèles de régularisation *composite* sur le module d'objets complexes [a.1], je développe en collaboration avec Jean-Christophe PESQUET, Lotfi CHAÂRI, Cyril POUPON et Cécile RABRAIT[26] des méthodes qui visent à restaurer un objet positif, ici une densité de protons. plusieurs alternatives sont envisagées, toutes pour l'instant s'appuient sur une régularisation par pénalisation.

Les approches les plus répandues dans la littérature sont relativement frustes dans la mesure où elles exploitent une régularisation de type Tikhonov [LIN et coll. 2004, LIN et coll. 2005], c'est-à-dire en norme L_2 sur la solution, ou s'appuient de façon voisine sur une décomposition tronquée en valeurs singulières de l'objet [HOGE et coll. 2004]. Parmi les travaux les plus aboutis en régularisation quadratique [LIN et coll. 2004], l'apprentissage du paramètre de régularisation est réalisé par la méthode de la « courbe en L » [HANSEN 1992]. Il pourrait tout autant être obtenu, sans doute plus rapidement, par maximisation de la vraisemblance gaussienne. Ce maximiseur local ou point stationnaire de la log-vraisemblance se calcule à l'aide d'un algorithme EM [DEMPSTER et coll. 1977].

L'objet finalement obtenu est une solution à valeurs complexes, relativement lisse spatialement et issue de la résolution d'un système linéaire. La positivité de la solution est finalement garantie *a posteriori*, par la prise du module sur l'image complexe restaurée. La Fig. III.20 illustre ce type de régularisation quadratique. Les résultats obtenus démontrent une bonne réduction des distorsions pour un facteur de réduction $R = 2$, en comparaison avec le cas d'une acquisition sans sous-échantillonnage de l'espace-k ($R = 1$).

---

[26]Cette collaboration est officialisée au sein du programme OPTIMED financé par l'ANR en 2006.
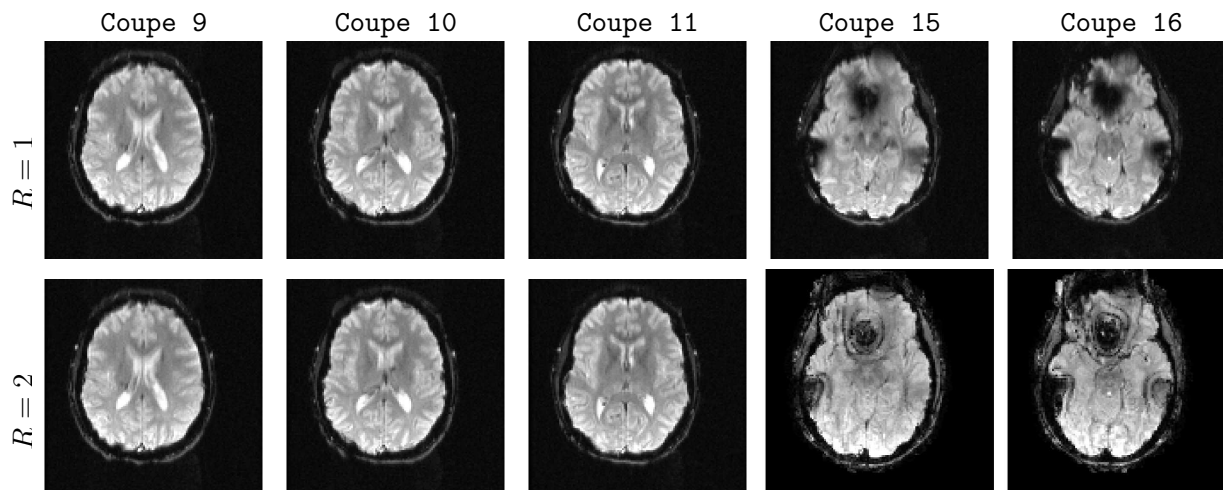
FIG. III.20 – Acquisition à haute résolution spatiale (matrice 128×128). **Récupération de signal** et **réduction des distorsions** pour $R = 2$ (coupes 9-11) par rapport à $R = 1$ en raison de la durée plus brève du train d'échos. Artefacts sur les coupes basses (15,16).

Afin de mieux préserver des discontinuités dans l'image cible, par exemple les interfaces entre les différents tissus cérébraux, nous proposons plutôt d'utiliser des fonctions de pénalisation du type variation totale, qui restaurent fidèlement des variations d'intensité au delà d'un certain seuil fixé [LI 1995]. Pour garantir la convexité du critère et simplifier sa minimisation, nous nous focalisons plutôt sur des pénalisations convexes (Huber, potentiel hyperbolique,...) [HUBER 1981, CHARBONNIER et coll. 1997, IDIER 2001].

Selon le type de critères à minimiser, plusieurs algorithmes peuvent être envisagés, mais dès lors que la pénalisation n'est plus quadratique, la solution ne résulte plus de la résolution d'un système linéaire. Le cadre que nous privilégions est celui des algorithmes semi-quadratiques [IDIER 2001, a.2], qu'on peut encore interpréter comme des méthodes de quasi-Newton à pas fixe [NIKOLOVA ET NG 2005, ALLAIN et coll. 2006]. Compte tenu des avancées théoriques récentes réalisées dans ce domaine au cours de la thèse de Christian LABAT (voir notamment [LABAT ET IDIER 2007]), je m'efforce depuis plus de dix huit mois de recruter un post-doctorant au sein du projet OPTIMED pour appliquer puis étendre les contributions de Christian LABAT au cas de données complexes. Après plusieurs échecs, il semble qu'un candidat biélorusse, Victor SHASHILOV puisse débuter en avril 2008 son stage post-doctoral sous ma responsabilité à NEUROSPIN.

III.5.3 Sous-thème 4b – Reconstruction auto-calibrée dans l'espace ondelette (depuis 2007)

Depuis octobre 2007, data à laquelle Lotfi CHAÂRI a obtenu une bourse doctorale de la région Ile de France, je coencadre officiellement son travail de thèse. Nous avons déjà pu obtenir des résultats plus aboutis notamment en reconstruction régularisée dans le domaine transformé en ondelettes. Après un premier travail en reconstruction supervisée, on s'est attaché à tenter de régler « le plus automatiquement possible » les hyperparamètres spécifiant le terme de régularisation. Pour ce faire, Lotfi CHAÂRI s'est attaché à développer une version *auto-calibrée* de l'algorithme de reconstruction, signifiant ainsi que les hyperparamètres sont estimés à partir d'une image de référence (en imagerie anatomique). Ces travaux préliminaires ont fait l'objet d'une soumission [c.23] à la prochaine conférence ISBI qui se tiendra à Paris en avril 2008. Bien

20 juin 2008

sûr dans une situation réaliste et plus intéressante, l'IRMf, il faudra développer une extension complètement non-supervisée, beaucoup plus coûteuse numériquement. Là encore, l'inférence pourra être conduite dans un cadre stochastique en considérant ces paramètres comme des variables aléatoires inobservées. Dans les lignes qui suivent, je décris succinctement les avancées réalisées en reconstruction dans l'espace ondelette en IRMp.

Les méthodes à base de décomposition sur lesquelles nous travaillons s'appuient typiquement sur des bases d'ondelettes 2D séparables et orthogonales qui opèrent coupe à coupe. Il s'agit ensuite de régulariser l'estimation des coefficients d'ondelette de l'image (coefficients d'approximation et de détails horizontaux, verticaux et diagonaux) en introduisant une pénalisation spécifique. Comme dans le cas de la reconstruction SENSE classique ou régularisée dans l'espace image, la solution dans le domaine transformé s'écrit comme le minimiseur d'un critère pénalisé dont on peut faire l'interprétation bayésienne pour lui adjoindre le qualificatif d'estimée au sens du Maximum *A Posteriori*. La solution dans l'espace image s'obtient alors naturellement par application de la transformation inverse à cette estimée du MAP.

En analysant la distribution empirique des parties réelle et complexe des coefficients de détail, on a pu montrer que leurs histogrammes s'ajustaient bien à une distribution gaussienne généralisée (*cf.* [p.13, c.23]). À la résolution la plus grossière, la distribution empirique des parties réelle et complexe des coefficients d'approximation s'apparente quant à elle à une densité gaussienne. Par ailleurs, une quasi décorrélation a pu être mise en évidence entre les parties réelle et imaginaire des coefficients d'ondelette. On a donc considéré qu'ils étaient indépendants et même identiquement distribués au sein de chaque sous-bande d'analyse.

Dans ce cadre, compte tenu de la pénalisation issue de la gaussienne généralisée, le critère à minimiser peut admettre une partie non différentiable. Dans ces conditions, l'algorithme d'optimisation mis en œuvre s'appuie sur la notion fondamentale en optimsation convexe d'*opérateur proximal* [DAUBECHIES et coll. 2004, COMBETTES ET WAJS 2005, CHAUX et coll. 2007], qui généralise l'opérateur de seuillage doux.

Dans [c.23], on démontre que cette méthode de reconstruction converge rapidement (150 itérations) sur un problème de taille standard (IRM T_1 de taille $256 \times 256 \times 192$). La Fig. III.21 illustre aussi qu'elle se comporte mieux que la reconstruction SENSE régularisée quadratiquement pour $R = 4$. Toutefois, même pour cette dernière certains artefacts demeurent au centre de l'image. L'augmentation du niveau de résolutions dans la pyramide d'analyse permet de lever ce problème en partie seulement comme illustré dans [c.23]. Nous travaillons actuellement à l'amélioration de la modélisation statistique *a priori* des coefficients de détail. Il semblerait qu'un modèle de mélange plus souple qu'une seule distribution soit potentiellement mieux adapté pour prendre en compte la dynamique très étalée de ces coefficients.
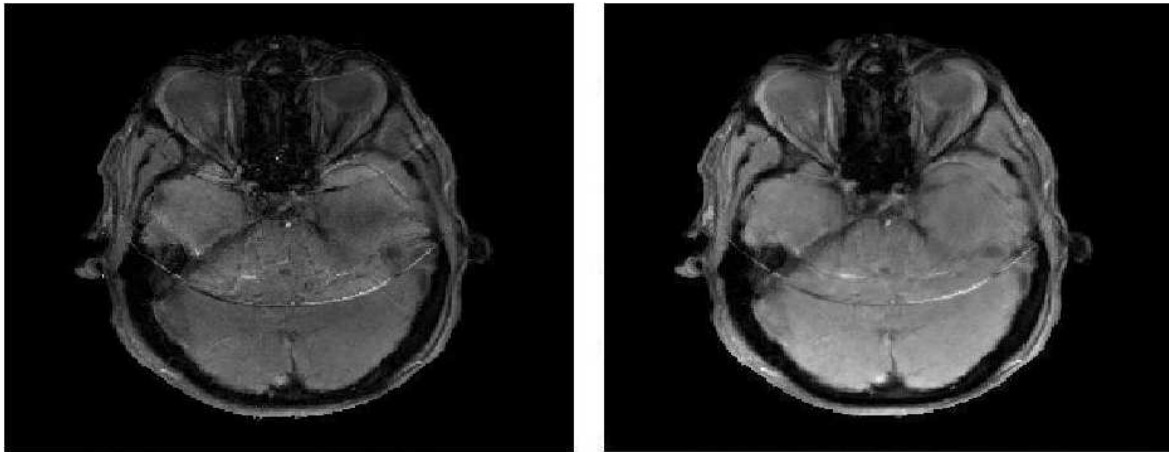
Fig. III.21 – Images reconstruites par régularisation quadratique (à gauche) et par régularisation dans l'espace ondelette pour des données acquises avec un facteur de réduction $R = 4$.

III.6   Thème 5 – IRM fonctionnelle « ultra-rapide »

La plupart des travaux en neuro-imagerie fonctionnelle par IRM s'appuient sur la séquence 2D d'acquisition rapide planaire dite EPI (*Echo Planar Imaging*), qui procède coupe par coupe[27], par balayage de l'espace-k bidimensionnel[28] pour acquérir le signal fonctionnel dans l'ensemble du volume cérébral avec une résolution spatiale planaire de 2 à 3 mm de côté et une résolution temporelle de 1 à 2 secondes. Ainsi, la séquence EPI est bien adaptée à la détection des activations cérébrales et à la cartographie des régions impliquées dans une expérience de neurosciences. Toutefois, avec l'avènement des protocoles événementiels rapides [ROSEN et coll. 1998] et l'intérêt croissant des neuroscientifiques pour les propriétés temporelles de la fonction de réponse hémodynamique, l'amélioration de la résolution temporelle des données acquises devient un enjeu majeur. Deux voies de recherche s'ouvrent alors :

I. développer une imagerie à haute résolution *intrinsèque*, c'est-à-dire qui poursuit un phénomène potentiellement neuronal donc plus rapide que l'effet BOLD [LE BIHAN et coll. 2006, LE BIHAN 2007, MILLER et coll. 2007] ;

II. améliorer la résolution temporelle *extrinsèque*, c'est-à-dire la vitesse d'imagerie du système d'acquisition pour échantillonner un volume en 100 à 200 ms environ.

Naturellement, ces deux pistes sont complémentaires, la deuxième pouvant servir les intérêts de la première. Nous discuterons de la voie intrinsèque dans les perspectives du chapitre IV au moment d'évoquer nos travaux en cours sur l'analyse des données d'IRM de *diffusion fonctionnelle*, sujet du stage post-doctoral de Laurent RISSER que j'encadre. Nous explorons maintenant la seconde voie.

La haute résolution extrinsèque peut être le fruit d'une séquence d'imagerie spécialisée, par exemple la séquence *Echo Volumar Imaging* (EVI), une extension 3D de l'EPI, dans laquelle l'espace de Fourier tridimensionnel est encodé à l'aide d'une seule impulsion radio-fréquence. Des premiers travaux ont déjà rapporté l'intérêt de cette séquence d'imagerie pour l'IRMf [MANSFIELD et coll. 1995, YANG et coll. 1997, ZWAAG et coll. 2006]. Il est utile de rappeler les propriétés les plus saillantes de l'EVI qui en font un outil de choix pour l'IRMf :

---

[27]Chaque coupe est acquise en 60 à 100 ms environ suite à une impulsion radio-fréquence.

[28]L'axe Ox de cet espace correspond à la direction d'encodage en fréquence tandis que l'axe Oy correspond à une direction d'encodage de phase.

1. la possibilité d'acquérir le signal de résonance pour un sous-volume du cerveau en 200 ms ;

2. l'insensibilité aux artefacts cardiaque et respiratoire due à une échantillonnage suffisamment rapide respectant les conditions de validité du théorème de Shannon-Nyquist ;

3. une acquisition 3D mono radio-fréquence (*single shot*) ne nécessitant pas de correction du délai d'acquisition inter-coupes, et minimisant les risques de mouvement du sujet à l'intérieur d'un volume acquis ;

4. la réduction des effets de flux vasculaire entrant (*inflow effect*) due à l'utilisation d'une séquence d'imagerie 3D [Frahm et coll. 1994]. Ces effets peuvent rendre contestable l'interprétation des résultats en IRMf notamment à champ magnétique relativement faible (*e.g.,* 1.5 Tesla) [Lu et coll. 2002].

Néanmois, l'EVI n'a été jusqu'ici que peu utilisée en IRMf en raison de la sollicitation forte des gradients qu'elle entraîne. Ainsi, en raison de contraintes matérielles, la longueur des trains d'échos demeure relativement importante en EVI et la bande passante le long de la direction de partition (deuxième direction d'encodage de phase) est très faible. Par conséquent, disposer d'une résolution spatiale relativement élevée et d'une bonne couverture du cerveau en EVI implique l'apparition d'importantes distorsions (artefacts de susceptibilité magnétique) et de pertes de signal dues à la relaxation en T_2^*. Pour dépasser ces limitations, il était nécessaire de trouver de nouveaux compromis. Une première solution a consisté à favoriser la direction d'encodage en fréquence, moins gourmande en temps d'acquisition, et ainsi acquérir des volumes très anisotropes. Une seconde, celle poursuivie, s'est attachée plutôt à réduire le champ de vue pour raccourcir les trains d'écho.

Récemment, une modification de la séquence EVI a été proposée par [Zwaag et coll. 2006]. Elle a conduit à une amélioration significative des images acquises et s'appuie sur des gradients de « rembobinage » pour réduire les artefacts de *ghosting* le long de la direction de partition. Elle inclut aussi des impulsions de saturation du signal à l'extérieur d'un volume donné pour réduire le champ de vue. Toutefois, ce travail reste limité à l'acquisition de volumes anisotropes (*e.g.,* matrices $64 \times 32 \times 28$ et $192 \times 96 \times 24$), avec une résolution spatiale planaire comparable à celle de l'EPI et un rapport signal-à-bruit élevé.

La thèse de Cécile Rabrait[29] visait à s'affranchir de ces problèmes. La séquence EVI modifiée qu'elle a proposée combine ces impulsions de saturation [Leroux et coll. 1998], une acquisition parallèle et une reconstruction Sense bidimensionnelle le long des directions d'encodage de phase et de partition[30]. En conséquence, les volumes acquis sont quasi-isotropes avec une large couverture du cerveau sans concession sur la résolution temporelle (200ms par volume) [c.13, a.12]. La résolution spatiale (au mieux $4.5 \times 4.5 \times 5$mm^3) est néanmoins un peu moins bonne que celle atteinte dans [Zwaag et coll. 2006]. Dans la suite, je précise les points sur lesquels je suis intervenu dans ce travail, qui a débouché sur une publication dans la revue *Journal of Magnetic Resonance Imaging* [a.12].

III.6.1 Sous-thème 5a – Caractérisation statistique des signaux *Echo Volumar Imaging* (depuis 2005)

A partir de l'automne 2005, J'ai supervisé le travail de Cécile Rabrait sur la caractérisation statistique de la séquence EVI utilisée en IRMf comme alternative à la séquence EPI. Nous avons ainsi mis en évidence sur des protocoles visuels événementiels lents, des différences significatives sur la sensibilité statistique de cette séquence par rapport à la séquence EPI, conduisant notamment à des topographies d'activation contrastées. La localisation anatomique du volume acquis

---

[29]doctorante au SHFJ puis à Neurospin sous la direction de Denis Le Bihan et sous l'encadrement conjoint de Franck Lethimonnier

[30]L'acquisition parallèle 2D permet raisonnablement d'exploiter un facteur de réduction jusqu'à 4, à 1.5 Tesla.

au cours de ce protocole visuel est décrite à la Fig. III.22. Elle met en relief la couverture du pôle occipital.
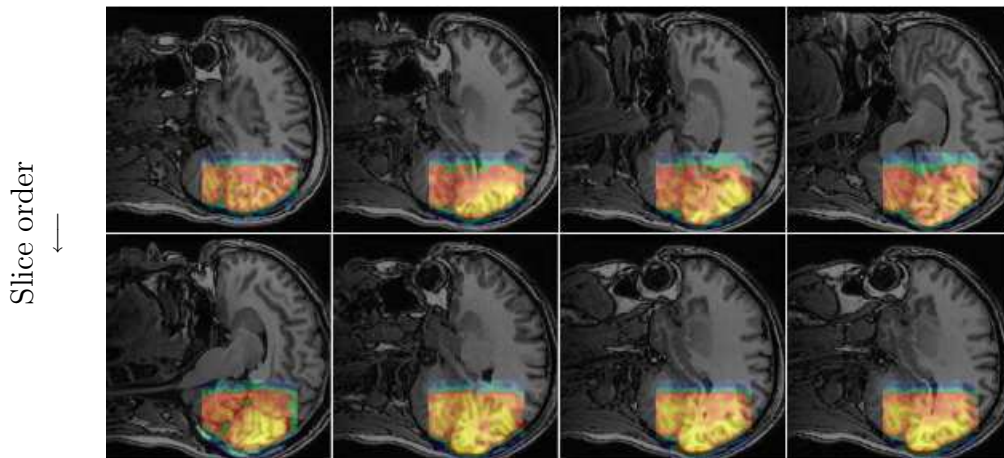
FIG. III.22 – Localisation anatomique du volume acquis lors de la séquence zoom EVI au cours de protocoles visuels.

Nos premiers résultats ont permis de démontrer que la séquence EVI parallèle localisée conduisait à détecter moins de voxels activés mais que ces activations étaient plus focales, c'est-à-dire moins étalées, à l'instar de la Fig. III.23.

Nous avons en outre évalué l'influence de la régularisation spatiale opérée au cours de la reconstruction SENSE. Les résultats obtenus sont présentés dans [a.12] et corroborent ceux déjà publiés dans [LIN et coll. 2005] : ils démontrent que cette régularisation améliore la puissance de détection et donc la significativité des activations vues en imagerie SENSE. Une de nos contributions originales a consisté à mesurer aussi cette influence sur la dynamique de la fonction de réponse hémodynamique estimée. Cette régularisation spatiale a pour effet d'homogénéiser spatialement les FRH estimées, les dynamiques devenant plus homogènes dans des voxels voisins.

Après cette phase de validation, nous nous sommes tournés vers une application cognitive en collaboration avec Ghislaine DEHAENE-LAMBERTZ afin de démontrer l'apport de cette séquence d'imagerie pour les neurosciences. Il s'agissait d'un protocole dédié au langage et à la cartographie des aires qui s'habituent ou non à la répétition des phrases. L'utilisation d'une acquisition sagitale (direction de partition) nous a permis de couvrir tout le cerveau dans l'axe gauche-droite (voir Fig. III.24) et de nous affranchir des impulsions de saturation visant à réduire le champ de vue. Le prix à payer est une légère diminution de la résolution spatiale (au mieux $5 \times 5 \times 6$mm^3).

Comme en EPI, nous avons pu mettre en évidence des activations dans les lobes temporaux (voir Fig. III.25), de façon prédominante à gauche dans le sillon temporal supérieur (STS). Les activations dans le gyrus de HESCHL ne montrent pas d'habituation tandis que certains voxels détectés dans le STS présentent une habituation de l'amplitude. Des travaux sont encore en cours pour mesurer la vitesse de cette habituation, qui pourrait être mieux précisée compte tenu de la résolution temporelle de l'EVI, par rapport aux précédents travaux en EPI [a.4]. Cette application figure dans l'article [a.12], qui est reproduit en Annexe IX.

(a) Activations en EVI   (b) Activations en EPI
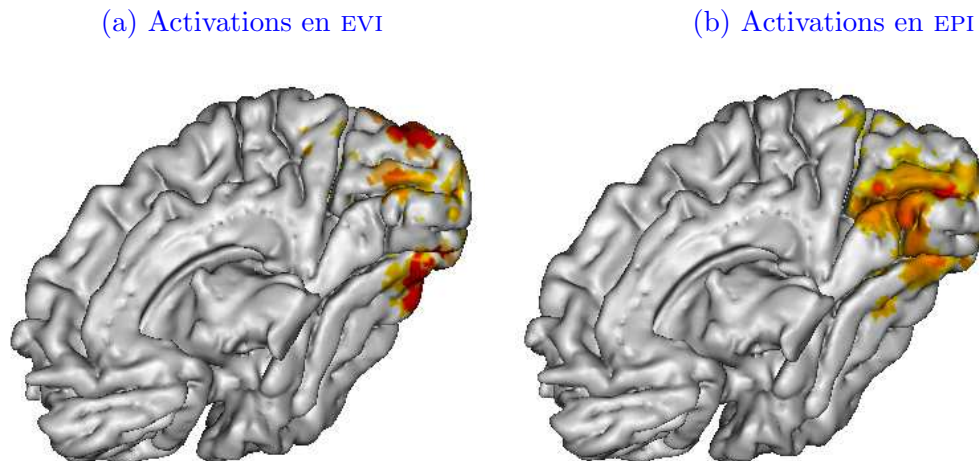
FIG. III.23 – Cartographie des activations détectées en EVI à gauche et en EPI à droite chez le même sujet en réponse à un stimulus visuel, montrant des différences d'extension spatiale. A gauche : superposition des cartes d'activation obtenues pour deux sessions EVI du même sujet ($p < 0.01$, non corrigé). A droite : carte d'activation obtenue pour une session EPI ($p < 0.005$, non corrigé).

FIG. III.24 – Position du volume EVI par rapport à l'anatomie cérébrale d'un sujet dans l'expérience auditive.

III.6.2 Sous-thème 5b – Caractérisation de la dynamique des signaux *Echo Volumar Imaging* (depuis 2005)

Nous avons aussi caractérisé la forme de la réponse hémodynamique à l'aide des outils présentés dans le **Thème 1** et pointé certaines différences reproductibles sur plusieurs sujets comme la présence d'un *undershoot* plus marqué en EVI ou une amélioration du contraste à bruit (doublé par rapport à l'EPI pour le même paradigme expérimental chez le même sujet),
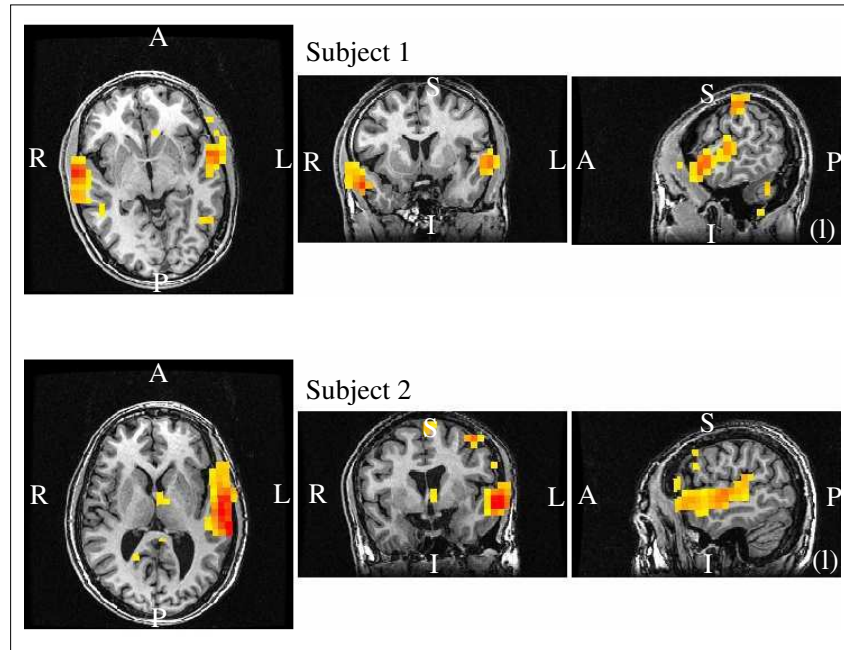
FIG. III.25 – Cartes d'activation superposées à l'anatomie cérébrale de deux sujets, et seuillées à $p = 10^{-4}$, valeur corrigée pour les comparaisons multiples.

comme le démontre la Fig. III.26.

(a) FRHs en EVI

(b) FRHs en EPI
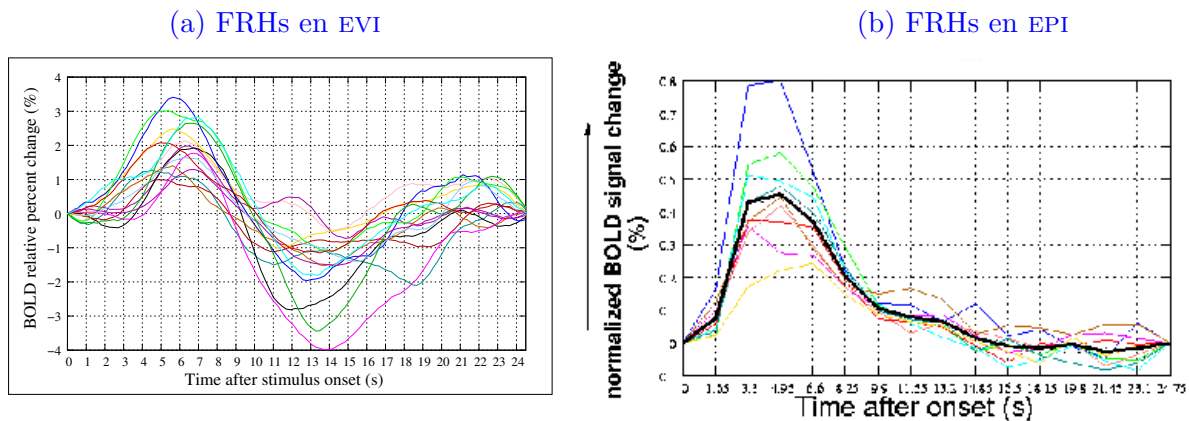
FIG. III.26 – Fonctions de réponse hémodynamique estimées en EVI à gauche et en EPI à droite chez le même sujet et dans plusieurs voxels d'un cluster d'activation visuelle. Les courbes moyennes, calculées sur l'ensemble des voxels, apparaissent en noir.

Nous avons par ailleurs testé et validé dans le système visuel l'hypothèse de stationarité de la réponse hémodynamique estimée à partir de données EVI acquises au cours d'un paradigme expérimental événementiel lent. Pour ce faire, nous avons évalué statistiquement la différence entre les courbes des FRHs estimées sur chaque session à l'aide des tests présentés au **Sous-thème 1d**. Ce travail a fait l'objet d'une présentation orale à la conférence internationale ISMRM en 2006 [c.13].

Enfin, notons que les avantages de cette séquence d'imagerie nous autorisent à envisager comme perspective de recherche l'imagerie néonatale, dans la mesure où les bébés sont des sujets particulièrement mobiles dans le scanner IRM, et qu'il est difficile de les empêcher de bouger.

Cette séquence EVI 3D permettant d'acquérir beaucoup plus vite le signal et ne nécessitant pas de correction du délai d'acquisition inter-coupes réduit ainsi substantiellement la durée des examens. En ce sens, elle est adaptée à l'imagerie du nourrisson. Ce thème a fait l'objet à l'automne 2007 d'un dépôt de sujet de thèse sélectionné pour la campagne de bourses IRTELIS [31] mises au concours au printemps 2008.

---

[31] Programme Doctoral International des Sciences du Vivant du CEA.

III.7 Thème 6 – Analyse exploratoire multifractale des données d'IRMf (depuis 2005)

Tous les travaux précédents s'appuient sur la définition d'un modèle explicatif de la réponse BOLD reliant le paradigme expérimental aux données observées. Les modèles évoqués, bien qu'étant de complexité croissante (linéaire, bilinéaire, non-linéaire/non-stationnaire), reposent toujours sur des hypothèses précises qui peuvent parfois être invalidées sans explication rationnelle *a priori*. Bien sûr, en considérant des modèles génératifs neurophysiologiques du signal BOLD, tels que ceux étudiés dans [BUXTON et coll. 1998, AUBERT et COSTALAT 2002, BUXTON et coll. 2004, AUBERT et coll. 2005], il est possible de relâcher certaines contraintes inhérentes aux modèles présentés dans les sections précédentes.

Toutefois, l'ajustement de ces modèles présuppose la connaissance d'un paradigme d'activation cérébrale comme donnée d'entrée. Ainsi, l'analyse de l'activité cérébrale de fond ou spontanée (acquise au repos [32]) demeure exclue des approches orientées modèle. Dans ces conditions, la seule voie pour analyser ces données est exploratoire. Elle ne repose sur aucune hypothèse *a priori* sur la forme du signal BOLD. En IRMf, comme dans d'autres domaines, les techniques exploratoires classiques telles que l'analyse en composantes principales ou indépendantes (ACP, ACI) sont couramment utilisées [KHERIF et coll. 2002, HIMBERG et coll. 2004, BECKMANN et SMITH 2004]. Ces approches *multivariées* sont à rapprocher des méthodes de *clustering* car elles permettent d'associer des ensembles de voxels à un profil de réponse temporelle particulier. Dans tous les cas, l'association des décours temporels et des régions aux différentes tâches expérimentales du paradigme expérimental se fait *a posteriori*.

Ces méthodes se basent sur des mesures de similarité fonctionnelle, linéaire ou non, des différents signaux acquis. Ces derniers sont analysés « dans leur ensemble », c'est-à-dire que le degré de ressemblance calculé est global et rend compte des séries temporelles complètes. Pour analyser et comparer le degré de régularité locale des signaux IRMf, il est nécessaire d'utiliser d'autres méthodes. Si l'on souhaite quantifier ce degré de régularité à l'aide d'un objet assimilable à un histogramme, il semble naturel de recourir à une décomposition spectrale de chaque signal, et de mesurer sa distribution en fréquence. Toutefois, les données de neuro-imagerie présentent des non-stationnarités temporelles, bien connues en EEG, mais visibles aussi en IRMf. Recourir à ce type de décomposition nécessite de travailler simultanément selon l'axe temporel et l'axe fréquentiel. Pour ses caractéristiques « localisationnistes », nous avons opté pour une décomposition *temps-échelle* plutôt que *temps-fréquence*, autrement dit une base d'ondelettes, capable d'appréhender des non-stationarités ou de refléter un phénomène d'*invariance d'échelle*[33] présent sur les signaux IRMf. Ce travail, encore en cours, a fait l'objet d'une étroite collaboration avec Patrice ABRY (DR CNRS) du laboratoire SiSyPhe de l'ENS de LYON. Des premiers résultats illustrant la modification d'attributs multifractaux entre des données de repos et d'activation ont été publiés dans [c.16]. Une version longue de ce travail a fait l'objet d'une réjection en octobre 2006 à la revue *NeuroImage* [rr.4] ; une version remaniée est présentée au chapitre V. La version anglaise correspondante sera soumise prochainement [s.1].

---

[32]Cette notion de *resting state* est mal définie au sens où elle admet plusieurs acceptions. Nous considérons ici qu'elle représente l'état d'activité basal du sujet accessible par la mesure lorsque celui-ci est allongé dans le scanner, les yeux fermés.

[33]L'invariance d'échelle se traduit par la persistance ou la rémanence d'un phénomène à travers toute une gamme d'échelles au sein d'une analyse multi-résolution.

III.8   Conclusion et perspectives

Le mode de présentation adopté dans les sections précédentes permet de distinguer nettement les thèmes aboutis et les thèmes en développement, et, pour ces derniers, mes préoccupations « prioritaires ». Ces travaux sont à l'origine de thèmes émergeants dans mon activité à savoir :

- la validation au niveau des études de groupes de la méthodologie mise en place pour l'analyse intra-sujet. Le chapitre IV est consacré à ces aspects. Y sont abordés également des perspectives concernant d'une part les analyses à effets mixtes et d'autre part la sélection de modèles génératifs des données IRMf, en tant qu'extensions des travaux déjà entrepris.

- le développement d'approches exploratoires multifractales, *uni-* et *multivariées*, s'affranchissant ainsi de tout paradigme expérimental. Un premier but est notamment de pouvoir attaquer des problématiques d'analyse simultanée de données de repos et d'activation. Un second concerne la mise en évidence, à l'aide de mesures de connectivité, de réseaux fonctionnels pouvant être modulés par l'attention par exemple au cours d'un protocole d'activation. Le développement de telles approches est concomitant à l'avènement de séquences d'imagerie ultra-rapides qui permettent d'acquérir des signaux plus finement échantillonnés, rendant ainsi les analyses multifractales plus fiables. Le chapitre V illustre ces aspects sur la base de travaux préliminaires déjà accomplis.

Deuxième partie

Quelques contributions autour de la dynamique cérébrale en neuro-imagerie fonctionnelle

20 juin 2008

IMPACT DE LA MODÉLISATION NON-PARAMÉTRIQUE DU SIGNAL BOLD SUR LES ANALYSES DE GROUPE

IV.1 Introduction

L'objectif de l'analyse de groupe en IRMf est d'extraire une bonne représentation de la relation entre la structure cérébrale et la fonction à travers une population cible de sujets. Elle consiste habituellemment à moyenner des réponses ou des effets BOLD individuels estimés après une étape de normalisation spatiale qui garantit la définition d'un espace anatomique commun à tous les sujets. Généralement, cette procédure de moyennage est mise en œuvre à l'aide d'une statistique standard de Student et s'appuie sur l'hypothèse que la distribution des effets BOLD, c'est-à-dire de l'activité cérébrale induite par la stimulation, est *normalement* distribuée à travers les sujets. Pour parvenir aux résultats d'une analyse de groupe, il est nécessaire préalablement de traiter les données IRMf acquises chez chaque individu afin de produire une estimée des effets BOLD. Ces effets BOLD sont typiquement fournis par l'ajustement d'un modèle linéaire généralisé. Les approches classiques considèrent une forme canonique, supposée connue et constante à travers tout le cerveau de *fonction de réponse hémodynamique* (FRH), autrement dit de réponse impulsionnelle du système neuro-vasculaire. Récemment, nous avons œuvré au développement d'une alternative reposant sur une modélisation régionale non-paramétrique du signal BOLD [a.8, c.11, c.15]. Cette approche introduite au chapitre III, **Thème 2**, opère une *détection-estimation conjointe* (DEC) de l'activité cérébrale au sens où elle révèle à la fois la localisation spatiale des activations au regard d'une condition ou d'un contraste choisi, et rend compte de leur cinétique. Dans sa version la plus aboutie, elle autorise même le traitement de données non lissées spatialement en tenant compte de la corrélation spatiale inhérente à celles-ci.

La suite de ce chapitre[1] est structurée comme suit. L'analyse intra-sujet classique est rapidement exposée en Section IV.2 avec une attention toute particulière sur la façon dont la souplesse de modélisation peut être envisagée dans ce cadre. En Section IV.3, la modélisation non-paramétrique du signal BOLD que nous promouvons dans un contexte de détection-estimation conjointe est résumée. En particulier, nous mettons en exergue le fait qu'elle s'appuie sur une parcellisation *a priori* – potentiellement multi-sujets — des données IRMf. Cette parcellisation a pour but de faire émerger une échelle de résolution spatiale, les *parcelles*, pertinente pour l'analyse, en agrégeant des voxels voisins sur la base de leur similarité fonctionnelle. Puis, la Section IV.4 est dévolue aux études de groupe. Les principes des analyses à *effets aléatoires* sont rappelés et une attention particulière est consacrée aux approches par test de permutations. Nous illustrons finalement l'impact de cette nouvelle modélisation intra-sujet sur une étude de groupe dans le cadre d'une expérience de calibration fonctionnelle développée par Philippe PINEL et Bertrand THIRION à NEUROSPIN. Enfin, la Section IV.6 élargit le débat en présentant quelques

---

[1]Une version anglaise courte de ce chapitre a été acceptée à la conférence ISBI'08 [**?**], tandis qu'une version longue sera prochainement soumise à la revue *IEEE Trans. on Medical Imaging*.

pistes de recherche méthodologique pertinentes pour l'avenir en vue de rendre les modèles explicatifs plus parcimonieux et capables d'appréhender d'autres sources de variabilité.

IV.2   Analyse intra-sujet classique en IRMf

IV.2.1   Modèle linéaire généralisé

Les méthodes classiques d'analyse des données d'IRMf sont « orientées modèle » au sens où elles s'appuient sur un modèle explicatif du signal BOLD, qui s'instancie dans le cadre des *modèles linéaires généralisés* (MLG). Un tel modèle se définit par une suite de régresseurs, en général tous construits selon le même procédé à partir de deux ingrédients : le paradigme expérimental et la donnée d'une fonction de réponse hémodynamique canonique $\boldsymbol{h}_c$. L'ensemble des MLGs se distinguent donc d'autres approches par la connaissance *a priori* d'une forme de réponse impulsionnelle du système neuro-vasculaire identique pour le cerveau tout entier [FRISTON 1994]. Ainsi, chaque régresseur ou colonne de la matrice de dessin expérimental $\boldsymbol{X}$ dérive de la convolution de $\boldsymbol{h}_c$ avec le signal de stimulation $\boldsymbol{x}^m$ associé à la $m^{\text{ième}}$ condition expérimentale. Le modèle s'écrit donc sous la forme suivante [FRISTON et coll. 1995] :

$$[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_J] = \boldsymbol{X}\,[\boldsymbol{a}_1, \ldots, \boldsymbol{a}_J] + [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_J]\,, \qquad (\text{IV.1})$$

où $\boldsymbol{y}_j$ désigne la série temporelle IRMf mesurée dans le voxel V_j aux instants $(t_n)_{n=1:N}$ et $\boldsymbol{a}_j \in \mathbb{R}^M$ définit le vecteur des effets BOLD en V_j pour toutes les conditions expérimentales $m = 1 : M$. Notons que la matrice $\boldsymbol{X}$ peut être complétée par des régresseurs modélisant des effets de non-intérêt à l'instar d'une base de fonctions en cosinus pour tenir compte de la présence de dérives basse fréquence dans les données. Le bruit $\boldsymbol{b}_j$ est habituellement modélisé comme un processus autorégressif du premier ordre (*i.e.,* AR(1)) afin de tenir compte de l'autocorrélation temporelle, fluctuant dans l'espace, des données IRMf [WORSLEY et coll. 2002] : $b_{j,t_n} = \rho_j\, b_{j,t_{n-1}} + \varepsilon_{j,t_n},\ \forall j,\ t,$ avec $\boldsymbol{\varepsilon}_j \sim \mathcal{N}(\boldsymbol{0}_N, \sigma_{\varepsilon_j}^2 \boldsymbol{I}_N)$. Ici $\boldsymbol{0}_N$ désigne un vecteur de longueur N et $\boldsymbol{I}_N$ définit la matrice identité de taille N. Le bruit $\boldsymbol{b}_j$ reste donc gaussien mais de matrice de covariance égale à $\sigma_{\varepsilon_j}^2 \boldsymbol{\Lambda}_j^{-1}$ où $\boldsymbol{\Lambda}_j$ est tridiagonale symétrique, avec $|\boldsymbol{\Lambda}_j| = 1 - \rho_j^2$, $(\boldsymbol{\Lambda}_j)_{1,1} = (\boldsymbol{\Lambda}_j)_{N,N} = 1$, $(\boldsymbol{\Lambda}_j)_{\ell,\ell} = 1 + \rho_j^2$ et $(\boldsymbol{\Lambda}_j)_{\ell+1,\ell} = -\rho_j$ pour $\ell = 2 : N - 1$.

Les amplitudes ou effets BOLD estimés $\widehat{\boldsymbol{a}}_j$ en V_j de la réponse hémodynamique sont calculés au sens du maximum de vraisemblance par :

$$\widehat{\boldsymbol{a}}_j = \underset{\boldsymbol{a} \in \mathbb{R}^M}{\arg\min}\, (\boldsymbol{y}_j - \boldsymbol{X}\,\boldsymbol{a}_j)^{\text{t}}\, \widehat{\sigma}_{\varepsilon_j}^{-2}\, \widehat{\boldsymbol{\Lambda}}_j\, (\boldsymbol{y}_j - \boldsymbol{X}\,\boldsymbol{a}_j)\,,$$

où $\widehat{\sigma}_{\varepsilon_j}^{-2}\widehat{\boldsymbol{\Lambda}}_j$ définit l'inverse de l'estimée de la matrice d'autocorrelation de $\boldsymbol{b}_j$; voir par exemple [WORSLEY et coll. 2002, PENNY et coll. 2003] pour les détails concernant l'identification de la structure de bruit. Plus récemment, des extensions qui intègrent des informations *a priori* sur $(\boldsymbol{a}_j)_{j=1:J}$ ont été développées dans un contexte bayésien [WOOLRICH et coll. 2004, PENNY et coll. 2005, FLANDIN et PENNY 2007]. Dans ces derniers cas, les vecteurs $(\widehat{\boldsymbol{a}}_j)_{j=1:J}$ sont calculés itérativement en utilisant des procédures plus coûteuses numériquement, qu'elles soient de type variationnelles ou stochastiques. Cependant, aucune de ces contributions fondées sur les MLGs ne remet en cause la définition d'un modèle, seul et unique, d'explication du signal BOLD, en considérant par exemple une forme fluctuante de la FRH, comme le recommandent les auteurs de [HUETTEL et MCCARTHY 2001, HANDWERKER et coll. 2004], au vu de la variabilité inter-régionale mise en évidence entre différentes aires corticales (motrice, visuelle, auditive).

IV.2.2   Gain en flexibilité au sein des MLGs

Bien que plus faible *a priori* que les fluctuations inter-individuelles, la variabilité régionale de la FRH canonique $\boldsymbol{h}_c$ est suffisamment grande pour être prise en compte. Un modèle linéaire généralisé peut en partie s'accomoder de cette variabilité au niveau de chaque voxel, par le biais d'ajout de régresseurs. Par exemple, la réponse impulsionnelle $\boldsymbol{h}_c$ peut être suppléée par l'ajout de ses dérivées première et seconde ($[\boldsymbol{h}_c \,|\, \boldsymbol{h}'_c \,|\, \boldsymbol{h}''_c]$) afin de modéliser des variations de forme vis-à-vis de l'instant du pic d'activation. Bien que puissante et élégante, la souplesse d'un MLG est obtenue au prix d'une perte de sensibilité consécutive à l'ajustement du modèle en vertu de la diminution du nombre de degrés de liberté du test statistique. Par ailleurs, il est important de remarquer que dans un MLG souple, l'effet BOLD attaché à chaque voxel V_j et chaque condition expérimentale m devient multivarié ($\boldsymbol{a}_j^m \in \mathbb{R}^P$). Ainsi, la statistique *univariée* de Student ne devrait plus être utilisée pour tester la significativité de différences potentielles $\boldsymbol{a}_j^m - \boldsymbol{a}_j^n$ entre les $m^{\text{ième}}$ and $n^{\text{ième}}$ types de stimuli. Au contraire, la statistique *multivariée non-signée* de Fisher constitue l'outil d'analyse de choix. Son usage rend les cartes d'activation plus difficiles à interpréter dans la mesure où une part non négligeable de la différence peut être captée par un régresseur autre que celui portant sur l'amplitude, et où l'information de signe est perdue.

IV.3   Analyse intra-sujet non paramétrique en IRMf

IV.3.1   Changement de résolution spatiale

La définition d'un modèle régional de FRH permettrait ainsi d'introduire de la souplesse dans le modèle à moindre coût puisqu'elle conduirait à conserver un *unique* régresseur par parcelle et validerait l'inférence reposant sur la comparaison univariée signée $\widehat{a}_j^m - \widehat{a}_j^n$. Le cadre de détection-estimation conjointe décrit au chapitre III – **Thème 2** se prête à une telle modélisation régionale pourvu qu'une parcellisation du cortex soit disponible. Un « pseudo-MLG » peut alors être ajusté localement aux données au sein de chaque parcelle. Il ne s'agit en effet pas d'un MLG classique dans la mesure où la fonction de réponse hémodynamique est considérée comme inconnue et à estimer, rendant ainsi le modèle *bilinéaire*. Pour des raisons d'efficacité et de parcimonie, nous considérons une *seule* forme de FRH par parcelle. Si cette hypothèse devait être remise en cause, il conviendrait de remettre en question la définition de la parcellisation. Nous reviendrons sur ce point au Chapitre VI du présent chapitre.

Pour définir la bonne résolution spatiale, nous recourons à des techniques de parcellisation telles que celle proposée dans [a.9], où le masque du cerveau est divisé en K *parcelles* selon des critères d'homogénéité fonctionnelle et de connexité spatiale. En d'autres termes, il s'agit de minimiser un critère composite où la mesure de similarité spatiale vise à connecter les voxels voisins dans le système de coordonnées de Talairach tandis que la mesure de similarité fonctionnelle s'appuie sur des paramètres qui résument localement le profil d'activité, par exemple les effets BOLD estimés $\widehat{\boldsymbol{a}}_j = (\widehat{a}_j^m)_{m=1:M}$[2] lors d'une analyse intra-sujet classique décrite en Section IV.2. Il est clair que s'appuyer sur une première analyse à base de MLG pour définir le socle spatial d'une modélisation non-paramétrique du signal BOLD, alors même que l'on souhaite s'en affranchir, peut constituer de prime abord une faiblesse de l'approche préconisée. Nous verrons comment s'affranchir de cette contrainte en Section IV.6.2.

Dans le type de parcellisation utilisée dans la suite, le nombre de parcelles K est fixé empiriquement : plus ce nombre est grand, plus l'homogénéité intra-parcelle est importante, mais plus les parcelles sont petites ce qui peut potentiellement engendrer une diminution du contraste

---

[2]ou les signaux IRMf eux-mêmes.

FIG. IV.1 – Vues sagitales d'une parcellisation multi-sujets. Chaque parcelle est codée par une couleur différente. À gauche, sujet 1 ; à droite : sujet 2.

signal-à-bruit pour l'estimation de la FRH. Pour choisir objectivement un nombre adéquat de parcelles, des critères informationnels tels que le BIC (*Bayesian Information Criterion*) et des techniques de validation croisée ont été utilisés dans [THYREAU et coll. 2006] sur une étude de plusieurs dizaines de sujets. Les auteurs ont montré une convergence pour $K \approx 500$ pour une analyse cerveau entier conduisant à des tailles de parcelle typiquement de quelques centaines de voxels.

Il est par ailleurs important de remarquer qu'il existe une correspondance univoque entre les parcelles des différents sujets puisque la parcellisation est calculée au niveau du groupe. Chaque parcelle a donc une instance chez chacun des sujets de l'analyse, comme illustré par un code de couleurs à la Fig. IV.1. Ceci semble nécessaire si l'on souhaite faire une inférence de groupe au niveau des parcelles, par exemple en moyennant les effets BOLD estimés au sein de chaque parcelle. Si en revanche, la parcellisation demeure un outil purement instrumental afin d'obtenir une estimée spatialement adaptative du filtre hémodynamique régional, il n'est pas rigoureusement nécessaire d'obtenir cette mise en correspondance. Ainsi, des parcellisations intra-sujets peuvent tout autant être utilisées.

IV.3.2 Modélisation « parcellique »du signal BOLD

Nous précisons ici la modélisation régionale non-paramétrique du signal BOLD, introduite dans [a.8, c.15], en vigueur au sein de chaque parcelle. Comme illustré sur la Fig. IV.2, une seule forme de FRH $\boldsymbol{h}$ est considérée par parcelle, tout en autorisant une modulation de l'amplitude de la réponse a_j^{m}[3] à travers l'espace (indice j) et les conditions expérimentales (indice m). Dans la suite, on note $\mathcal{P} = (V_j)_{j=1:J}$ une parcelle constituée d'un ensemble de voxels connexes. Le modèle du signal BOLD dans $\mathcal{P}$ s'écrit :

$$\boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}^m \boldsymbol{h} + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j, \quad \forall j,\ V_j \in \mathcal{P}. \tag{IV.2}$$

où cette fois, $\boldsymbol{X}^m$ est une matrice binaire de dimensions $N \times (D + 1)$ qui code les onsets du $m^{\text{ième}}$ type de stimuli. Le vecteur $\boldsymbol{h} \in \mathbb{R}^{D+1}$ représente la forme inconnue de la FRH dans $\mathcal{P}$.

---

[3]aussi appelée niveau de réponse neurale (NRN) parce que modélisable comme l'entrée du système dont le filtre est la FRH.

20 juin 2008

Le terme $\boldsymbol{P}\boldsymbol{\ell}_j$ modélise la dérive basse fréquence qui permet de prendre en compte le repliement spectral des artéfacts physiologiques (battements cardiaques, respiration), échantillonnés trop lentement pour respecter les conditions de SHANNON-NYQUIST. Le terme $\boldsymbol{b}_j \sim \mathcal{N}(\boldsymbol{0}_N, \sigma_{\varepsilon_j}^2 \boldsymbol{\Lambda}_j^{-1})$ rend compte du bruit d'acquisition et des erreurs de modélisation et comme précédemment est modélisé comme une réalisation d'un processus AR(1).

Il est intéressant de remarquer qu'en agglomérant les différents vecteurs $\left[\boldsymbol{X}^1\boldsymbol{h}, \ldots, \boldsymbol{X}^M\boldsymbol{h} \right]$ dans (IV.2) et en unifiant la matrice résultante avec $\boldsymbol{P}$ on retrouve la matrice $\boldsymbol{X}$ décrite en Eq. (IV.1).
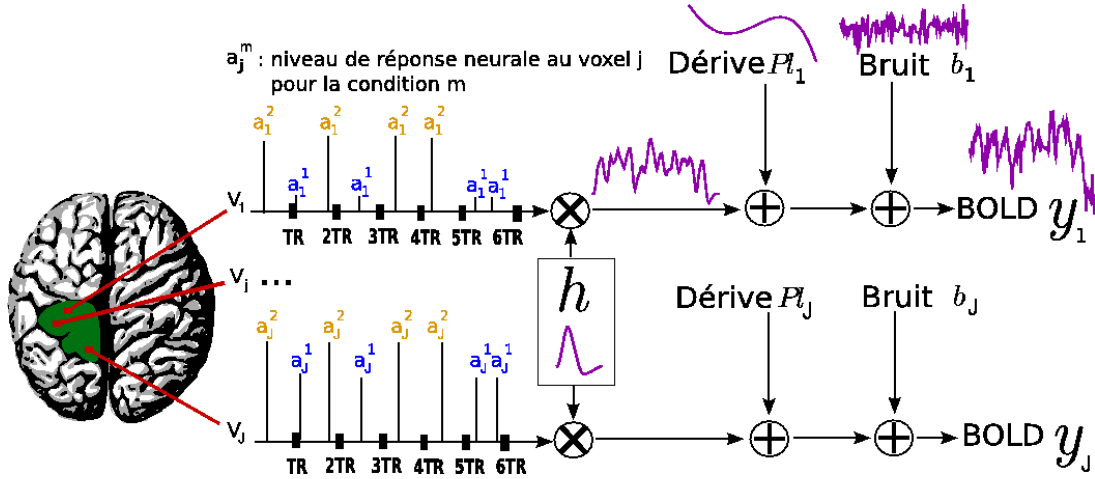
FIG. IV.2 – Modèle régional du signal BOLD dans le formalisme DEC. Les niveaux de réponse neurale (NRNs) sont notés a_j^m et varient d'un voxel à l'autre et d'une condition expérimentale à une autre, tandis que la FRH $\boldsymbol{h}$ caractérisant la parcelle $\mathcal{P}$ est de forme unique.

Ici, nous négligeons la corrélation spatiale du bruit considérant qu'il est plus important de modéliser celle du signal d'intérêt induit par la stimulation. Les signaux $\mathbb{y} = (\boldsymbol{y}_j)_{j=1:J}$ sont donc supposés indépendants dans l'espace :

$$p(\mathbb{y} \mid \boldsymbol{h}, \mathbb{a}, \mathbb{1}, \boldsymbol{\theta}_0) = \prod_{j=1}^{J} p(\boldsymbol{y}_j \mid \boldsymbol{h}, \boldsymbol{a}_j, \boldsymbol{\ell}_j, \rho_j, \sigma_{\varepsilon_j}^2) \tag{IV.3}$$

$$\propto \prod_{j=1}^{J} |\boldsymbol{\Lambda}_j|^{1/2} \, \sigma_{\varepsilon_j}^{-N} \exp\Big(-\sum_{j=1}^{J} \frac{\tilde{\boldsymbol{y}}_j^{\mathrm{t}} \boldsymbol{\Lambda}_j \tilde{\boldsymbol{y}}_j}{2\sigma_{\varepsilon_j}^2}\Big)$$

où $\boldsymbol{\theta}_0 = (\rho_j, \sigma_{\varepsilon_j}^2)_{j=1:J}$ et $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h} - \boldsymbol{P}\boldsymbol{\ell}_j = \boldsymbol{y}_j - = \boldsymbol{X} \left[\boldsymbol{a}_j \mid \boldsymbol{\ell}_j\right]^{\mathrm{t}}$.

IV.3.3   Inférence bayésienne

Informations *a priori*

Fonction de réponse hémodynamique.   La forme $\boldsymbol{h}$ et les Niveaux de Réponse Neuronale (NRNs) ou effets BOLD $(\boldsymbol{a}_j)_{j=1:J}$ sont conjointement estimés au sein de chaque parcelle $\mathcal{P}$. Puisqu'aucun modèle paramétrique n'est considéré pour $\boldsymbol{h}$, depuis [a.6, a.3] nous préconisons l'usage d'une contrainte de régularité temporelle sur la dérivée seconde de $\boldsymbol{h}$ afin de régulariser son estimation.

Modèles de mélanges spatiaux (MMS).   Le formalisme DEC a aussi pour vocation de *détecter* des activations c'est-à-dire d'identifier au sein de chaque parcelle $\mathcal{P}$ les voxels dont

le décours temporel est corrélé au paradigme expérimental. Dans ce but, nous avons introduit dans [c.15] une modélisation *a priori* sur les NRNs $(\boldsymbol{a}^m)_{m=1:M}$ que nous rappelons brièvement ici.

Nous supposons que différents types de stimuli induisent des NRNs ou des effets BOLD statistiquement indépendants *i.e.*, $p(\mathbb{a} \,|\, \boldsymbol{\theta_a}) = \prod_m p(\boldsymbol{a}^m \,|\, \boldsymbol{\theta}^m)$ avec $\mathbb{a} = (\boldsymbol{a}^m)_{m=1:M}$, $\boldsymbol{a}^m = (a_j^m)_{j=1:J}$ et $\boldsymbol{\theta_a} = (\boldsymbol{\theta}^m)_{m=1:M}$. $\boldsymbol{\theta}^m$ contient l'ensemble des hyper-paramètres inconnus reliés au $m^{\text{ème}}$ stimulus. Nous définissons un modèle de mélange spatial en introduisant des variables binaires indicatrices q_j^m ou *étiquettes*, qui établissent si un voxel $V_j \in \mathcal{P}$ s'active ($q_j^m = 1$) ou non ($q_j^m = 0$) en réponse au stimulus de type m. Conditionnellement à ces étiquettes, les NRNs $\boldsymbol{a}^m$ sont distribués spatialement selon une loi normale dont les paramètres fluctuent selon la parcelle : $(a_j^m \,|\, q_j^m = i) \sim \mathcal{N}(\mu_{i,m}, v_{i,m})$, avec $i = 0, 1$. Nous imposons $\mu_{0,m} = 0$ pour la classe des voxels non-activés, de ce fait : $\boldsymbol{\theta}^m = [v_{0,m}, \mu_{1,m}, v_{1,m}]$.

Contrairement à [a.8], nous introduisons des variables q_j^m corrélées spatialement entre elles impliquant $\Pr(q_j^m = 1) = \lambda_{1,j}^m$ au lieu de $\Pr(q_j^m = 1) = \lambda_1^m$, $\forall m$. Le modèle de corrélation retenu est un champ de MARKOV binaire sur les variables q_j^m, alors que les NRNs restent indépendants entre eux conditionnellement à $\boldsymbol{q}^m$. Plus précisément, un champ de ISING *symétrique* est introduit :

$$\Pr(\boldsymbol{q}^m \,|\, \beta^m) = Z(\beta^m)^{-1} \exp\big(\beta^m \sum_{j \in \mathcal{P}} \sum_{k \in \mathrm{N}_j} \omega_{jk} I(q_j^m = q_k^m)\big),$$
$$\propto \exp\big(\beta^m \sum_{j \sim k} \omega_{jk} I(q_j^m = q_k^m)\big), \tag{IV.4}$$

où $I(A) = 1$ si A est vraie et $I(A) = 0$ sinon. La notation N_j définit le voisinage de V_j tandis que celle plus compacte, $j \sim k$, a la même signification que la double somme mentionnée, c'est pourquoi nous la privilégierons dans la suite. Les paramètres ω_{jk} sont des constantes prédéfinies qui pondèrent les interactions entre voxels (V_j, V_k) : $\omega_{jk} \propto 1/d(V_j, V_k)$ où $d(V_j, V_k)$ est la distance entre V_j et V_k. Le paramètre $\beta^m > 0$ dans $\Pr(q_j^m \,|\, \beta^m)$ contrôle le niveau de corrélation spatiale. Ici, il est fixé empiriquement. La constante de normalisation du champ $Z(\beta^m)$, encore appelée *fonction de partition* dépend de β^m. Notons que ce champ est *caché* car $\mathbb{q} = (\boldsymbol{q}^m)_{m=1:M}$ n'apparaît pas dans (IV.2). En combinant ces informations, nous obtenons un modèle de mélange spatial pour chaque type de stimulus :

$$p(\boldsymbol{a}^m \,|\, \boldsymbol{\theta}^m) = \sum_{\boldsymbol{q}^m} \Big(\prod_{j=1}^J p(a_j^m \,|\, q_j^m, \boldsymbol{\theta}^m)\Big) \Pr(\boldsymbol{q}^m \,|\, \beta^m). \tag{IV.5}$$

Loi *a posteriori* conjointe

A l'aide de la règle de Bayes et des autres *a priori* définis dans [a.8], la loi *a posteriori* $p(\boldsymbol{h}, \mathbb{q}, \mathbb{a}, (\boldsymbol{\ell}_j), \boldsymbol{\Theta} \,|\, \mathbb{y})$ s'écrit :

$$p(\boldsymbol{h}, \mathbb{a}, \mathbb{l}, \boldsymbol{\Theta} \,|\, \mathbb{y}) \propto \sigma_{\boldsymbol{h}}^{-D} \sigma_{\ell}^{-JQ} \prod_{j=1}^J \Big(\frac{(1 - \rho_j^2)^{1/2}}{\sigma_{\varepsilon_j}^{N+1}} \, \mathbb{1}_{(-1,1)}(\rho_j)\Big)$$
$$\times \exp\Big(-\frac{\boldsymbol{h}^{\mathrm{t}} \boldsymbol{R}^{-1} \boldsymbol{h}}{2\sigma_{\boldsymbol{h}}^2} - \sum_{j=1}^J \Big(\frac{1}{2\sigma_{\varepsilon_j}^2} \tilde{\boldsymbol{y}}_j^{\mathrm{t}} \boldsymbol{\Lambda}_j \tilde{\boldsymbol{y}}_j + \frac{1}{2\sigma_{\ell}^2} \|\boldsymbol{\ell}_j\|^2\Big)\Big) \prod_{m=1}^M \Big(p(\boldsymbol{\theta}^m) p(\boldsymbol{a}^m \,|\, \boldsymbol{\theta}^m)\Big) \tag{IV.6}$$

avec $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \boldsymbol{S}_j \boldsymbol{h}$ et $\boldsymbol{S}_j = \sum_j a_j^m \boldsymbol{X}^m$.

Nous simulons des réalisations de (IV.6) par échantillonnage de Gibbs pour approcher des estimateurs de type espérance *a posteriori* $\widehat{\mathbb{q}}^{\mathrm{MP}}$, $\widehat{\mathbb{a}}^{\mathrm{MP}}$ et $\widehat{\boldsymbol{h}}^{\mathrm{MP}}$. Nous déduisons ainsi à travers

$\widehat{\mathbb{q}}^{\text{MP}}$ quels sont les voxels activés, et $(\widehat{a}_j^m)^{\text{MP}}\widehat{\boldsymbol{h}}^{\text{MP}}$ estime la réponse hémodynamique du voxel j pour le stimulus m. Ces différents estimateurs sont donnés par $\widehat{x}^{\text{MP}} = (K-I)^{-1}\sum_{k=I+1}^{K} x^{(k)}$, $\forall\, x \in \{\boldsymbol{h}, \mathbb{a}, \mathbb{q}, \boldsymbol{\Theta}\}$ où I est la longueur de la *période de chauffe* de la chaîne de MARKOV. Nous ne présenterons dans la suite que la formulation de la loi conditionnelle *a posteriori* de $(\mathbb{a}, \mathbb{q})$. Le lecteur pourra se référer à [a.8] pour les autres quantités d'intérêt.

D'un point de vue de l'analyse de la convergence du schéma MCMC et du taux de mélange des chaînes de Markov générées, nous avons considéré les critères usuels basés sur l'échantillonnage parallèle, qui consiste à simuler plusieurs chaînes en parallèles, en partant d'initialisation différentes, puis à mesurer la variabilité intra- et inter-chaînes sur différentes variables d'intérêt et finalement à en déduire un scalaire, défini à partir du rapport des deux, qui tend vers un lorsque la convergence est atteinte [GELMAN et coll. 2004].

Mélange spatial *a posteriori*

La distribution *a priori* sur les NRNs considérée ici étant un mélange gaussien, et l'expression de la vraisemblance à $\boldsymbol{h}$ fixée demeurant gaussienne, la densité conditionelle *a posteriori* des NRNs est aussi un mélange gaussien par un argument de conjugaison. De (IV.6), on peut montrer que chaque a_j^m est une réalisation de :

$$p(a_j^m \mid \boldsymbol{y}_j, \boldsymbol{h}, \boldsymbol{\theta}^m, \epsilon_j^2, a_j^{m'\neq m}, q_{k\in\mathrm{N}_j}^m) = \sum_{i=0,1} \lambda_{i,j}^m \mathcal{N}(\mu_{i,j}^m, v_{i,j}^m),$$

en posant $\mathrm{N}_j = \{V_k \mid k \sim j\}$. La simulation de ces variables se décompose alors en trois étapes : (i) identifier les paramètres $(\lambda_{i,j}^m, \mu_{i,j}^m, v_{i,j}^m)$; (ii) échantillonner le label binaire q_j^m en accord avec $\lambda_{i,j}^m$ et finalement (iii) simuler a_j^m contionnellement à q_j^m selon $\mathcal{N}(\mu_{i,j}^m, v_{i,j}^m)$ pour $i = 0, 1$; *cf.* [c.11] :

$$v_{i,j}^m = \left(v_{i,m}^{-1} + \frac{\boldsymbol{g}_m^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{g}_m}{\sigma_{\varepsilon_j}^2}\right)^{-1} , \; \mu_{i,j}^m = v_{i,j}^m\left(\frac{\boldsymbol{g}_m^{\mathrm{t}}\boldsymbol{\Lambda}_j\boldsymbol{e}_{m,j}}{\sigma_{\varepsilon_j}^2} + i\,\frac{\mu_{i,m}}{v_{i,m}}\right)$$

où $\boldsymbol{g}_m = \boldsymbol{X}^m\boldsymbol{h}$ et $\boldsymbol{e}_{m,j} = \boldsymbol{y}_j - \boldsymbol{P}\boldsymbol{\ell}_j - \sum_{m'\neq m} a_j^{m'}\boldsymbol{g}_{m'}$. La probabilité $\lambda_{i,j}^m$ de l'événement $(q_j^m = i)$ s'écrit :

$$\lambda_{i,j}^m = \left(1 + \frac{r_{1-i,j}^m}{r_{i,j}^m}\frac{\pi_{1-i,j}^m}{\pi_{i,j}^m}\right)^{-1} \tag{IV.7}$$

avec $r_{i,j}^m = \left(v_{i,j}^m/v_{i,m}\right)^{1/2}\exp\left((\mu_{i,j}^m)^2/v_{i,j}^m - i(\mu_i^m)^2/v_i^m\right)$ et $\pi_{i,j}^m = \mathrm{Pr}(q_j^m = i \mid q_{k\in\mathrm{N}_j}^m, \beta^m)$. Pour calculer (IV.7), nous devons évaluer :

$$\pi_{1-i,j}^m/\pi_{i,j}^m = \exp\left(2\beta^m(2i-1)\sum_{k\in\mathrm{N}_j} w_{jk}(2q_k^m - 1)\right),$$

qui ne dépend que des étiquettes dans le voisinage N_j de V_j.

Pour l'instant, le paramètre β^m est fixé empiriquement, rendant ainsi la régularisation spatiale supervisée. Son estimation nécessite l'approximation de la fonction de partition du champ de Ising $Z(\beta^m)$ en amont de la boucle d'échantillonnage, qui pourra être précalculée et sauvegardée dans une table. Nous discutons ce point en guise de perspective méthodologique dans la dernière section de ce chapitre.

À titre d'information, puisque nous considérons les données des différentes parcelles indépendantes entre elles, l'inférence au sein de chaque parcelle est parallélisable. Ainsi, une analyse cerveau entier prend environ une heure pour des données de longueur $N = 125$ et une parcellisation de $K = 500$ parcelles sur une machine quadri-processeurs à double cœur.

<div align="right">20 juin 2008</div>

IV.4   Analyses de groupe

IV.4.1   Inférence paramétrique classique

On considère en général que les S sujets ayant subi l'expérience d'IRMf ont été tirés au hasard au sein d'une population d'intérêt. Comme nous l'avons vu dans les sections précédentes, les deux types d'analyse intra-sujet produisent en un voxel particulier V_j de l'*espace normalisé* — habituellement fourni par le *template*[4] du MNI — et pour chaque sujet s, un vecteur d'effets BOLD estimés $\widehat{\boldsymbol{a}}_{j,s}$ ou de NRNs selon la terminologie employée. La comparaison de conditions expérimentales passe habituellement par la définition de *contrastes*. Compte tenu des efforts consentis à la Section IV.3 pour préserver des *effets* BOLD *scalaires*, nous pouvons nous restreindre à des contrastes unidimensionnels c'est-à-dire à des différences signées de la forme $\widehat{d}_{j,s}^{m-n} = \widehat{a}_{j,s}^m - \widehat{a}_{j,s}^n$ pour distinguer les effets BOLD estimés pour les $m^{\text{ième}}$ et $n^{\text{ième}}$ conditions. Dans la suite, pour alléger les notations, l'indice j des voxels et celui du contraste considéré, *i.e.*, $m - n$, sont supprimés.

Alors que la différence estimée $\widehat{d}_s$ diffère généralement de la valeur inconnue du contraste d_s non observé, nous supposons dans un premier temps une estimation parfaite à l'échelle individuelle c'est-à-dire : $\widehat{d}_s = d_s \quad \forall s = 1 : S$. Le problème se résume donc ainsi. Étant donné un échantillon $(d_1, \ldots, d_S)$ tiré aléatoirement selon une densité de probabilité inconnue $f(d)$, qui décrit la distribution des effets d'intérêt dans la population, nous cherchons à inférer sur un paramètre scalaire de localisation de cette distribution (*e.g.*, moyenne, médiane, mode, ...). Supposons pour l'instant que nous souhaitons tester l'hypothèse *nulle* que la *moyenne* de la population est négative :

$$H_0 : \qquad \mu_{\text{G}} = \int d\, f(d)\, dd \leq 0$$

où l'indice $_{\text{G}}$ rappelle la quantité rattachée au groupe. Dans ce but, nous pouvons avoir recours au test t de Student. Le calcul de la statistique de Student s'écrit :

$$t = \frac{\hat{\mu}_{\text{G}}}{\hat{\sigma}_{\text{G}}/\sqrt{S}}, \text{ avec : } \quad \hat{\mu}_{\text{G}} = \frac{\sum_s d_s}{S}, \quad \hat{\sigma}_{\text{G}}^2 = \frac{\sum_s (d_s - \hat{\mu}_{\text{G}})^2}{S-1} \qquad \text{(IV.8)}$$

Dès lors, nous rejetons H_0, et donc nous acceptons l'hypothèse *alternative* $H_1 : \mu_{\text{G}} > 0$ si la probabilité sous H_0 d'atteindre la valeur de t observée est inférieure à un taux de faux positifs[5] fixé à l'avance. Sous l'hypothèse que la distribution $f(d)$ est gaussienne, il est bien connu que cette probabilité s'obtient à l'aide de la distribution de Student à $S - 1$ degrés de liberté. Dans ce contexte paramétrique, il est possible de démontrer que la statistique t de Student est de sensibilité statistique optimale (techniquement, au sens du test non-biaisé uniformément le plus puissant, voir par exemple [GOOD 2005]).

IV.4.2   Populations non-gaussiennes

En revanche, si l'hypothèse gaussienne n'est pas réaliste, la distribution de Student ne demeure valide que dans la limite des grands échantillons (*i.e.*, grosse cohorte de sujets) et peut donc s'avérer fournir un contrôle inexact du taux de faux positifs pour des tailles d'échantillons usuelles en IRMf (dix à vingt sujets). Ce problème peut être contourné à l'aide de schémas de calibration non-paramétrique tels que celui fourni par les tests de permutations, qui autorise une inférence exacte sous des hypothèses plus faibles, en l'occurrence la symétrie de la densité

---

[4]En français, on parlerait de gabarit mais cette traduction n'est pas usitée.

[5]On dit aussi taux de fausses alarmes ou d'erreur de première espèce, consistant à déclarer comme significative une différence qui en réalité ne l'est pas.

$f(d)$ autour de zéro. Il est donc recommandé d'utiliser des tests de permutations. Toutefois, ces tests ne fournissent qu'une stratégie alternative au seuillage d'une carte statistique donnée, et à ce titre, résolvent un problème de *spécificité*.

Mais le fait que la distribution d'échantillonnage $f(d)$ puisse s'écarter de la *normalité* pose aussi une question de *sensibilité* puisque la statistique de Student peut s'avérer sous-optimale dès lors que l'hypothèse de normalité n'est plus vérifiée. Sans connaissance *a priori* sur la forme de $f(d)$, un choix par défaut raisonnable pour le test statistique consiste à maintenir de bonnes performances de détection sur une large gamme de densités possibles. Une telle statistique doit donc s'avérer robuste, non pas au sens classique en étant *résistante* à des données atypiques ou aberrantes, mais plutôt dans le sens d'être compatible avec des distributions qui tendent à produire de telles données (*e.g.,* distributions à queue lourde ou multimodales). Dans la suite, nous utiliserons la statistique du rang signé de Wilcoxon (RSW) qui s'avère fournir un bon compromis sensibilité-spécificité entre des versions plus robustes comme la statistique du signe et d'autres qui le sont moins comme celle de Student. La statistique RSW se calcule comme suit :

1. classer en ordre croissant les effets *absolus* ;
2. sommer les rangs modulés par les signes des effets correspondants produisant

$$T_{\mathrm{rsw}} = \sum_{s=1}^{S} \mathrm{rang}(|d_s|) \times \mathrm{sign}(d_s). \qquad (\text{IV.9})$$

Remarquons que comme la statistique du signe, la statistique T_{rsw} est indépendante des données. Ainsi la distribution marginale de T_{rsw} calculée par permutations peut être estimée à l'avance.

D'autres choix de statistiques sont évidemment possibles notamment celle reposant sur le rapport de vraisemblance empirique. Le lecteur intéressé pourra consulter [Mériaux et coll. 2006, Roche et coll. 2007] pour de plus amples détails.

IV.4.3   Problème de comparaisons multiples

Dans le contexte de l'IRMf, il est important de remarquer qu'un test est effectué en chaque voxel du volume de recherche. Il est alors calculatoirement efficace de seuiller la carte statistique ainsi produite à l'aide d'une valeur uniforme, qui peut être réglée de façon à contrôler le taux de faux positifs à un niveau fixé par l'expérimentateur. Le seuillage uniforme produit une bonne puissance de détection pourvu que la distribution de la statistique du test soit raisonnablement stationnaire dans l'espace. Dans le cas de tests spatialement indépendants, la valeur de ce seuil uniforme est obtenue par la correction de Bonferroni, qui consiste à diviser le taux fixé par l'expérimentateur par le nombre de tests réalisés dans le cerveau. En pratique, cette solution fournit des résultats trop conservatifs et une sensibilité de détection trop faible. De plus, en IRMf, les tests sont spatialement corrélés à la fois par l'effet du lissage spatial réalisé sur les données mais aussi par la corrélation anatomo-fonctionnelle réellement existante, par exemple au sein de chaque gyrus. La résolution du problème des comparaisons multiples par une correction moins sévère s'avère donc nécessaire.

La théorie des champs aléatoires développée principalement par Keith Worsley dans le contexte de la neuro-imagerie [Worsley 1994], a permis de pallier cette difficulté en autorisant d'autres types de corrections, basées sur des mesures topologiques d'une carte statistique, vue comme la réalisation d'un champ aléatoire paramétrique (par exemple gaussien). Ces corrections intègrent donc naturellement la prise en compte de la corrélation spatiale du champ mais en pratique, elles ne fournissent des bonnes approximations du taux de faux positifs et donc des p-valeurs exploitables qu'au delà d'une certaine régularité spatiale du champ nécessitant un

lissage important des données. De plus, ces techniques de correction supposent la stationnarité du champ aléatoire.

Les travaux menés ces dernières années au laboratoire par Alexis ROCHE et ses étudiants, Sébastien MÉRIAUX et Merlin KELLER, ont permis de résoudre aisément, grâce aux tests de permutations le problème des comparaisons multiples sous des hypothèses d'échangeabilité multivariée d'une façon beaucoup plus puissante que le font les tests paramétriques et les approximations sous-tendant la théorie des champs aléatoires [WORSLEY 1994]. Ainsi, dans la suite de [HOLMES et coll. 1996, BULLMORE et coll. 1999, NICHOLS ET HOLMES 2002, HAYASAKA ET NICHOLS 2003], ils ont fourni des p-valeurs corrigées pour le taux d'erreur de première espèce en calibrant la *statistique du maximum* sur le volume de recherche. De la même façon, ils sont capables de rapporter des p-valeurs corrigées au niveau cluster après seuillage de la distribution, obtenue par permutations, de la statistique de l'extension maximale d'un cluster. Dans la suite, les résultats présentés en Section IV.5 sur les données expérimentales ont été obtenus grâce à leurs travaux, synthétisés pour l'essentiel dans [MÉRIAUX et coll. 2006, ROCHE et coll. 2007] pour les analyses à effets aléatoires. La combinaison de nos analyses intra-sujet et des analyses à effets mixtes, qu'ils ont aussi développées, sont évoquées en perspectives à la fin de ce chapitre.

IV.5 Illustrations sur des données de calibration fonctionnelle

Les données d'IRMf ont été collectées lors d'une expérience de calibration fonctionnelle, systématiquement réalisée chez tous les sujets au SHFJ puis maintenant à NEUROSPIN. Ici, un groupe d'une quinzaine de sujets a été extrait et les données sont issues du scanner à 1.5 T (Signa GE) du SHFJ. Pour chaque sujet, une seule session d'acquisition comportant $N = 125$ volumes fonctionnels est disponible. Le temps de répétition ou délai inter-acquisition est de $TR = 2, 4$ sec conduisant à une durée d'examen de 5 mn. La résolution spatiale des volumes est de 3 mm dans les trois directions. Le but principal de cette expérience est de cartographier très rapidement à la fois des aires sensori-motrices (aires visuelles, auditives, motrices) mais aussi des régions associées à des fonctions cognitives plus évoluées telles que le calcul ou la lecture. Dans la suite on s'intéresse uniquement à un seul contraste, c'est-à-dire à la comparaison des conditions expérimentales *auditive et visuelle*, référencé A−V.

IV.5.1 Comparaison des deux analyses intra-sujet

Tout d'abord, nous comparons les effets BOLD estimés par les deux types d'analyse intra-sujet présentés dans les sections IV.2-IV.3, respectivement. Ces effets ont été calculés à partir des données non-lissées spatialement. Certains pourraient s'étonner d'une comparaison sur données non-lissées. Toutefois, dans le cadre d'une analyse individuelle, eu égard aux efforts consentis par les physiciens pour améliorer la résolution spatiale des images tout en conservant un RSB acceptable, il apparaît inopportun de détruire une partie de l'information contenue dans les données. Pour ma part, le filtrage spatial n'est qu'une façon de masquer la variabilité inter-individuelle en étendant spatialement les activations de chaque sujet et en favorisant ainsi le recouvrement des activations à travers les sujets. Ce lissage spatial peut donc se justifier pour les analyses de groupe mais perd de son sens pour une analyse individuelle.

La Fig. IV.3 souligne clairement pour le contraste A−V que le formalisme DEC atteint une meilleure sensiblité de détection (avec des activations bilatérales) par rapport à l'inférence de type MLG. En effet, les effets BOLD $\widehat{d}_j^{\mathrm{A.-V.}}$ ont des valeurs plus élevées sur la Fig. IV.3(b) et apparaissent plus contrastés. Deux principales raisons expliquent ces différences. La première

concerne la modélisation de la corrélation spatiale à l'aide de modèles de mélanges spatiaux dans le formalisme DEC alors que l'approche MLG traite de façon univariée chaque voxel. La seconde, plus intéressante, repose sur une différence chez le sujet considéré entre sa réponse impulsionnelle neuro-vasculaire et la fonction de réponse hémodynamique canonique $\boldsymbol{h}_c$, utilisée dans la définition du MLG. La Fig. IV.3(c)-[rouge] démontre que la dynamique cérébrale estimée $\widehat{\boldsymbol{h}}_{\mathcal{P}}$ sur la parcelle associée à l'activation la plus forte s'écarte notablement de la forme canonique $\boldsymbol{h}_c$ rappelée en Fig. IV.3(c)[vert]. En ce sens, elle conforte la deuxième hypothèse. Ce résultat intéressant et reproductible sur d'autres sujets démontre évidemment que la modélisation de la forme de la réponse impacte fortement l'amplitude de l'effet estimé, autrement dit que ces deux quantités sont inter-dépendantes. Même si l'on peut douter du double rebond en fin de décours, le seul fait de bien estimer le pic d'activation avec un délai retardé de 2 sec. permet de mieux ajuster le modèle aux données et de restaurer ainsi une carte d'effets BOLD plus contrastée.

La comparaison des cartes de variances associées à ces effets est aussi très pertinente (résultats non montrés). Dans le cadre du MLG, on montre que l'erreur commise sur l'effet peut s'avérer corrélée à sa taille : plus l'activation est forte, plus son incertitude est grande ! cette situation peut surprendre mais se produit fréquemment en pratique dès lors que les variables explicatives du signal IRMf introduites dans le modèle ont du mal à capter les fluctuations liées à l'activation. Dans ce cas, une partie de l'effet BOLD est modélisée par le terme de bruit, dont la variance se met à croître inconsidérément.
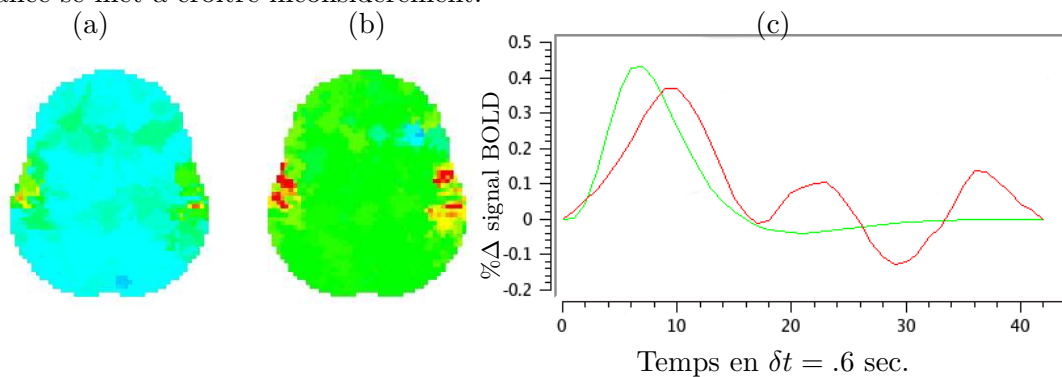
FIG. IV.3 – Cartes de contraste BOLD estimé $(\widehat{d}_j^{\mathrm{A.-V\cdot}})_j$ chez un sujet donné entre les conditions auditive et visuelle. **(a)** : résultats obtenus à l'aide du MLG et de la FRH canonique $\boldsymbol{h}_c$. **(b)** : résultats obtenus par l'approche de détection-estimation conjointe, *i.e.,* le modèle (IV.2). **(c)** : Comparaison des formes de FRH estimées dans la parcelle la plus activée $\mathcal{P}$: $\boldsymbol{h}_c$ et $\widehat{\boldsymbol{h}}_{\mathcal{P}}$ apparaissent respectivement en vert et rouge. Ecart entre les pics d'activations d'environ 1,2 sec.

IV.5.2   Analyses à effets aléatoires

Pour asseoir l'impact de notre comparaison au niveau du groupe vis-à-vis des chaînes de traitements classiques en IRMf (*e.g.,* SPM, FSL), les données utilisées pour ajuster le modèle (IV.1) ont été filtrées spatialement à l'aide d'un lissage gaussien isotrope à 6 mm. Dans le formalisme DEC en revanche, nous considérons toujours les images *non lissées* mais normalisées afin de travailler dans un référentiel commun à tous les sujets. La Fig. IV.4 précise le cheminement des deux types d'analyse et rappelle que notre approche s'appuie sur une parcellisation anatomo-fonctionnelle des données. Ici, nous avons considéré une parcellisation multi-sujets de façon à garantir pour chaque parcelle une instance chez chacun des sujets. Toutefois, ceci pourra être remis en cause à l'avenir (voir Section IV.6.2). En effet, les analyses de groupe menées ici sont massivement univariées et donc réalisées dans chaque voxel, au contraire des inférences de

groupe parcelliques présentées dans [a.11]. A titre d'information, les deux chaînes de traitements présentées à la Fig. IV.4 seront disponibles très prochainement dans la nouvelle version de la plate-forme logicielle `BrainVisa` (http://brainvisa.info) en mars 2008.
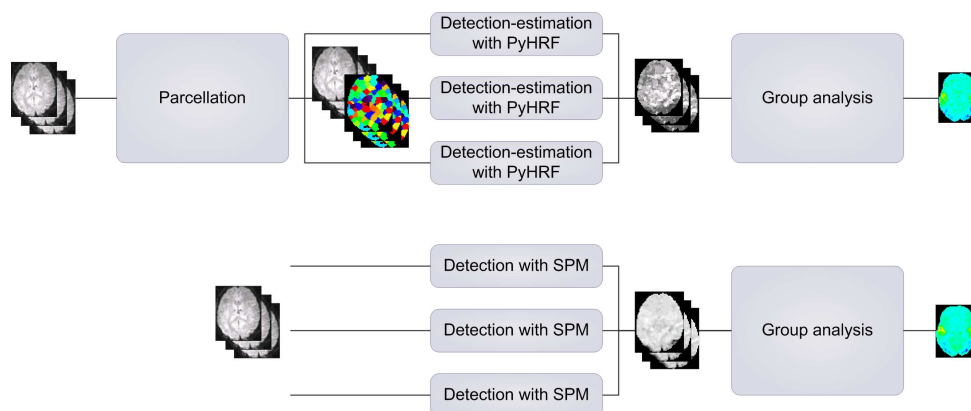
FIG. IV.4 – Chaînes de traitements associées aux deux types d'analyse intra-sujet. Au dessus : l'approche s'appuyant sur une parcellisation multi-sujets, le formalisme DEC et une analyse de groupe non-paramétrique par tests de permutation. En dessous : l'approche classique reposant sur un MLG et la même analyse de groupe qu'au dessus.

La Fig. IV.5 fournit les cartes statistiques d'analyse à effets aléatoires obtenues à l'aide de la statistique de Wilcoxon T_{rsw}, corrigées pour les comparaisons multiples dans le contexte des tests de permutations. Les coupes représentées coïncident avec les activations les plus significatives. Les cartes statistiques dérivées au niveau du groupe à l'aide du formalisme DEC démontrent une bonne sensibilité malgré l'absence de lissage, et même par endroit une meilleure sensibilité sans doute due à une estimée plus précise du contraste BOLD en raison de la modélisation non-paramétrique et spatialement adaptative de la FRH ; cette différence est prégnante notamment dans l'hémisphère gauche où des activations dans l'aire de Broca (en haut à droite sur la Fig. IV.5(a)) ont été rapportées de façon significative à l'aide du formalisme DEC. Concernant les activations bilatérales massives, les deux méthodes se différencient peu sur le plan de la significativité statistique. La Table IV.1 confirme quantitativement ces résultats. Elle souligne aussi que les foyers d'activation rapportés sur les Figs. IV.5(a-c-e) par notre approche sont d'étendue plus petite que ceux présentés sur les Figs. IV.5(b-d-f) obtenus avec l'approche MLG. Ceci n'est qu'une conséquence du lissage introduit sur les données dans le cadre de l'analyse de groupe réalisée à partir des effets estimés par MLG.

TAB. IV.1 – Foyers significatifs d'activation pour la statistique de décision T_{rsw}.

	Niveau cluster p_{corr}	Taille cluster (voxels)	Niveau voxel p_{corr}	Coords. pic x	y	z
DEC	**0.002**	**1151**	$1e-06$	8	30	26
	0.003	876	0.0007	47	27	30
SPM	0.0022	1788	0.0001	5	29	28
	0.0028	**1680**	0.0001	45	27	27

La Fig. IV.6 illustre pour les quatre sujets les plus homogènes parmi les quinze les FRHs identifiées dans la parcelle associée à l'activation maximale dans l'étude de groupe, située dans le lobe temporal supérieur gauche (première ligne de la Table IV.1). Cette parcelle ne coïncide pas avec celle rapportée pour les résultats présentés en analyse intra-sujet. On observe ici une
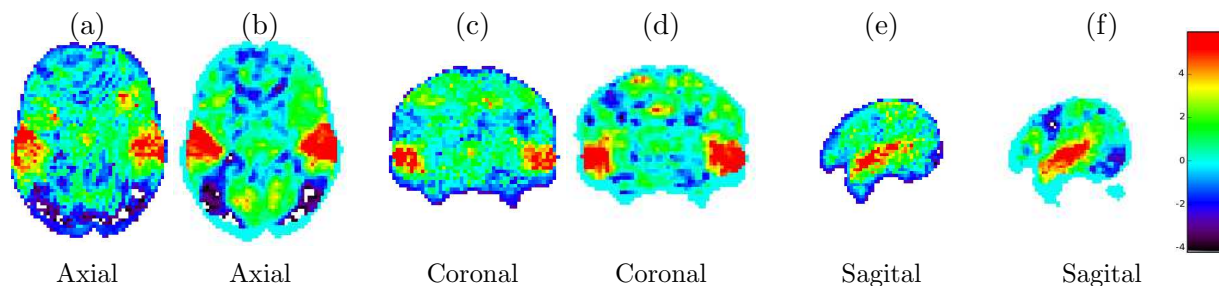
FIG. IV.5 – Résultats d'une analyse à effets aléatoires s'appuyant sur la statistique du rang signé de Wilcoxon. Les coupes correspondent aux activations les plus significatives et sont orientées selon la convention radiologique (la gauche est à droite). les résultats **(a)-(c)-(e)** ont été obtenus dans le cadre du formalisme DEC tandis que ceux présentés en **(b)-(d)-(f)** dérivent d'une analyse SPM standard au niveau intra-sujet.

variabilité évidente de l'instant du pic d'activation (jusqu'à 1,5 sec.) tout comme des différences dans la déplétion consécutive à l'activation (*final undershoot*).
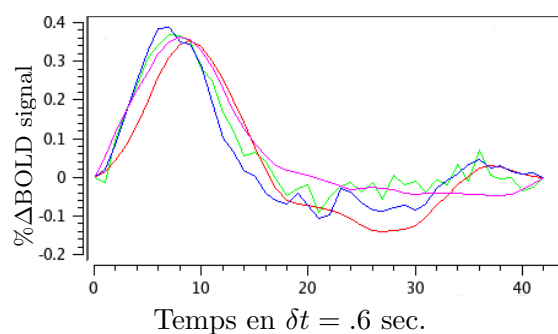
FIG. IV.6 – Fonctions de réponse hémodynamique estimées pour différents sujets (une couleur par sujet) dans la parcelle correspondant à l'activation la plus significative, c'est-à-dire à la statistique T_{rsw} maximale.

IV.5.3   Conclusion

Dans ce chapitre, nous avons mis en évidence que les statistiques de groupe en IRMf sont influencées par le type d'analyse conduites au niveau de chaque individu. En particulier, nous avons démontré que le formalisme de détection-estimation conjointe est capable de fournir au neuroscientifique des résultats d'analyse à *effets aléatoires* pertinents et robustes qui peuvent même s'avérer plus sensibles dans certaines situations sans dégrader la résolution spatiale originale des données IRMf (pas de lissage spatial). L'interprétation de ces différences est imputable semble-t-il à la variabilité spatiale de la réponse impulsionnelle du système neuro-vasculaire. Des travaux futurs s'attacheront à valider davantage ces résultats notamment en travaillant sur des données de nouveaux nés chez qui il est désormais connu que la forme de ce filtre hémodynamique est différente de celle mesurable chez l'adulte [D'ESPOSITO et coll. 1999, JACOBS et coll. 2008].

Par ailleurs, ces travaux ont rendu possible et réalisable l'analyse de la variabilité inter-individuelle de la dynamique cérébrale au cœur des foyers d'activation par le biais d'une modélisation non-paramétrique appropriée du signal BOLD. Nous allons maintenant dans la section ultime de ce chapitre évoquer les nombreuses perspectives de ce travail.

IV.6 Problèmes ouverts en guise de perspectives

Les perspectives discutées ici sont uniquement d'ordre méthodologique et visent à l'amélioration des méthodes existantes.

IV.6.1 Régularisation spatiale adaptative

Avant d'envisager un réglage automatique du paramètre de corrélation spatiale, nous avons commencé par vérifier la criticité du réglage de ce paramètre, c'est-à-dire en quel sens la valeur retenue conditionne la sensibilité et la spécificité de détection (voir [d.7] pour les détails). En substance, nous avons constaté une bonne robustesse de détection uniquement sur un intervalle restreint de valeurs de ce paramètre : $\beta^m \in [0,3\,;0,5]\,, \forall m$. Il semble donc difficile d'échapper au développement de cette extension non-supervisée.

Par souci de simplicité, on considère dans un premier temps $\beta^m = \beta$, $\forall m$. Afin de régler automatiquement ce paramètre du champ de ISING défini pour une parcelle fixée $\mathcal{P}$, il est nécessaire d'estimer la constante de normalisation de ce champ, en vue d'ajouter dans l'échantillonneur de Gibbs global une étape de MÉTROPOLIS-HASTINGS sur le paramètre β vis-à-vis de ce paramètre. En effet, la simulation selon la densité conditionnelle *a posteriori* de β au sein de $\mathcal{P}$ fait intevenir la loi

$$p(\beta \mid \boldsymbol{q}^1, \ldots, \boldsymbol{q}^M) \propto p(\boldsymbol{q}^1, \ldots, \boldsymbol{q}^M \mid \beta)\,p(\beta) \propto \left[\prod_m \Pr(\boldsymbol{q}^m \mid \beta) \right] p(\beta)$$

$$\propto Z_{\mathcal{P}}(\beta)^{-M} \prod_m \exp(\beta U_{\mathcal{P}}(\boldsymbol{q}^m))\,p(\beta)$$

$$\text{avec} \quad U_{\mathcal{P}}(\boldsymbol{q}) = \sum_{j \sim k, j \in \mathcal{P}} \omega_{jk} I(q_j = q_k). \tag{IV.10}$$

Pour parvenir à l'expression (IV.10), on a utilisé le fait que $Z_{\mathcal{P}}(\beta)$ est indépendant de la condition expérimentale m quelle que soit la parcelle considérée $\mathcal{P}$. En revanche, $Z_{\mathcal{P}}(\beta)$ est fonction de la géométrie de la parcelle et notamment de sa taille. Ainsi, la valeur estimée de $Z_{\mathcal{P}}(\beta)$ pourra évoluer selon $\mathcal{P}$ et conduire potentiellement à des valeurs de β fluctuant à travers les parcelles. Le paramètre β étant positif, nous pourrons considérer par exemple $p(\beta) = \boldsymbol{I}_{\mathbb{R}_+}(\beta)$ ou tout autre loi à support sur $\mathbb{R}_+$.

D'après (IV.10), la simulation de la distribution $p(\beta \mid \boldsymbol{q}^1, \ldots, \boldsymbol{q}^M)$ requiert donc la connaissance de la fonction de partition $Z_{\mathcal{P}}(\beta)$ du champ, qui s'écrit :

$$Z_{\mathcal{P}}(\beta) = \sum_{\boldsymbol{q} \in \{0,1\}^J} \exp(\beta U_{\mathcal{P}}(\boldsymbol{q})). \tag{IV.11}$$

L'échantillonnage *a posteriori* de β autorisera donc la prise en compte de non-stationarités spatiales dans la carte d'activations.

L'approximation de la fonction de partition (IV.11) peut être envisagée selon une approche d'*intégration thermodynamique* [GELMAN ET MENG 1998, HIGDON et coll. 1997] rappelée ci-dessous. Notons tout d'abord que $Z_{\mathcal{P}}(\beta)$ est une fonction décroissante de β telle que $Z(0) = 2^J$

pour un champ de ISING si Card $[\mathcal{P}] = J$[6]. Il suffit ensuite d'écrire

$$
\begin{aligned}
\frac{dZ_{\mathcal{P}}(\beta)}{d\beta} &= \sum_{\boldsymbol{q} \in \{0,1\}^J} U_{\mathcal{P}}(\boldsymbol{q}) \exp(\beta U_{\mathcal{P}}(\boldsymbol{q})) \\
&= Z_{\mathcal{P}}(\beta) \sum_{\boldsymbol{q} \in \{0,1\}^J} U_{\mathcal{P}}(\boldsymbol{q}) \frac{\exp(\beta U_{\mathcal{P}}(\boldsymbol{q}))}{Z_{\mathcal{P}}(\beta)} \\
&= Z_{\mathcal{P}}(\beta) \, \mathrm{E}[U_{\mathcal{P}}(\boldsymbol{Q} = \boldsymbol{q} \,|\, \beta)]
\end{aligned}
$$

d'où l'on déduit :

$$
\frac{d \ln Z_{\mathcal{P}}(\beta)}{d\beta} = \mathrm{E}[U_{\mathcal{P}}(\boldsymbol{Q} = \boldsymbol{q} \,|\, \beta)].
$$

Par conséquent, on peut écrire :

$$
\ln Z_{\mathcal{P}}(\beta) = J \ln 2 + \int_0^\beta \mathrm{E}[U_{\mathcal{P}}(\boldsymbol{Q} = \boldsymbol{q} \,|\, \beta') \, \mathrm{d}\beta']. \tag{IV.12}
$$

L'équation (IV.12) est la base d'une méthode relativement directe — dans son principe — pour estimer $\ln Z_{\mathcal{P}}(\beta)$. L'espérance mathématique peut être remplacée par une moyenne empirique

$$
\mathrm{E}[U_{\mathcal{P}}(\boldsymbol{Q} = \boldsymbol{q} \,|\, \beta)] = \frac{1}{K} \sum_{k=1}^K U_{\mathcal{P}}(\boldsymbol{q}_\beta^{(k)}) \tag{IV.13}
$$

où les K échantillons $\boldsymbol{q}_\beta^{(k)}$ sont des réalisations du champ de Ising à la tempétature $1/\beta$. Toute la qualité de l'approximation (IV.13) dépend de notre capacité à simuler ces champs de Ising pour des valeurs de paramètres croissantes : $\beta \in \{\beta_1, \beta_2, , \ldots, \beta_G\}$. En particulier, une question sensible concerne la façon efficace de générer ces réalisations pour des valeurs élevées de β correspondant à des configurations spatialement très corrélées. L'échantillonnage efficace de tels champs dans le cadre général des champs de POTTS soumis à un champ extérieur est une question qui n'a connu des avancées théoriques et pratiques que très récemment [CHANG et coll. 2004, LUCZAK ET VIGODA 2005]. Pour les champs de ISING symétriques, la procédure la plus connue et la plus efficace est celle de SWENDSEN-WANG [SWENDSEN ET WANG 1987], qui procède par bloc. Ces échantillonneurs peuvent être encore accélérée grâce à des stratégies de *Simulation tempérées* [GEYER ET THOMPSON 1995, BROOKS et coll. 2006] plus efficaces qu'un simple schéma MCMC pour « sauter » d'un minimum local à un autre correspondant tous deux à la même énergie.

Dans le but de densifier la grille discrète de valeurs de β *i.e.,* $\{\beta_1, \beta_2, , \ldots, \beta_G\}$ sur laquelle est estimée $Z_{\mathcal{P}}(\beta)$ sans augmenter significativement le coût de calcul, il est aussi possible d'utiliser la stratégie suivante :

$$
\forall p \in \mathbb{N}_P^*, \forall \epsilon \in \mathcal{B}(0, r), \beta_g' = \beta_g + \epsilon \,|\, \{\boldsymbol{q}^{(k)}\}_{k=1}^K \sim \Pr(\boldsymbol{q}; \beta_g), \frac{Z_{\mathcal{P}}(\beta_g')}{Z_{\mathcal{P}}(\beta_g)} \approx \frac{\sum_{k=1}^K \exp(\beta_g' U_{\mathcal{P}}(\boldsymbol{q}^{(k)}))}{\sum_{k=1}^K \exp(\beta_g U_{\mathcal{P}}(\boldsymbol{q}^{(k)}))}. \tag{IV.14}
$$

Sinon, il est toujours envisageable de recourir à des techniques d'interpolation classique.

La stratégie décrite ci-dessus pourra aussi être utilisée pour l'échantillonnage de la loi *a posteriori* jointe vis-à-vis de β. En effet, l'étape MÉTROPOLIS-HASTINGS concernant le paramètre β

---

[6]Pour la généralisation aux champs de POTTS, on a $Z(0) = C^J$ si C désigne le nombre d'étiquettes possibles que peut prendre chaque variable d'état q_j.

pourra être définie comme une marche aléatoire exploitant la relation (IV.14) dont la probabilité d'acceptation sera exactement $f(\beta')/f(\beta) = Z_{\mathcal{P}}(\beta')/Z_{\mathcal{P}}(\beta)$. Toutefois, il est important de noter que cette approche n'est pas stable numériquement si $r > 0$ s'avère trop grand : il faut vraiment travailler au voisinage des points de la grille, c'est-à-dire pour r suffisamment petit.

Compte tenu de la dépendance de β vis-à-vis de $\mathcal{P}$, un compromis réaliste devra être trouvé pour éviter l'estimation de cette fonction sur toutes les régions. A partir d'une table de valeurs précalculées pour des configurations extrêmale, on pourra procéder par extrapolation de la valeur de la fonction de partition $Z(\beta)$ en fonction de la taille de la parcelle, c'est-à-dire du nombre de cliques et/ou de sites présents au sein de chaque parcelle.

Enfin, il sera aussi envisageable de rendre le champ de ISING *a priori* informé anatomiquement en considérant comme dans [SMITH et coll. 2003, OU ET GOLLAND 2005] une énergie externe c'est-à-dire un terme privilégiant un état parmi les deux en fonction de critères anatomiques, issus par exemple de la segmentation matière grise/matière blanche/liquide céphalorachidien obtenue à partir de l'IRM anatomique du sujet. Pour ce faire, il suffira d'introduire une composante *séparable* qui ne tient compte que des cliques *singletons*.

IV.6.2   Choix de la parcellisation : comment la faire évoluer ?

Notre préoccupation concerne la possibilité d'améliorer la robustesse de nos analyses vis-à-vis du choix de la parcellisation. Il s'agit d'abord de s'affranchir d'une parcellisation *a priori* calculée sur la base d'effets BOLD estimés par l'ajustement d'un MLG. En effet, comme démontré dans [**?**] par une analyse de sensibilité, le choix de la parcellisation demeure critique au sens où l'utilisation d'une parcellisation « aléatoire », comprenant des parcelles peu homogènes du point de vue du signal fonctionnel, engendre une variabilité importante de l'estimation des effets BOLD et de leur incertitude. De plus, il serait aussi intéressant de limiter le nombre de parcelles de non-intérêt c'est-à-dire dans lesquelles peu ou pas d'effet BOLD n'est détecté, ceci afin de limiter les problèmes d'identifiabilité des paramètres et de réduire la dimension du problème. Il est donc nécessaire de choisir au mieux cette parcellisation en recherchant une solution évolutive et plus parcimonieuse.

Dans l'approche retenue jusqu'à maintenant, la parcellisation était définie à partir d'un mélange de distributions gaussiennes multivariées, tenant compte à la fois des trois dimensions de l'espace (coordonnées des voxels) mais aussi de dimensions fonctionnelles (données par les effets BOLD estimés). Il semble souhaitable à l'avenir de remettre en cause la définition de cette segmentation, pour d'une part réduire le nombre de parcelles de non-intérêt mais aussi parce que spatialement, rien ne garantit que la configuration optimale des parcelles ne suive une loi normale.

Dans le futur, j'envisage une démarche *adaptative* dans laquelle la parcellisation pourra évoluer au cours de l'algorithme de détection-estimation conjointe selon des critères de similarité des réponses impulsionnelles estimées dans des parcelles connexes. L'idée sera de mettre en œuvre une approche de type *clustering* agglomératif, d'abord sous forme supervisée et déterministe selon des critères de distance bien définis, puis dans un cadre probabiliste en tant qu'inférence sur un modèle. Dans ce cas, il sera envisageable de recourir à des algorithmes stochastiques à sauts reversibles qui procèderont par fusion ou dissociation de parcelles existantes au delà d'une certaine taille fixée par exemple [SAINT PIERRE 2003]. Une alternative pourra aussi être envisagée dans le cadre des statistiques bayésiennes non-paramétriques (*e.g.,* mélanges de processues de DIRICHLET, MPD) [FERGUSON 1973, ANTONIAK 1974], en considérant le processus de parcellisation comme un problème de segmentation non-supervisée, *i.e.,* où le nombre de classes reste à déterminer. Dans ce cadre, pour garantir la connexité des classes *i.e.,* des voxels de même

étiquette, le couplage d'un MPD à un champ de Markov sur les étiquettes peut s'avérer fort judicieux [Orbanz et Buhmann 2008].

Ces travaux feront l'objet au printemps 2008 du stage de master recherche d'Anne-Laure Fouque. Ils entraîneront évidemment un coût de calcul supplémentaire mais les efforts algorithmiques que nous avons consentis ces dernières années permettront d'exploiter une mise en œuvre en partie parallélisable sur une ferme de calculateurs. A terme, l'idée est de définir les parcelles comme des entités différentes du fonctionnement cérébral, qui pourront être contraintes spatialement par des frontières anatomiques bien définies telles que des sillons étiquetés [Rivière et coll. 2002]. Ces travaux seront donc menés en étroite collaboration avec Bertrand Thirion[7] ainsi qu'avec les autres membres du lnao spécialiste d'anatomie cérébrale.

IV.6.3 Sélection de modèles génératifs

En vue de franchir une étape importante dans la phase d'« analyse intra-sujet » des données irmf et afin de déterminer au sein d'une région cérébrale fixée le meilleur modèle explicatif des données au regard des conditions expérimentales définies dans le paradigme par le neuro-psychologue, je présente ici quelques pistes de recherche. Dans les approches classiques présentées en Section IV.2-IV.3, l'ensemble des conditions expérimentales manipulées dans le paradigme est introduit dans un modèle linéaire ou bilinéaire, en prenant soin de tenir compte des temps d'arrivée de chaque type de stimulation pour différencier ces conditions. Aucun choix n'a donc été opéré sur le bon nombre de conditions à modéliser pour expliquer des activations régionales spécifiques si ce n'est un choix *maximal* par défaut. Dans l'approche classique par MLG, la pertinence de la modélisation est évaluée *a posteriori* en contrastant différentes conditions voire différents modèles à l'aide de tests statistiques judicieusement choisis (tests de Student ou de Fisher,...). Dans le formalisme DEC, nous envisageons le problème comme celui de la recherche d'une dimension au sein d'une famille emboîtée.

Vers une famille de modèles ...

La poursuite de ces travaux cherchera à s'écarter de ce paradigme d'analyse conventionnelle. Comme dans d'autres domaines, il paraît judicieux d'envisager des modèles plus parcimonieux vis-à-vis du choix optimal du nombre de conditions à introduire dans le modèle, ou vis-à-vis de la modélisation de certaines non-linéarités ou non-stationarités de la réponse bold. En particulier, si l'on considère une approche régionale, il semble clair que toutes les conditions expérimentales ne vont pas jouer le même rôle dans une région donnée. Il semble donc naturel non seulement de rechercher le nombre adéquat de conditions à introduire dans chaque modèle régional, mais il paraît aussi pertinent de s'interroger sur le choix ou la sélection des bons types de stimuli (meilleure combinaison possible) à dimension de modèle fixée.

Par ailleurs, le modèle explicatif des données bold peut évoluer à travers une dynamique différente selon les régions (variabilité de la réponse impulsionnelle [Hernandez et coll. 2002], variabilité des non-linéarités [Soltysik et coll. 2004]). Là encore, comme nous l'avions déjà pressenti dans [c.10], certaines régions présenteront par exemple des réponses très variables à travers les essais successifs d'un même stimulus tandis que d'autres seront d'une reproductibilité sans faille.

Pour réaliser cette sélection, nous nous placerons dans le cadre bayésien et nous envisagerons deux approches : une *off-line* basée sur le calcul de cotes *a posteriori* ou de facteurs de Bayes

---

[7]chercheur inria et responsable de la jeune équipe Parietal à Neurospin, principal investigateur des techniques de parcellisations à Neurospin.

pour comparer deux modèles déjà ajustés aux données, l'autre *on-line*, plus complexe, visant à identifier de façon non-supervisée le bon modèle en même temps que ses paramètres.

Exemple : Choix des conditions expérimentales

Dire que seul un sous-ensemble des stimulations liées au paradigme expérimental induisent des activations cérébrales dans une parcelle donnée revient à considérer que pour certaines valeurs de m (*e.g.*, condition auditive dans une parcelle $\mathcal{P}$ visuelle), l'ensemble des états cachés $\mathbb{q}^m$ peut satisfaire : $q_j^m = 0, \forall j \in \mathcal{P}$. Dans ce cas, il apparaît inutile et même sous-optimal (principe de parcimonie ou rasoir d'OKAM [MACKAY 2003]) de modéliser la condition m. Outre ces aspects, dans le formalisme de détection-estimation conjointe décrit dans ce chapitre, citons aussi le problème « plus technique et mécanistique » de l'échantillonnage des hyper-paramètres qui en cas d'« inactivation totale » nécessite de recourir à des hyper-*a priori* propres sur la moyenne et la variance de la classe d'activation afin de ne pas risquer de dégénérescence de la vraisemblance *a posteriori*. En d'autres termes, on est facilement confronté à des problèmes d'*identifiabilité* de paramètres et d'hyper-paramètres lorsque l'on ne se s'appuie pas sur le principe de parcimonie[8].

C'est la raison pour laquelle il est pertinent de considérer le choix du bon nombre de conditions à modéliser $C \in \{1, \ldots, M\}$, et à la détermination de la meilleure combinaison de conditions $\{m_1, \ldots, m_C\} \in \mathbb{N}_C^*$ une fois C fixé. Il s'agit donc d'un problème de sélection de modèle au sein d'une famille emboîtée de modèles

$$\mathcal{M} = \bigcup_{c=2}^{M} \{c\} \bigotimes_{m_1, \ldots, m_c} M_{\boldsymbol{m}}^c,$$

$$M_{\boldsymbol{m}}^c \triangleq \left\{ \exists (m_1, \ldots, m_c) \in (\mathbb{N}_M^*)^c \,|\, \forall n \in \{m_1, \ldots, m_c\} \, \exists j \in \mathbb{N}_J^*, \, q_j^n \neq 0 \right\}, \qquad \text{(IV.15)}$$

à savoir, quelle est la combinaison la plus vraisemblable de conditions $\boldsymbol{m}$ de dimension fixée c compte tenu des données observées $\mathbb{y}$ pour l'ensemble des voxels de la parcelle $\mathcal{P}$. Nous supposons donc ici que tous les voxels de la parcelle courante répondent au même modèle défini par d conditions $\{m_1, \ldots, m_c\}$:

$$\boldsymbol{y}_j = \sum_{n=m_1}^{m_c} a_j^n \boldsymbol{X}^n \boldsymbol{h} + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j, \quad \forall V_j \in \mathcal{P} \qquad \text{(IV.16)}$$

Ce problème est donc double puisqu'il s'apparente d'abord à la recherche de la dimension du modèle n, puis à l'identification du modèle $M_{\boldsymbol{m}}^c$ c'est-à-dire de la meilleure combinaison possible de conditions $\boldsymbol{m}$ de dimension c. Remarquons que $M_{\boldsymbol{m}}^c$ peut aussi être défini par complémentation à partir de (IV.15) :

$$M_{\boldsymbol{m}}^c = \mathbb{N}_M^* \setminus \left\{ n \in \mathbb{N}_M^* \,|\, n \not\ni \boldsymbol{m} \Rightarrow q_j^n = 0, \, \forall j \in \mathcal{P} \right\}.$$

Cette formulation permet peut-être de mieux faire ressortir le fait que tous les voxels associés à une condition m non répertoriée dans l'ensemble $\boldsymbol{c}$ du modèle $M_{\boldsymbol{m}}^c$ sont inactivés dans $\mathcal{P}$.

Au delà d'un exemple

Ce cadre de travail de la sélection de modèles dépasse largement le seul exemple de l'exhibition des conditions expérimentales pertinentes pour la détection des activations et l'estimation de leur dynamique. Il autorise aussi la comparaison de modèles non-linéaires et/ou non-sationaires

---

[8]On parle aussi d'*overfitting* dans la littérature de l'apprentissage statistique [SCHÖLKOPF ET SMOLA 2002].

tels que ceux proposés dans [rr.3, p.9], qui visent à appréhender des phénomènes neuronaux tels que l'habituation à la répétition de stimuli ou comprendre la nature de la variabilité inter-essais de la réponse hémodynamique en relâchant l'hypothèse de repdroductibilité à travers les essais pour les voxels générant une activation.

Aspects algorithmiques

Dans la communauté, la problématique de sélection de modèles n'a été abordée jusqu'à maintenant qu'au travers de l'utilisation de critères informationnels (Akaike – AIC, BIC,DIC, ...) [Akaike 1974, Schwartz 1978, Speigelhalter et coll. 2002, Woolrich et coll. 2004], dont on connaît les limites, ou par le biais de comparaisons d'*évidence* de modèles calculées à l'aide de distances de Kullback et d'approximations du type *Variational Bayes* [Beal 2003, Penny et coll. 2003, Penny et coll. 2004]. Toutefois, dans le formalisme DEC, la bilinéarité du modèle (IV.16) rend plus imprécises les techniques d'approximation variationnelle. En effet, la bilinéarité rend impossible l'intégration analytique conjointe vis-à-vis des effets BOLD et de la réponse impulsionnelle. La seule façon de contourner cette diffculté consiste à supposer une indépendance conditionnelle *a posteriori* de la taille et de la forme des effets, ce qui peut sembler irréaliste d'un point de vue physiologique.

Récemment, une technique de sélection du nombre de conditions par comparaison de fonctions de vraisemblance estimées par échantillonnage pondéré, a été proposée dans [Donnet et coll. 2006] : par rapport aux deux approches envisagées, il s'agit d'une méthode *off-line*. Bien que séduisante, cette approche souffre d'une limitation intrinsèque, qu'on pourrait qualifier d'« instabilité par emboîtement ». En effet, il n'est pas exclu dans ce type de méthodes que lorsque la dimension de la famille de modèles évolue (passage de c à $c + 1$), des conditions différentes soient sélectionnées. Pour gérer ce types de difficultés, les approches de type Monte Carlo par chaînes de Markov (MCMC) à *sauts réversibles* constituent un outil d'investigation de choix [Green 1995, Andrieu et Doucet 1999, Richardson et Green 1997, Robert et coll. 2000, Green et Richardson 2002] dans la mesure où elles fournissent les éléments pour naviguer à travers des modèles de dimension différente tout en expliquant par des processus bien définis les mécanismes d'apparition, de suppression ou d'échanges de conditions au sein du modèle. Toutefois, la convergence de ces algorithmes dépend fortement du bon choix de lois de proposition permettant d'explorer au mieux l'espace constitué des différents modèles, comme illustré dans [Brooks et coll. 2003].

IV.6.4 Analyses à effets mixtes

A court terme, nous souhaitons dépasser le cadre des analyses à effets aléatoires classiques pour nous concentrer sur des analyses à effets mixtes afin de prendre également en compte l'incertitude sur les effets estimés. L'idée que l'on souhaite tester est d'évaluer le gain réalisé sur l'estimation de cette incertitude en passant d'une analyse intra-sujet par MLG à une analyse de type detection-estimation conjointe où l'on autorise une variabilité régionale de la fonction de réponse hémodynamique. Evidemment, on s'attend à gagner en sensibilité chez des populations où l'on sait clairement que la forme du filtre hémodynamique s'écarte de la version canonique, que ce soit chez des bébés ou des populations de malades, par exemple atteints d'épilepsie [Bï¿½ar et coll. 2003, Gotman et coll. 2004, Jacobs et coll. 2008].

En effet, nous avons vu en Section IV.4 que les analyses à effets aléatoires considèrent l'estimation des effets BOLD comme parfaite en ce sens que la variance de l'estimateur des effets $\widehat{d_s}$ au niveau de chaque individu est négligée. Dans les analyses à effets mixtes, il s'agit toujours d'établir une inférence sur le paramètre global de la population, par exemple la moyenne ou la

médiane à partir des effets estimés mais en tenant compte du caractère imparfait et potentiellement inhomogène à travers les sujets de l'estimation. Il est en effet courant, pour des questions de mouvement par exemple, que l'effet mesuré chez certains sujets soit beaucoup plus incertain que chez d'autres. Classiquement dans ce cadre [WOOLRICH et coll. 2004, ROCHE et coll. 2007], le paramètre inconnu est relié aux observations par le modèle hiérarchique linéaire à deux niveaux suivant :

Premier niveau intra-sujet

Pour chaque sujet, le contraste d'effets estimés $\widehat{d}_s$ est relié au vrai contraste inconnu recherché à travers un modèle linéaire conditionnellement gaussien :

$$\forall s, \qquad \widehat{d}_s | d_s \sim g_s(\widehat{d}_s) = \frac{1}{\sqrt{2\pi r_s}} e^{-\frac{(\widehat{d}_s - d_s)^2}{2r_s}} \tag{IV.17}$$

De façon équivalente, cela revient à écrire $\widehat{d}_s = d_s + \varepsilon_s$, où ε_s est un bruit blanc gaussien indépendant de d_s de variance r_s. Le bruit est dit *homoscédastique* lorsque r_s est constant à travers les sujets et *hétéroscédastique* sinon. En pratique, r_s est estimée à partir des résidus du MLG dans le cas d'une estimation classique (*cf.* Section IV.2) ou au sens de la moyenne *a posteriori* si l'on adopte le formalisme DEC présenté en Section IV.3. Dans le premier cas, le modèle gaussien néglige la précision d'estimation sur r_s ou de façon équivalente assigne un nombre de degrés de liberté infini à chaque sujet. Dans le second, cette contrainte peut être relachée, puisque la variance est probabilisée et son estimateur bayésien est disponible. Dans la littérature classique du domaine, certains auteurs recommandent de remettre en cause l'hypothèse d'un bruit gaussien et considèrent par exemple une distribution de Student [WOOLRICH et coll. 2004, FRISTON et coll. 2005]. Toutefois, cette hypothèse induit un coût de calcul beaucoup plus élevé. Par ailleurs, le fait de supposer un bruit gaussien ne contamine pas la spécificité du test lorsque celui-ci est calibré par permutations ; le seul impact porte sur sa sensibilité.

Second niveau inter-sujet

La distribution des vrais effets dans la population est modélisée comme une distribution $f \in \mathcal{F}$, où $\mathcal{F}$ est une famille :

$$\forall s, \qquad d_s \sim f(d_s) \tag{IV.18}$$

Les premiers travaux menés en neuro-imagerie sur les analyses à effets mixtes ont restreint $\mathcal{F}$ à la famille paramétrique gaussienne [WORSLEY et coll. 2002, BECKMANN et coll. 2003, MÉRIAUX et coll. 2006]. Sous l'angle des modèles linéaires hiérarchiques, cela correspond à définir au niveau inter-sujet la relation

$$d_s = d_{\mathrm{G}} + \epsilon_s \quad \text{avec} \quad \epsilon_s \sim \mathcal{N}(0, \sigma_{\mathrm{G}}^2) \tag{IV.19}$$

$$d_s \,|\, d_{\mathrm{G}} \sim f \equiv \mathcal{N}(d_{\mathrm{G}}, \sigma_{\mathrm{G}}^2). \tag{IV.20}$$

Identifier f revient donc à estimer les paramètres $(d_{\mathrm{G}}, \sigma_{\mathrm{G}}^2)$ de cette distribution.

En intégrant les vrais effets d_s hors du problème, on voit que chaque observation $\widehat{d}_s$ a pour densité de probabilité p_s qui s'écrit

$$p_s(\widehat{d}_s, \widehat{r}_s \,|\, d_{\mathrm{G}}) = \int g_s(\widehat{d}_s, \widehat{r}_s \,|\, d_s) \, f(d_s \,|\, d_{\mathrm{G}}) \, \mathrm{d}d_s$$

qui tient compte de la variabilité composite resultant à la fois des composantes de variance intra- et inter-sujets. Cela prouve que les observations ne sont pas identiquement distribuées à moins que le bruit soit effectivement homoscédastique, comme c'est le cas dans une analyse de groupe standard s'appuyant sur la statistique t de Student. Le modèle hiérarchique peut alors se résumer à travers sa fonction de vraisemblance qui, sous des hypothèses usuelles d'échantillonnage indépendant s'écrit :

$$L(f) = p(\widehat{d}_1, \widehat{r}_1 \ldots, \widehat{d}_S, \widehat{r}_S \,|\, f) = \prod_{s=1}^{S} p_s(\widehat{d}_s, \widehat{r}_s \,|\, d_{\mathrm{G}}) \qquad \text{(IV.21)}$$

Ici La distribution gaussienne f peut être estimée au sens du maximum de vraisemblance comme celle maximisant $L(f)$ dans (IV.21) sur l'espace $\mathcal{F}$ des densités admissibles. Dans le cas d'une famille des distributions gaussiennes, l'algorithme EM (*Expectation Maximization*) proposé dans [MÉRIAUX et coll. 2006] permet de converger vers un point stationnaire de la vraisemblance $L(f)$.

Les travaux récents menés au laboratoire par Alexis ROCHE ont permis d'élargir la recherche de f aux distributions non-paramétriques c'est-à-dire où $\mathcal{F}$ est l'espace de dimension infinie des densités de probabilités sur $\mathbb{R}$.

Les effets mixtes dans un cadre non-paramétrique

Dans le cas non-paramétrique, [LINDSAY 1983] a montré que la solution s'écrit comme un mélange d'au plus S masses de Dirac, et sa recherche consiste à explorer un espace de dimension finie :

$$f(d) = \sum_{s=1}^{S} w_s \delta(d - z_s), \qquad \text{(IV.22)}$$

où les poids positifs w_s du mélange sont inconnus et satisfont une contrainte unitaire $\sum_s w_s = 1$, et où les points supports z_s sont également à identifier. Bien sûr, la solution peut avoir un nombre de composantes strictement inférieur à S si certains points supports z_s coïncident ou si certains poids w_s s'annulent. Ce théorème de représentation généralise la propriété bien connue et centrale pour les méthodes de vraisemblance non-paramétrique que la solution du maximum de vraisemblance sous des observations exactes est fournie par la fameuse distribution empirique [OWEN 2001] :

$$\hat{f}_e(d) = \frac{1}{S} \sum_{s=1}^{S} \delta(d - \widehat{d}_s), \qquad \text{(IV.23)}$$

qui correspond à des poids uniformes, *i.e.,* $\forall s \; w_s = 1/S$, et à des points supports localisés sur les observations. Puisqu'aucune expression explicite n'est disponible dans le cas d'observations entâchées d'erreur, *i.e.,* dans le cas des effets mixtes, Alexis ROCHE a développé une version d'un algorithme EM [ROCHE et coll. 2007] permettant de déterminer un couple solution $(\boldsymbol{w}, \boldsymbol{z})$. Cette solution n'est rien d'autre qu'un maximum local de la vraisemblance, cette dernière pouvant s'avérer multimodale dans le contexte où la variabilité inter-sujet est dominée par la variabilité intra-sujet. Il peut alors devenir intéressant de recourir à des extensions stochastiques de l'EM telles que le SAEM [DELYON et coll. 1999, KUHN ET LAVIELLE 2004] ou d'autres approches par simulation — éventuellement tempérée — afin de rechercher le maximum global de la vraisemblance. C'est l'approche préconisée dans [KELLER et coll. 2008] en sus de la prise en compte d'une incertitude sur la localisation spatiale de l'effet BOLD liée aux erreurs de normalisation.

20 juin 2008

En outre, la recherche de f est un problème classique dans le cadre des statistiques bayésiennes non-paramétriques. Nous souhaitons à l'avenir l'aborder dans ce cadre avec Alexis Roche et Merlin Keller, par exemple à l'aide des processus de Dirichlet et de Pitman-Yor, qui constituent des outils de choix pour définir un modèle *a priori* sur un objet de type densité de probabilité [Ferguson 1973, Escobar et West 1995, Ishwaran et James 2001].

CHAPITRE V

ANALYSE EXPLORATOIRE MULTIFRACTALE DES DONNÉES D'IRMf

Alors que les analyses classiques — *orientées modèle* — en IRMf s'appuient sur une prise en compte explicite du paradigme expérimental pour localiser l'activité cérébrale induite par la stimulation, les approches exploratoires visent à extraire des tendances expliquant les fluctuations des signaux mesurés pour en dégager des composantes intelligibles et explicables *a posteriori*. Dans la plupart des cas, ces tendances sont extraites sur la base d'indices reproductibles, notamment spatialement, on parle alors de méthodes *multivariées* (*e.g.,* analyse en composantes principales ou indépendantes, ...). Ici, on s'intéresse plutôt à une alternative mettant en correspondance de façon *univariée* des signaux acquis soit au repos soit lors d'un protocole d'activation. Il s'agit ainsi d'interpréter l'activité évoquée au regard de l'activité spontanée, sans spécifier la nature du lien entre ces signaux. Il devient alors intéressant de dépasser la recherche d'une relation linéaire et stationnaire entre ces différents types d'activité. Répondre à cette question nous oriente vers le type de méthodes à mettre en œuvre, en prémice d'ailleurs à de futures investigations sur les relations spatiales pouvant sous-tendre cette mise en correspondance. Ces méthodes doivent en effet mettre en exergue des variations locales du signal BOLD présentant par exemple des non-stationnarités ou reflétant un phénomène d'*invariance d'échelle*[1].

Ce chapitre dévoile quelques travaux que j'ai menés en collaboration avec Patrice ABRY et Herwig WENDT de l'équipe SISYPHE de l'ENS de Lyon, sur l'utilisation des méthodes exploratoires univariées dites d'*analyse multifractale* en vue de mieux comprendre de façon non-supervisée les modifications trajectorielles liées à l'activation cérébrale. L'obtention de premiers résultats dans ce domaine sur des séries temporelles suffisamment longues[2] pour obtenir des résultats robustes sur le plan statistique, ouvre des perspectives de recherche intéressantes (études de connectivité fonctionnelle, ...) pour la compréhension du lien entre l'activité de fond et l'activité évoquée par la stimulation. Ce travail a donné lieu à un article de conférence internationale [c.16] et à un article rejeté en 2006 lors d'une soumission à la revue *NeuroImage* [s.1]. La version revue et corrigée présentée ici a été soumise au numéro spécial de *IEEE Special Topics on Signal Processing* dédié à l'imagerie cérébrale le 15 mars 2008.

V.1   Motivation : scaling in FMRI ?

While classical data analysis procedures in functional Magnetic Resonance Imaging (fMRI) rely on a model of the expected BOLD (Blood oxygen Level Dependent) response for localizing evoked activity *i.e.,* specific brain regions involved in the performance of cognitive or behavioral

---

[1]L'invariance d'échelle se traduit par la persistance ou la rémanence d'un phénomène à travers toute une gamme d'échelles au sein d'une analyse multi-résolution.

[2]acquises au moyen d'une séquence EVI parallèle localisée développée par Cécile RABRAIT au cours de sa thèse.

tasks, exploratory analysis aims at finding significant components in the data that explain most of the fluctuations in a model-free manner. A post-hoc classification of these components is often used to discriminate physiological trends related for instance to cardiac beat or breathing rate from stimulus-induced (i.e. evoked) activity. In most cases, one is interested in extracting spatial components using multivariate methods like Principal or Independent Components Analysis techniques [McKeown 2000, Calhoun et coll. 2001, Beckmann et Smith 2004]. In the present work, we rather concentrate on a univariate *model-free framework* which is able to clearly disentangle true scaling behaviors from non-stationary trends superimposed to the BOLD signal irrespective of the presence of an external stimulation (resting state or evoked activity).

The presence of scale invariance in fMRI data has been considered as confound or noise for a long time. Indeed, fMRI time series are known to have a colored noise structure, the majority of which occurs at low frequency. Preliminary evidence that fMRI time series have long memory in time or $1/f$ spectral properties has been demonstrated on "resting state" motion-corrected datasets [Zarahn et coll. 1997, Bullmore et coll. 2004]. Previous studies have shown that head movement is a common source of long memory noise caused by slow rotation or translation of the subject's head through an imperfectly homogeneous magnetic field [Birn et coll. 1998, Birn et coll. 1999, Diedrichsen et Shadmehr 2005]. Physiological factors such as cardiac beat or breathing cycle may also contribute to this scaling phenomenon since they may contaminate the BOLD signal with properties depending on the sampling period of data (*i.e.,* short/long time of repetition (TR)) [Purdon et Weisskoff 1998, Frank et coll. 2001, Kruger et Glover 2001]. Early investigations therefore considered these space-varying low frequency components as noise, which are responsible for potential non stationarities [Bullmore et coll. 2001, Meyer 2003, Long et coll. 2005]. Hence, to fulfill the assumptions underlying the classical model-based localization techniques of brain activaty, most neuropsychologists resort to high-pass filtering to remove these trends. Interestingly, in the last few years, the Generalized Linear Model estimation method has been extended to account for $1/f$ (or fractional Gaussian) noise using wavelet decomposition [Fadili et Bullmore 2002]. More recently, it has been demonstrated that a Bayesian approach incorporating nonstationary noise models outperforms the classical GLM-based techniques in terms of activation detection [Huaien et Puthusserypady 2007]. The latter actually loses crucial dynamic features in the data.

Other authors have pointed out that the BOLD signal itself contains power at virtually all frequencies, notably in randomized event-related designs [Burock et coll. 1998]. Interestingly, recent studies have reported that low-frequency spatial fluctuations in cortical BOLD signals may be indicative of synchronized long memory neuronal oscillations rather than merely noise [Leopold et coll. 2003, Thurner et coll. 2003, Maxim et coll. 2005]. Concomitantly, greater persistence during brain activation has been found in normal subjects in [Thurner et coll. 2003]. This confirms that high-pass filtering may potentially remove parts of the signal of interest. Also, higher predictability summarized in terms of scaling exponent has been reported in patients with Alzheimer disease or with major depressive disorder, especially in brain regions implicated in the early stages of the degeneracy process [Maxim et coll. 2005, Linkenkaer-Hansen et coll. 2005].

Inspired by the connection between $1/f$ and long range dependence, several groups have argued that the analysis of fMRI time series should be performed in the wavelet domain [Fadili et Bullmore 2002, Meyer 2003, Bullmore et coll. 2004, Maxim et coll. 2005, Huaien et Puthusserypady 2007]. Then, a first attempt to identify stimulus-induced signal changes from scaling parameters has been proposed in [Thurner et coll. 2003, Shimizu et coll. 2004]. These authors have developed voxel-based fluctuation analysis (FA) and wavelet multiresolution analysis (*e.g.,* the wavelet

transform modulus maxima method [ARNEODO et coll. 1993, MUZY et coll. 1993]), respectively to high temporal resolution fMRI data. Interestingly, they have shown that the fractal feature of voxel time series can be utilized to separate active and inactive brain regions [THURNER et coll. 2003, SHIMIZU et coll. 2004]. Also, to decide whether fractal analysis can help distinguish motion artifacts from true BOLD responses, complementary analyses have been conducted in [LEE et coll. 2005]. They are based on detrended fluctuation analysis (DFA) and conclude that DFA succeeds in distinguishing among three types of voxels, noise, motion artifacts, and true BOLD responses when classical FA fails to robustly recognize which active regions in the brain are truly involved in certain tasks. However, it has been argued in [TORRES ET ABRY 2003] that wavelet tools perform better than DFA. Therefore, in the present contribution, we describe a new type of voxel-based scaling analysis based on two conceptual changes : First, mono-parameter self-similar (or long memory) models are replaced with multiple parameter multifractal ones ; Second, standard multiresolution quantities such as increments or (discrete or continuous) wavelet coefficients are abandoned to the benefit of wavelet Leaders, new multiresolution quantities bringing significant robustness and gain in estimation performance.

The goals and contributions of the present paper are of different natures. First, we intend to give a brief yet precise and tutorial introduction to the concepts tied to scale invariance : Intuitions and definitions, stochastic models ($1/f$, long range dependent, self-similar and multifractal processes) and relations between them, introduction to the most up-to-date and efficient techniques to perform an actual multifractal analysis on real fMRI data. Instead of the commonly used continuous wavelet transform coefficients based approach [ARNEODO et coll. 2002], we favor the use of new multiresolution quantities, *i.e.*, the *wavelet Leaders* [JAFFARD 2004], based on a discrete wavelet transform and which were recently introduced in the literature and shown to bring significant improvements to both theoretical and empirical multifractal analysis. Also, instead of the classical *structure function* based procedures [ABRY et coll. 2000, ARNEODO et coll. 2002], we elaborate on the use of multifractal (MF) attributes estimated from the logarithm of the wavelet Leaders. These theoretical elements are introduced in Section V.2. Second, we apply this scaling analysis to a recently proposed kind of BOLD fMRI data acquired during a 3D parallel localized Echo Volumar Imaging (EVI) sequence. This is described in Section V.3. Our goal is to analyze whether such data collected during resting state and under a slow event-related visual protocol do possess scaling or not and what their nature is (self-similarity or multifractal). Results are reported in Section V.4.1. It is shown that this *log-cumulant* wavelet-leader based approach enables us to bring evidence for the existence of scaling in fMRI time series for both kinds of data and to characterize their nature. Also, we illustrate that such scaling cannot be confused with putative non stationary trends superimposed to data. A set of statistical tests enable us to conclude that activation induces a clear and systematic increase of the long memory parameter with a (less strong and systematic) decrease in multifractality. This is discussed in Section V.5.

V.2   Scaling and multifractal : theory

V.2.1   Scale invariance (or scaling)

The concept of scale invariance, or scaling, refers to systems or signals for which no scale of time (or space) can be singled out or identified as playing a characteristic role. Equivalently, it amounts to saying that all scales are of equal importance. Historically, scale invariance had been tied to $1/f$ 2nd-order stationary stochastic processes. Let X denote the signal under analysis, and $\hat{\Gamma}_X(f, t_k)$ any standard spectrum estimation procedure, such as Welch estimator or such as

average over sliding time windows (centered around times t_k) of the square ($q = 2$) of smoothed periodograms : Scale invariance is associated to a power law behavior of the spectrum estimate with respect to a wide range of frequencies, $f \in [f_m, f_M], f_M/f_m >> 1$:

$$\frac{1}{n} \sum_{k=1}^{n} \hat{\Gamma}_X(f, t_k) \simeq C|f|^{-\gamma}, \gamma > 0. \tag{V.1}$$

However, it is now well established (eg., [ABRY et coll. 1995, ABRY et coll. 2000]) that $1/f$-processes and standard spectral estimation have limited capacities both for modeling and analyzing scaling. For instance, the power law behavior of the spectrum cannot simultaneously hold in both limits ($|f| \to 0$ and $|f| \to +\infty$) and a $1/f$ spectrum constitutes only a rough definition for a class of models. Therefore, it is now commonly admitted that scale invariance can fruitfully be modeled with self-similar [SAMORODNITSKY ET TAQQU 1994] and/or multifractal processes [RIEDI 2003]. Along the same line, standard spectral analysis tools provide the practitioners with poor estimates of the scaling exponent γ. They also turn out to encounter severe limitations in enabling to distinguish between true power law spectral behavior and low frequency (or slow) non stationarities, trends or drifts that may also exist in data. More specifically, in the context of fMRI, low frequency drifts have been measured in phantoms and cadavers [SMITH et coll. 1999]. It has also been observed that high frequency physiological components can be aliased in low frequencies. Therefore, it has often been thought that the scaling properties that are observed in data actually only correspond to artifacts caused by non stationary behaviors. Hence, wavelet transforms are nowadays used as tools providing relevant analysis of scaling properties with robustness against non stationarity. This has been detailed in e.g. [ABRY et coll. 1995, ABRY ET VEITCH 1998, VEITCH ET ABRY 1999, ABRY et coll. 2000, VEITCH ET ABRY 2001, ABRY et coll. 2002].

Such new models (self-similarity and multifractality) and analysis techniques (wavelets) imply two major changes in paradigms : First, the definition of scaling involves a whole range of (positive, negative and fractional) statistical orders q, and no longer only the second order ($q = 2$) ; Second, standard spectral estimates $\hat{\Gamma}_X(f, t_k)$ are replaced with multiresolution quantities, labeled $T_X(a, t)$, i.e., quantities describing the content of X around a time position t, and a scale a. Qualitatively, the analysis scale a acts as the inverse of a frequency : $a \sim f_0/f$ (f_0 being a constant that depends on the details of the definition of $T_X(a, t)$). Standard examples for $T_X(a, t)$ are given by wavelet, increment or box-aggregated coefficients. Therefore, scale invariance is now commonly and operationally defined as the power law behaviors of the (time average of the $q-$th power of the) $T_X(a, t)$, with respect to the analysis scale a, for a given (large) range of scales $a \in [a_m, a_M], a_M/a_m >> 1$:

$$\frac{1}{n_a} \sum_{k=1}^{n_a} |T_X(a, k)|^q \simeq c_q a^{\zeta(q)}. \tag{V.2}$$

The $\zeta(q)$ are referred to as the scaling exponents and their estimation is often the major goal in scaling analysis. The estimated exponents are in turns used for data identification or classification.

Let us now introduce wavelet coefficients and Leaders, self-similar and multifractal processes seen as stochastic models for scale invariance together with scaling and multifractal analysis procedures.
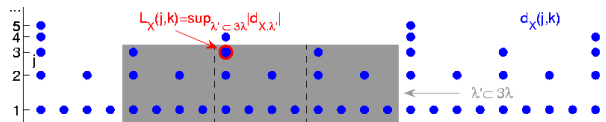
FIG. V.1 – **Definition of Wavelet Leaders.**

V.2.2   Multiresolution quantities

Wavelet coefficients

Let $\psi_0(t)$ denote the *mother-wavelet* : an elementary function, characterized by a fast exponential decay and a strictly positive integer $N \geq 1$, the *number of vanishing moments*, defined as $\forall k = 0, 1, \ldots, N-1$, $\int_{\mathcal{R}} t^k \psi_0(t) dt \equiv 0$ and $\int_{\mathcal{R}} t^N \psi_0(t) dt \neq 0$. The mother-wavelet is moreover chosen such that the collection of templates of ψ_0 dilated to scales 2^j and translated to time positions $2^j k$, $\{\psi_{j,k}(t) = 2^{-j} \psi_0(2^{-j}t - k), j \in \mathcal{Z}, k \in \mathcal{Z}\}$, forms an orthonormal basis of $L^2(\mathcal{R})$. The discrete wavelet transform (DWT) of X is defined through its coefficients :

$$d_X(j,k) = \int_{\mathcal{R}} X(t) \, 2^{-j} \psi_0(2^{-j}t - k) \, dt. \tag{V.3}$$

Note the choice of the L^1-norm (as opposed to the more common L^2-norm choice) that better matches scaling analysis. For further details on wavelet transforms, the reader is referred to e.g., [MALLAT 1998].

Wavelet Leaders

Let us further assume that $\psi_0(t)$ has compact time support and define dyadic intervals as $\lambda = \lambda_{j,k} = [k2^j, (k+1)2^j)$. Also, let 3λ denote the union of the interval λ with its 2 adjacent dyadic intervals : $3\lambda = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}$. Following [JAFFARD 2004], we define the wavelet Leaders by :

$$L_X(j,k) \equiv L_\lambda = \sup_{\lambda' \subset 3\lambda} |d_{\lambda'}|. \tag{V.4}$$

Fig. V.1 illustrates this definition.

Structure functions

For fixed scales $a = 2^j$, the time averages of (the q-th powers of) the $d_X(j,k)$ (resp., the $L_X(j,k)$) are referred to as the structure functions (with n_j the number of $d_X(j,k)$ or $L_X(j,k)$ available at scale 2^j). For convenience, in the remainder of the text, all quantities based on the $d_X(j,k)$ (resp., the $L_X(j,k)$) are labeled with superscript ^{d} (resp., ^{L}) :

$$S^d(j,q) = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)|^q, \tag{V.5}$$

$$S^L(j,q) = \frac{1}{n_j} \sum_{k=1}^{n_j} L_X(j,k)^q. \tag{V.6}$$

20 juin 2008

V.2.3 Scaling models : Long range dependence, Self-similarity, Multifracta-lity

Long range dependence (LRD)

As mentioned in Section V.1, the oldest intuition associated to scaling or scale invariance traces back to $1/f$-processes, i.e., to (second-order) stationary processes Y whose spectrum behaves, for some (wide) range of frequencies, as,

$$\Gamma_Y(f) \simeq C|f|^{-\gamma} \tag{V.7}$$

with $\gamma > 0$. In the particular case where the power law behavior holds in the limit $f \to 0$ and for $0 \le \gamma \le 1$, the process Y is said to be long range dependent or to possess long memory [SAMORODNITSKY ET TAQQU 1994]. The wavelet coefficients $d_Y(j, k)$ of Y are related to its spectrum via

$$2^j \mathbb{E}|d_Y(j,k)|^2 = \int_{\mathcal{R}} \Gamma_Y(f)|\Psi_0(2^j f)|^2 df, \tag{V.8}$$

where Ψ_0 denotes the Fourier transform of ψ_0 [ABRY et coll. 2000]. For LRD processes, it implies the following asymptotic power law behavior of the square wavelet coefficients :

$$2^j \mathbb{E}|d_Y(j,k)|^2 = 2^j \mathbb{E}S_Y(j,2) \sim C_\psi 2^{j\gamma}, \ 2^j \to +\infty. \tag{V.9}$$

Self-similarity (SS)

Long range dependence involves only the description of the statistical second-order ($q = 2$) of data. Self-similar processes provide us with a better mathematically grounded framework to model scale invariance. A process X is said to be self similar when it satisfies $\forall c > 0, \{X(t), t \in \mathcal{R}\} \overset{fdd}{=} \{c^H X(t/c), t \in \mathcal{R}\}$, where $\overset{fdd}{=}$ means equality of all finite dimensional distributions [SAMORODNITSKY ET TAQQU 1994]. For finite variance and stationary increment self similar processes (of interest here and denoted $H-$sssi), the self similarity parameter H is restricted to $H \in (0, 1)$, and one can show that [ABRY et coll. 2000, JAFFARD et coll. 2006] :

$$\mathbb{E}S^d(j,q) = C_q 2^{jqH}, \forall 2^j, q \in (-1, +\infty), \tag{V.10}$$

$$\mathbb{E}S^L(j,q) = C_q 2^{jqH}, \forall 2^j, q \in (-\infty, +\infty). \tag{V.11}$$

When $1/2 < H < 1$, the increments Y of X are LRD, with $\gamma = 2H - 1$. This is why SS and LRD properties are often confused one with the other. Note, however, that Y is stationary when X is not and that LRD involves only the second statistical order and the limit of coarse scales ($a = 2^j \to +\infty$) when SS implies all orders and all scales, from finest to coarsest ($0 < a = 2^j < +\infty$).

Self-similarity consists of a demanding model with respect to empirical data as it requires first, that the scaling property holds for all scales and second, that the single parameter H controls all the statistical properties (i.e., for all q) of the data. From a practical perspective, this is often too severe a limitation and multifractal models are preferred.

Multifractality

Multifractal (MF) processes are often considered as a further extension to model scale inva-riance since they reproduce a form of scaling often observed on empirical data : for some orders

q and a range of scales 2^j, the structure functions of MF processes exhibit power law behaviors with respect to scales, of the form [JAFFARD et coll. 2006] :

$$S^d(j,q) = C_q^d 2^{j\zeta(q)}, \text{ for a range of } 2^j, \ q \in [0, q_*+], \tag{V.12}$$

$$S^L(j,q) = C_q^L 2^{j\zeta(q)}, \text{ for a range of } 2^j, \ q \in [q_*^-, q_*+], \tag{V.13}$$

where the $\zeta(q)$ depart from a linear behavior qH, the signature of self-similarity. This, hence, offers versatility in actual data modeling. Multifractal scaling theoretically only hold in the limit of the finest scales and are necessarily bounded by a coarse scale.

V.2.4 Multifractal theory

Let us now detail multifractal analysis theory.

Hölder Exponent and Singularity Spectrum

Multifractality is related to the irregular behavior of the sample path of X. Its regularity is commonly studied via Hölder exponents, defined in t_0 as :

$$h(t_0) = \sup\{\alpha : X \in C^\alpha(t_0)\}, \quad \text{for } \alpha \geq 0$$

such that in a neighbourhood of t_0 :

$$|X(t) - P(t - t_0)| \leq C|t - t_0|^\alpha$$

with $C > 0$ and the polynomial P satisfying $deg(P) < \alpha$. A simple example is provided by the *cusp-type* function $X(t) = A + B|t - t_0|^h$, for which $h(t_0) = h$ when h is not even. The fluctuations, along time t, of the Hölder exponent h are usually described through the *singularity (or multifractal) spectrum*, labelled $D(h)$ and defined as the Hausdorff dimension of the set of points where the Hölder exponent takes the value h. For details and a complete introduction to MF analysis, the reader is referred to *e.g.,* [JAFFARD 2004, RIEDI 2003].

Wavelet Leader Multifractal formalism

A key practical issue consists in obtaining $D(h)$ from a single observation of finite duration. This is the goal of the MF formalism. Wavelet Leaders have been shown to exactly reproduce the Hölder exponent of X at t_0, insofar as $L_X(j,k) \simeq C2^{jh(t_0)}$, for $2^j \to 0$ and for j, k such that $2^{-j}k = t_0$ [JAFFARD 2004]. This theoretical result fully justifies Eqs. (V.12) and (V.13) above and also connects $\zeta(q)$ and $D(h)$ via the Legendre transform $D(h) = \inf_{q\neq0}(1 + qh - \zeta(q))$. This equality is referred to as the Wavelet Leader Multifractal formalism (WLMF). The full structure function based MF analysis procedure is detailed *e.g.,* in [JAFFARD et coll. 2006]. An example of the resulting estimations for $\zeta(q)$ and $D(h)$ (detailed below) is reported in Fig. V.7 for a representative voxel.

Log-Cumulant

An alternative analysis has been proposed and developed in [CASTAING et coll. 1993, ARNEODO et coll. 2002], we extend it here to wavelet Leaders. For a large subclass of MF processes, Eq. (V.13) is equivalent to $\mathbb{E}|L_X(j,\cdot)|^q = F_q 2^{j|\zeta(q)}$, which can be straightforwardly rewritten as

20 juin 2008

$$\ln \mathbb{E} e^{q \ln |L_X(j,\cdot)|} = \sum_{p=1}^{\infty} C^L(j,p) \frac{q^p}{p!} = \ln F_q + \zeta(q) \ln 2^j, \tag{V.14}$$

where $C^L(j,p)$ stands for the cumulant of order $p \geq 1$ of the random variable $\ln |L_X(j,\cdot)|$. This immediately yields that the coefficients $C^L(j,p)$ have to be of the form :

$$\forall p \geqslant 1, C^L(j,p) = c_{0,p}^L + c_p^L \ln 2^j \tag{V.15}$$

and thus that :

$$\zeta(q) = \sum_{p=1}^{\infty} c_p^L \frac{q^p}{p!}. \tag{V.16}$$

This shows that the knowledge of $\zeta(q)$ and hence of $D(h)$ can be rephrased by that of the *log-cumulants* c_p^L. This is of practical interest as monofractal processes or self-similar processes would be characterized by $\forall p \geq 2 : c_p^L \equiv 0$ (or $c_p^d \equiv 0$), while for multifractal processes of interest $c_2^L \neq 0$ (or $c_2^d \neq 0$). For the multifractal case, when time series are short as in fMRI, c_1^L, c_2^L gathers most of the information practically available from $D(h)$ or $\zeta(q)$: c_1^L characterizes the location of the maximum of $D(h)$, while c_2^L is related to its width. In the systematic analysis of the voxels reported below, we will therefore concentrate on these parameters and thus of approximations of the multifractal characterization of the forms : $\zeta(q) \simeq c_1^L q + c_2^L q^2/2$ and $D(h) \simeq 1 - (h - c_1^L)^2/(2c_2^L)$.

V.2.5 Practical scaling or multifractal analysis

Let us first emphasize that multifractal analysis can be applied to any kind of data or processes, be they multifractal or not, just as Fourier analysis can be applied to any signal be they pure harmonic tones or not. Practical multifractal analysis conducted on empirical data mostly consists of tracking straight lines in $\log_2 2^j$ vs. $\log_2 S^d(j,q)$ ($\log_2 S^L(j,q)$, respectively) plots, referred to as logscale diagrams (LD) (Leader logscale diagrams, (LLD), respectively) and of estimating the corresponding slopes. This is detailed below.

Linear regressions

Practical scaling analysis or empirical MF analysis mostly amounts to measuring the *scaling exponents* $\zeta(q)$, or equivalently the *log-cumulants* c_p, from the observed data. Essentially, analysis and estimation procedures consist in tracking straight lines and estimating slopes in log-log plots, as suggested by Eqs. (V.13) or (V.15) above.

From n_j wavelet coefficients $d_X(j,k)$, wavelet Leaders are computed (according to Eq. (V.4)). Then, standard estimation procedures are used to obtain the estimates $S^L(j,q)$ and $\hat{C}^L(j,p)$. From these, the corresponding $\hat{\zeta}^L(q)$ and c_p^L can then be estimated by linear regression (*cf.* Eq. (V.15)),

$$\hat{\zeta}^L(q) = \sum_{j=j_1}^{j_2} w_j \log_2 S^L(j,q), \tag{V.17}$$

$$\hat{c}_p^L = \log_2 e \sum_{j=j_1}^{j_2} w_j \hat{C}^L(j,p). \tag{V.18}$$

Weights

The weights w_j have to satisfy the constraints $\sum_{j_1}^{j_2} j w_j \equiv 1$ and $\sum_{j_1}^{j_2} w_j \equiv 0$ and can be expressed as $w_j = b_j^{-1} \frac{V_0 j - S_1}{V_0 V_2 - V_1^2}$ with $V_i = \sum_{j_1}^{j_2} j^i b_j^{-1}$, $i = 0 : 2$. As in [ABRY et coll. 2000], we chose $b_j = n_j^{-1}$ to perform weighted linear fits.

V.2.6   Wavelet coefficients versus Leaders

The estimation procedures can be rewritten directly using the absolute value of the wavelet coefficients $|d_X(j, k)|$ instead of Leaders $L_X(j, k)$, yielding mutatis mutandis the estimates $\hat{\zeta}^d(q)$, $\hat{c}_d^L(p)$. This is referred to as the Wavelet Coefficient Multifractal formalism (WCMF).

Until a recent past, this WCMF was the only MF formalism available. However, it suffers from two major drawbacks. First, structure functions $S^d(j, q)$ can obviously not be used for $q < 0$. Coefficients $d_X(j, k)$ hence do not probe the entire multifractal spectrum of X (the largest Hölder exponents are beyond analysis). Second, coefficients $d_X(j, k)$ reproduce the Hölder exponents only for subclasses of MF processes. It has been shown that wavelet Leaders $L_X(j, k)$ efficiently solve these two issues : They accurately measure the entire multifractal spectrum for all MF processes and should hence be preferred (see [JAFFARD 2004, JAFFARD et coll. 2006] for details and proofs). It has also been recently observed that Leader based estimation procedures have far better statistical performance than those based on coefficients [WENDT et coll. 2006].

An   earlier solution to the negative q issue had been proposed (see *e.g.,* [ARNEODO et coll. 2002]). It is based on the Modulus Maxima of the CWT. However, wavelet Leaders provide us with a better mathematically grounded solution, exhibit a far lower computational cost and can be readily extended to processes of higher dimensions (to images for instance).

V.3   Data acquisition

V.3.1   Why do we use EVI fMRI datasets ?

Irrespective of the retained approach, analysis of scaling behavior implies long enough signals. In [SHIMIZU et coll. 2004], the authors tested their MF analysis on Echo Planar Imaging (EPI) fMRI data, which temporal resolution was decreased down to 200ms for partial brain volume acquisition to get up to 1500 time points. Here, we resort to a new imaging technique called parallel localized EVI, recently validated on the human brain [c.13, a.12]. This imaging procedure offers a very high temporal resolution (one volume every 200ms) and thus enables acquisition of a larger number of brain volumes (not only a single slice) in a standard acquisition period (here 2210 scans). This offers the possibility to perform reliable scaling analyses. The present study therefore aims at exploring the benefit of this new leader based MF analysis in combination with EVI brain images, which are highly resolved in time, as described in the next part.

V.3.2   Echo-Volumar Imaging technique

The principle of EVI has been introduced in [MANSFIELD 1977]. Faster than EPI, EVI allows 3D single-shot acquisition at very high scanning rates. Nevertheless, this acquisition technique requires very high performances from the MR hardware in order to avoid signal losses and geometric distortions due to B_0 inhomogeneities and long echo train durations. Thus, only a few

attempts at using EVI in fMRI have been performed until now, focusing mainly on very aniso-tropic brain volumes [Mansfield et coll. 1995, Zwaag et coll. 2006]. Due to improved gradient hardware and magnet homogeneity, and especially to the application of parallel acquisition and reconstruction, we demonstrated the feasibility of acquiring large, isotropic, brain volumes with EVI, at usual fMRI spatial resolution, in about 200 ms [a.12]. Localized Parallel EVI relies on the use of outer volume suppression pulses and parallel acquisition with undersampling by a factor of 2 along two directions, in order to reduce the echo train durations by a factor of 4, as described in Fig. V.2. Consequently, a $120 \times 120 \times 144 \text{mm}^3$ brain volume can be acquired in 200 ms, with a distortion level comparable to EPI. Parallel reconstruction was performed using an in-house-developed multidimensional SENSE algorithm, which requires one sensitivity map for each acquisition channel. In order to improve signal stability in the reconstructed time series, parallel reconstruction was regularized, as pioneered in [Lin et coll. 2005, Ribés et coll. 2007]. Tikhonov weighting was applied, with a regularization condition minimizing the magnitude of the MR signal in the reconstructed volumes, as proposed in [King 2001]. The relative importance of the regularization term was modulated by a regularization parameter λ^2, empirically set.

All experiments were performed on a 1,5 T GEHC scanner (40 mT/m, 150T/m/s slew rate gradient, 8 channel head coil array). EVI acquisitions have been performed using the following parameters : orientation= sagittal plane, TE/TR = 40/200 ms, flip angle (FA) = 35°, BW = 62.5 kHz, FOV = $80 \times 80 \times 100 \text{ mm}^3$, acquired/reconstructed matrices = $20 \times 10 \times 10 / 20 \times 20 \times 20$, echo train duration = 60.5 ms. Sensitivity maps : sagittal plane, TE/TR = 10/500 ms, FA = 30°, BW = 62.5 kHz, FOV = $240 \times 240 \times 100 \text{ mm}^3$, matrix $60 \times 60 \times 20$. A high resolution T1-weighted 3D volume was also acquired for each subject for anatomical localization ($256 \times 256 \times 128$ matrix, voxel size = $0.9 \times 0.9 \times 1.2 \text{ mm}^3$) as shown in Fig. V.2(b)-(c).

V.3.3   fMRI data analysis

For ease of interpretation and validation, we implemented a classical slow event-related fMRI paradigm which studies occipital responses to presentation of alternative contrast checkerboard.

The five healthy subjects gave their written informed consent and this study was approved by a local ethical committee for biomedical research. Two sessions of a slow visual event-related paradigm were acquired for each subject. The stimulus was a black and white contrast reversing checkerboard with a 20-ms period, which appears during 80 ms, followed by a 24.67-ms rest period (ISI = 24.75 s). One session consisted of 20 trials of the stimulus. All series were corrected for subject motion with SPM2 (`www.fil.ion.ucl.ac.uk`). No spatial smoothing was performed. Response magnitudes for each voxel were estimated using a general linear model with a canonical Hemodynamic Response Function (HRF) and its first derivative as regressors. A Fisher (F) test was performed to assess significance. 3D superimpositions with anatomical data were obtained with Anatomist (`http://brainvisa.info`) as shown in Fig. V.2(b). For illustrative purpose, we report activations for one subject in Fig. V.2(c) that were detected both in occipital cortex and cerebellum. We obtained quite reproducible activations across subjects. Fig. V.3 also depicts voxel-based HRF estimates that were computed by selective averaging in voxels eliciting an activation from the raw time courses.

In parallel EPI fMRI acquisitions, regularization of parallel reconstruction generally increases the functional Contrast to Noise Ratio, through a noise reduction larger than the contrast reduction. This increase of CNR mechanically results in an increased sensitivity to the BOLD contrast. This effect has also been observed in parallel EVI : median statistical scores increase when the regularization parameter is increased. Nevertheless, this improvement slows down at high regularization factors, as also observed in [Ribés et coll. 2007], since the regularization
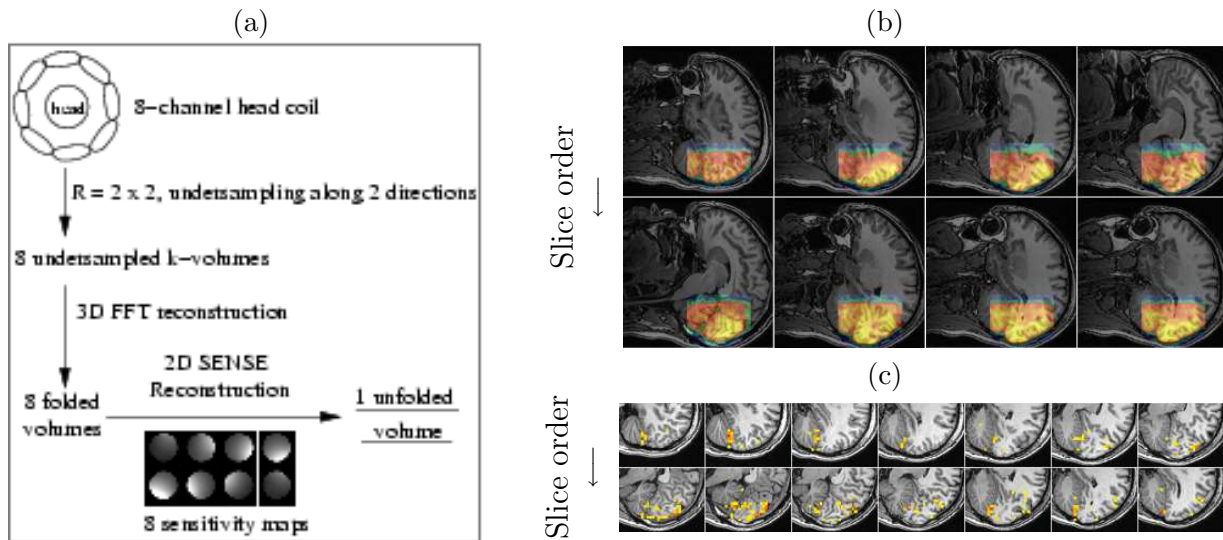
FIG. V.2 – **Localized Parallel EVI imaging sequence. (a)** : sketch summarizing the parallel reconstruction stragegy based on 2D SENSE unwarping technique. **(b)** : anatomical localization of zoomed EVI BOLD data. **(c)** : visual activations detected in EVI datasets (single subject) superimposed on anatomical data.
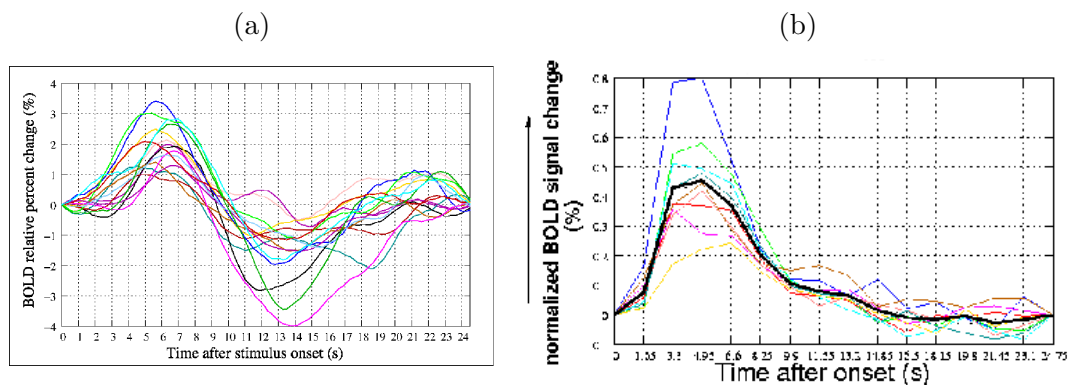
FIG. V.3 – **Comparison of brain dynamics seen in EVI and EPI BOLD sequences.(a)** : voxel-based HRF estimates from the parallel localized EVI sequence. **(b)** : corresponding HRF estimates from the EPI sequence for the same subject.

condition becomes predominant over the accuracy of the reconstruction. Therefore an optimal value of λ^2 has been determined from the GLM-based analysis ($\lambda^2 \approx 0.01$).

V.4   Results

V.4.1   Scaling and Multifractal : fMRI data analysis

In the present work, the ensuing goal is twofold : First, validation of the existence of scaling in the analyzed data and estimation of the MF parameters ; Second, differentiation of evoked and ongoing brain activity in terms of scaling behaviors. This has never been addressed in the fMRI literature using the WLMF approach. For doing so, we statistically compare MF parameters estimated from raw motion-corrected fMRI time series acquired during activation and resting runs from selected regions of interest (ROI). Importantly, these ROIs have been identified on each subject separately from the uniquely relevant F-contrast $c = [1, 1, 0]$ to detect activations

in the above mentioned SPM2 analysis. The extracted SPM clusters have been corrected for multiple comparisons and thresholded below 5 % in corrected p-value and above 5 voxels in spatial extent (up to 25 voxels). Note that the number of SPM clusters R varies from one subject to the other.

Scaling in fMRI data ?

The first major issue lies in assessing whether the data possess scaling properties or not. To do so, one can naturally compare the power spectral density (PSD) estimate (obtained from an averaged periodogram) and the so-called log-scale diagram (LD) defined as j *vs* $\log_2(2^j S(j, q = 2))$ (cf. [ABRY et coll. 2000] and Section V.2). Using Eq. (V.8), both estimates can be superimposed for comparison purposes. PSD, LD and superimposition are plotted in Fig. V.4 together with the time course of a voxel chosen in R_2 (and labeled $Y(t)$ for consistency with notations in Section V.2.3). Both the LD and PSD plots show clear power law behaviors existing over times ranging from the second to the minute. These are clear evidences in favor of the existence of scaling properties in data over this range of scales. Equivalent plots and conclusions can be drawn for each voxel of the different ROIs confirming that scaling actually exists in the analyzed data.
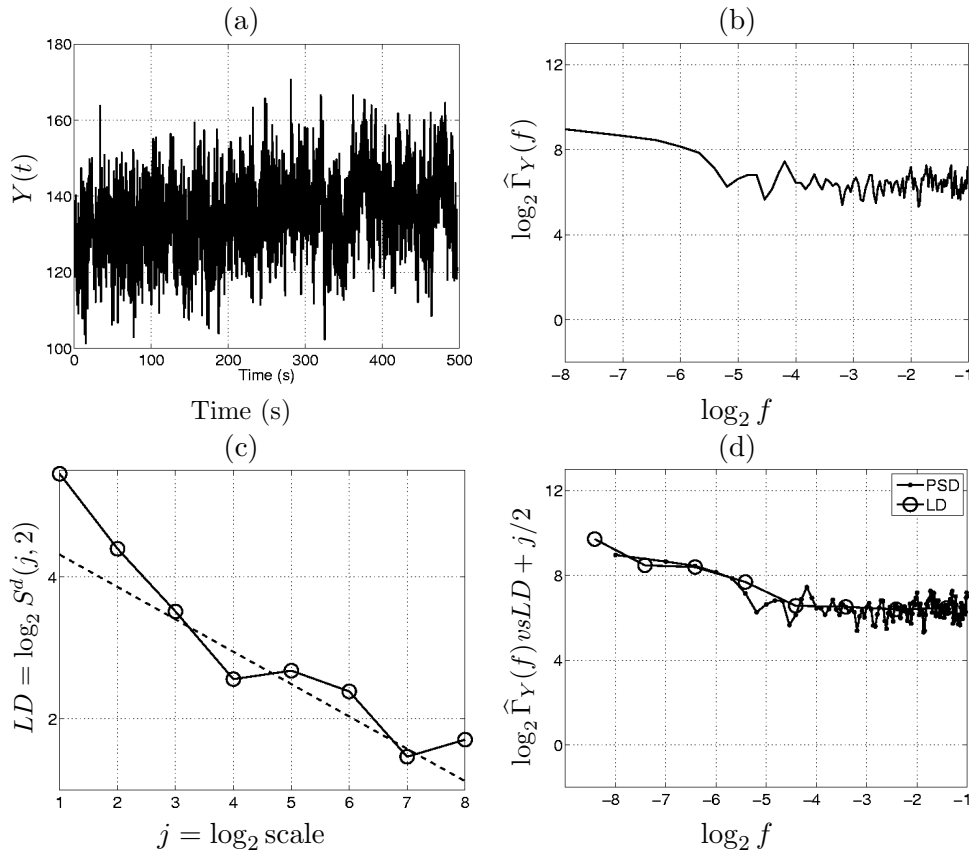
FIG. V.4 – **Scaling in data. (a)** : typical time course $Y(t)$ of the fluctuations recorded in a voxel of R_1 during the visual paradigm ; **(b)** : corresponding PSD estimate ; **(c)** : corresponding wavelet LD plot ; **(d)** : comparison of the PSD with the suitably renormalized LD highlighting the superimposition of the power law behaviors and of the scaling ranges.

Scaling versus Non-Stationarity ?

However, Fig. V.4(a) and Fig. V.5(a) show that $Y(t)$ also possesses a low frequency trend. Such a very slow oscillation exists for most of the voxels under analysis. Their orders of magnitude may however significantly vary in space as may the involved range of low frequencies. These trends can be read as non stationary shifts of the mean superimposed on stationary fluctuations. Because very low frequency oscillations can naturally be modeled as $1/f$-processes, it has been claimed in the literature that the scaling observed on data may only correspond to spurious artifacts caused by non stationarities. However, it has been previously shown that wavelets brings substantial robustness in the analysis of scaling and in disentangling true scaling behavior from slow non stationarities or trends. More precisely, it enables to correctly estimate the scaling parameters when non stationarities are superimposed to a true scaling behavior and conversely to discriminate non stationarities that may be confused with scaling when using standard spectral analysis tools. This has been explained in detail in other contexts in [ABRY ET COLL. 1995, ABRY ET VEITCH 1998, VEITCH ET ABRY 1999, ABRY ET COLL. 2000, VEITCH ET ABRY 2001] and will therefore not be further discussed here. Instead, we chose to further examine the issue of trend versus scaling discrimination by making use of a recent tool, the *Empirical Mode Decomposition* (EMD) [HUANG ET COLL. 1998]. The precise definition and details of EMD cannot be given here for sake of space, the reader is referred to the excellent tutorial paper [FLANDRIN ET P.GONÃ§ALVÃ¨S 2004]. The main output of EMD consists of the fact that the analyzed data can be split into various components.

We chose to split data into 3 components : a very low frequency trend (LF), a medium frequency signal (mF) and a high frequency noise (HF), as illustrated in Fig. V.4(a). Both the LF trend and the HF noise may either alter the analysis of scaling or be confusingly associated to scaling, as commonly speculated in the literature.

LDs are computed independently for the entire original time series, and the three components. They are superimposed in Fig. V.5(b), clearly showing that the scaling, and scaling range, observed on the entire time series are neither caused by the LF nor by the HF components but are rather entirely due to fluctuation (mF component). The trend in itself also exhibits a power law behavior in the LD plot, but in a much coarser range of scales (beyond the minute) and with a clearly different scaling exponent. Moreover, when N (the number of vanishing moments of ψ_0) is increased the scaling exponent that can be estimated from the trend varies in a proportional and trivial manner. This is not the case for the scaling exponent estimated from the fluctuation (mF) component that remains (quasi-) constant when N is varied, a significant experimental proof in favor of the actual existence of scaling in data (see [ABRY ET COLL. 2000] for further details). Moreover, we can check that the scaling exponents estimated on the entire time series and on the mF component closely match. These analyses unambiguously disentangle true scaling properties from non stationary superimposed trends or high frequency noise corruption.

The analyses reported here clearly shows that the fMRI data under analysis possesses scaling properties, within scales that range from the second to the minute and that are not related to slow trends, whatever their origins.

Multifractality ?

Now that the existence of scaling in data is established, let us proceed, in a second step, to a more detailed and quantitative multifractal analysis. We apply the WLMF analysis procedure, as described in Section V.2.5, to the voxel-based mF signal component for all voxels in the identified SPM clusters. Because the Leader approach requires a positive regularity of the time series (*cf.* [JAFFARD 2004, JAFFARD ET COLL. 2006]), analysis procedures are applied to the *cumulated*
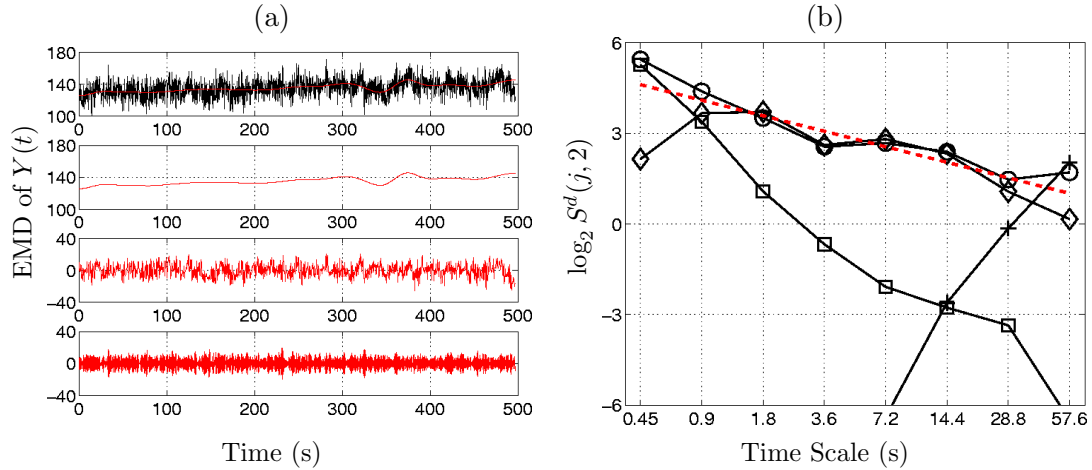
FIG. V.5 – **Scaling _vs_ Non-Stationarity Analysis.** (a) : fMRI time series for an activating voxel, and its EMD based separation into (from top to bottom) low, medium and high frequency components. (b) : the corresponding LDs (time series : 'o', LF : '+', mF : '◇', HF : '□').

sum of the observed time series (_i.e.,_ labeled as X in the notations of Section V.2) rather than to the time series Y itself.

The corresponding Leader-based LDs (LLDs) show clear scaling behaviors holding for $3 \leqslant j \leqslant 6$, i.e., for time scales ranging from 1.5 to 15s that slightly differ between rest and activation datasets. This is illustrated in Fig. V.6 for a voxel (chosen as representative) both during the visual experiment and resting state. The corresponding scaling exponents $\zeta(q)$ and MF spectrum $D(h)$ are compared in Fig. V.7(a)-(b), respectively. It shows that, for both the visual experiment and the resting state, the departure of the $\hat{\zeta}(q)$s from a linear behavior in q is weak, yet clear (or equivalently that the $\hat{D}(h)$ is narrow, yet not collapsing onto one single point). This hence leads to the conclusion that the analyzed data possess a weak yet significant and measurable multifractality. This is further confirmed by the analysis of the plots of the log-cumulant structure functions, $C^L(j,p)$, defined in Eq. (V.15), in Fig. V.8 (from which $\hat{c}_1^L$ and $\hat{c}_2^L$ are estimated). The estimated $C^L(j,2)$ clearly possesses a (negative) non zero slope hence yielding a negative $\hat{c}_2^L$. Let us emphasize that the more negative $\hat{c}_2^L$, the stronger the experimental evidence in favor of multifractality.

Now that an experimental scaling range has been conclusively determined, $1.5 \leq 2^j T_s \leq 15$s, a systematic estimation of the multifractal parameters can be performed for each voxel for both the the visual experiment and resting state. Scaling exponents $\zeta(q)$, the multifractal spectrum $D(h)$ or the log-cumulant c_p accounting theoretically equivalently for the scaling content of the analyzed data, we chose to concentrate on the log-cumulant c_1^L, c_2^L only : c_1^L mostly amounts to the self-similarity characterization, while c_2^L measure the deviation from pure self-similarity hence the impact of the multifractal component of the data.

These systematic analyses yield the following conclusions. For parameter c_1, we observe that it consistently takes values in the range $0.50 \leqslant c_1 \leqslant 1$ (cf. Fig. V.10), both for on-going and evoked brain activity, hence confirming the relevance of the LRD paradigm to characterize fMRI time series correlations. Also, we note that activation systematically (for all subjects and all ROIs) results in an increase in c_1, from the range $0.50 \leqslant c_1 \leqslant 0.75$ for ongoing activity to the range $0.70 \leqslant c_1 \leqslant 0.95$ for evoked activity. Activation hence induces an increase of the LRD strength and impact. This is consistent with findings reported in [SHIMIZU et coll. 2004]. For parameter c_2, the situation is more intricate, partly due to the fact that estimation is by far more difficult [WENDT et coll. 2006]. Nevertheless, we observe that, in some cases, activations coincide with an increase in c_2, from negative to close to 0 values, hence with a decrease in

multifractality (cf. Fig. V.11), while in others there is no significant modification in c_2 between both data types.

Therefore, a key result of these scaling analyses consists of the fact that compared to the resting state, the visual experiment induces a clear, sharp and systematic increase in the LRD parameter, while it either produces a decrease in the degree of multifractality or does not alter it. The statistical significance of these results is further assessed by means of statistical tests in Section V.4.2.
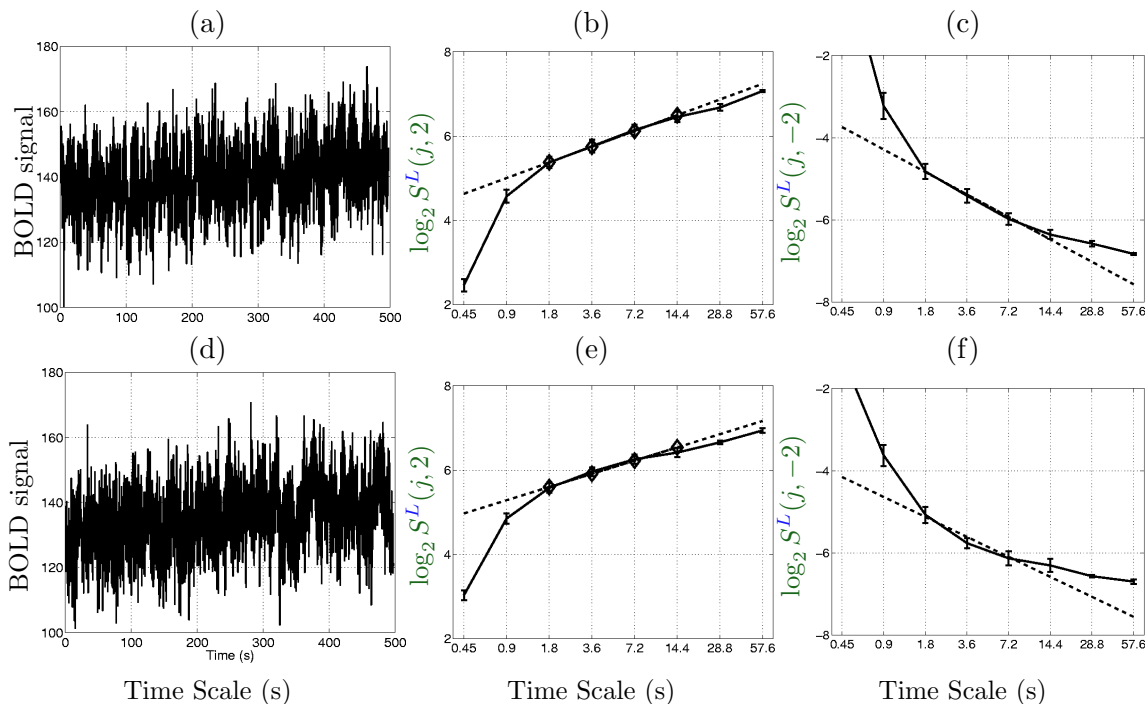
Fig. V.6 – **Leader based Multifractal Analysis. (a)** : fMRI signal acquired during the visual experiment. **(b)-(c)** : The log-scale diagrams computed for $q = 2$ **(b)** & $q = -2$ **(c)** from the time series depicted in (a). They show a clear scaling range, from 1.5 to 15s. **(d)** : fMRI signal acquired during resting state in the same voxel as in (a). **(e)-(f)** : Corresponding Log-scale diagrams computed for $q = \pm 2$ showing also a scaling phenomenon ranging from 1.5 to 15s.

Wavelet coefficients vs. Leaders

To further discuss whether one should prefer to use wavelet coefficients or wavelet Leaders (i.e., WCMF or WLMF) for scaling analysis, one can compare estimations for c_1 and c_2 in Fig. V.9. For c_1, one observes that the dispersion (or confidence interval size) for the $\hat{c}_1^d$ is larger than that of the $\hat{c}_1^L$ but not significantly so. For c_2, the situation is very different, the dispersion of $\hat{c}_2^L$ is dramatically diminished (by one order of magnitude) compared to that of $\hat{c}_2^d$. This is in perfect agreement with the analyses reported in [WENDT et coll. 2006] and yields the following conclusions of major practical importance. As long as scaling are modeled in terms of self-similarity only (i.e., one assumes that the sole parameter c_1 describes scaling), statistical performance are slightly better for Leaders compared to coefficients, at the price though of extra difficulties in selecting the range of scales where to perform the linear regressions. Therefore, Leaders and coefficients should be used jointly and collaboratively. When it comes to estimate c_2, hence scaling related to multifractal properties, only the WLMF should be used as the confidence interval sizes obtained with WCMF are so large that no conclusion can be drawn. Practically, in the statistical tests performed in Section V.4.2 below, this induces that it is strictly not possible to detect any change in c_2 when one uses wavelet coefficients, while Leader based tests clearly show a number of changes. Relevant and accurate estimations of c_2, with reduced confidence

20 juin 2008
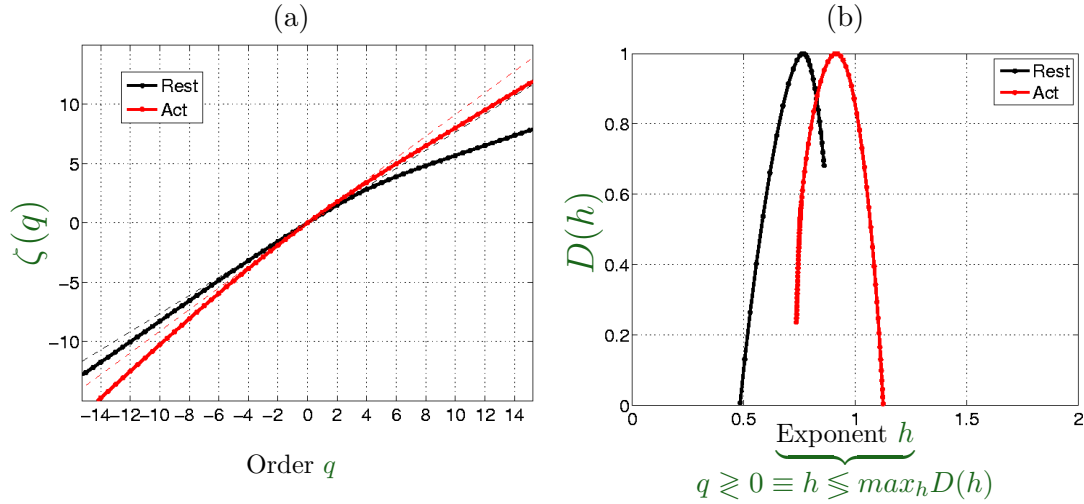
FIG. V.7 – **Leader based Multifractal spectrum**. Computation of the pair $(\zeta(q), D(h))$ from two fMRI time series corresponding to activation (red) and resting state (black) datasets in the same voxel.
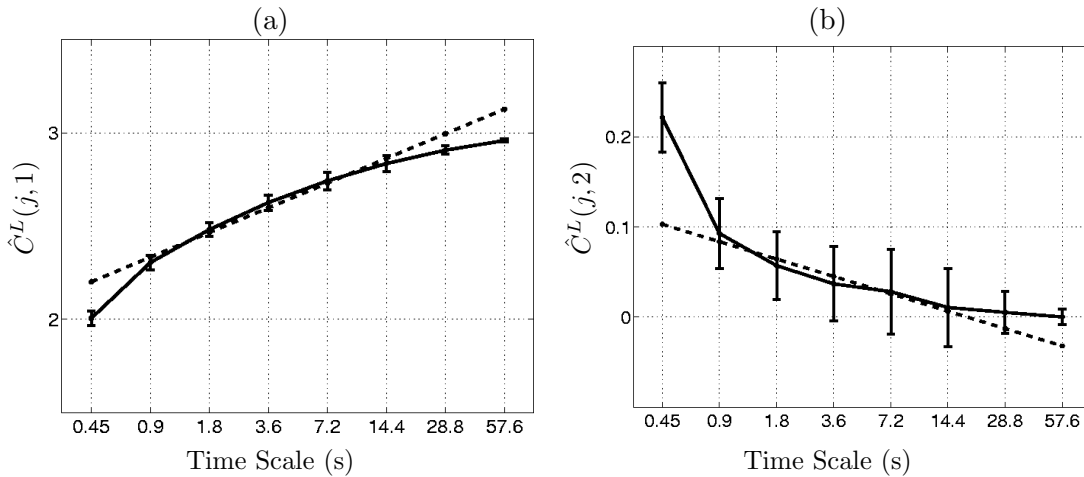
FIG. V.8 – **Log-Cumulants :** Estimated log-Leaders Cumulant functions $\hat{C}^L(j, p)$ for $p = 1$ (a) and $p = 2$ (b). From these linear behaviors in j, one estimates the classifiers $\hat{c}_1^L, \hat{c}_2^L$.

interval sizes, therefore constitute the major benefits of the use of Leaders in this context. Also, let us note that tests based on wavelet coefficients would miss a number of changes in c_1, despite their effect being net, that are clearly seen using Leader based tests.

V.4.2   Region-based hypothesis testing

The next goal consists in assessing the statistical significance of the observed difference in every cluster R_i between $\text{med}[\hat{c}_p^r]$ and $\text{med}[\hat{c}_p^v]$. We use nonparametric tests and robust statistics as there is no evidence that the scaling parameters are normally distributed across voxels for a given ROI. In such a case, one usually resorts to robust decision statistics (*e.g.,* to the Wilcoxon's signed rank (WSR) statistic), whose correct specificity control (control of false positives) can be correctly handled in the permutation testing framework [MÉRIAUX ET COLL. 2006, NICHOLS ET HAYASAKA 2003]. Here, robustness means that the influence of outliers on statistics remains bounded. Precisely, we perform the following *two-sided*

tests :

$$H_{0,p}^{\text{r}\neq\text{v}} : \text{med}[\widehat{\boldsymbol{c}}_p^{\text{r}}] = \text{med}[\hat{\boldsymbol{c}}_p^{\text{v}}], \quad \forall\, p = 1 \text{ or } 2, \tag{V.19}$$

which amount to testing whether the difference between the matched samples $\widehat{\boldsymbol{c}}_p^{\text{r}}$ and $\widehat{\boldsymbol{c}}_p^{\text{v}}$ comes from a distribution whose *median* $\text{med}[\widehat{\boldsymbol{c}}_p^{\text{r}-\text{v}}] = \text{med}[\widehat{\boldsymbol{c}}_p^{\text{r}} - \widehat{\boldsymbol{c}}_p^{\text{v}}]$ is zero.

Fig. V.10 shows the corresponding WSR statistic p-values and validates that the observed increase in c_1 is quasi-systematically significant across subjects. Again, for c_2, the results reported in Fig. V.11 are less clear, as significance of the changes varies with ROIs and subjects. However, results, over the entire datasets indicate a slight shift tendency in c_2 from negative to close to 0 values, when the test is significant. This confirms a global effect of reduction of multifractality under activation.

V.5 Discussion

In this contribution, we have analyzed the scaling and multifractal properties of EVI fMRI data acquired during a resting state and a slow visual event-related activation protocol. To do so, we made use of the most recent theoretical and practical developments in multifractal analysis. They are based on wavelet Leaders instead of wavelet coefficients. Also, we clearly detailed and entangled the often confused relations between the most common stochastic process models used to account for scaling, namely $1/f$-, long range dependent, self-similar and multifractal processes. The various parameters commonly involved in scaling description are also clearly related, and the analysis and estimation procedures are detailed. The use of wavelet coefficients and wavelet Leaders is compared to the benefits of the latter.

Using this careful methodology and powerful estimation tools, we provided significant evidence for the presence of scaling in the analyzed fMRI datasets, for the five participants, that can in no way be confused with non stationarities or low frequency oscillating trends, which were also present in data, yet superimposed to true scaling. Also, we established that this scaling properties should be related to both long memory in time and multifractality (in agreement with [Fadili et Bullmore 2002, Shimizu et coll. 2004]). Given our acquisition parameters, especially the deliberate choice of a large scan number, scaling parameters are estimated with a high degree of precision and thus are reliable.

As concluded from Section V.4.2, we demonstrated the existence of relationships between the scaling parameter estimates and the congruence between the ROIs and the data under consideration. Our findings show that activation induces a clear and systematic increase in the long range dependence (or self-similarity) c_1 parameter together with a (less clear and less systematic) decrease of the multifractality c_2 parameter. Interestingly, the former is connected to additive random walk and linear filtering while the latter is rather related to non-linear mechanisms. Therefore, the parameter c_2 that characterizes deviations from self-similarity can be thought of as a measure of the importance of non-linear effects in neurophysiological mechanisms. Our results suggest that activation tends to reduce their impact : this could be expected given the very simple nature of our paradigm. Also, in future works, we will take advantage of the use of non parametric bootstrap in multifractal analysis that enables not only to produce an estimate but also a confidence interval for each voxel independently [Wendt et coll. 2007]. This is likely to significantly improve the accuracy of the change detection tests and make feasible the use of mixed effect models for region-based inference.

V.6   Perspectives

This preliminary study needs to be further developed if one wants to consider multifractal attributes as putative classifiers of brain activity. In particular, to achieve a reliable model-free analysis, which is able to detect and localize task-related activity, we must explore true contrasts in the sense of comparing brain activity induced by different external stimuli on the same time series using these multifractal attributes as stimulus markers. In contrast to [Shimizu et coll. 2004], this should be investigated in event-related protocols, for which the BOLD signal fluctuations in time and frequency is much more complex. In such cases, we could find out more multifractal situations, for instance in regions eliciting habituation or learning phenomena as those described in [a.4]. To this end, we are currently exploring whole brain analysis blind to the use of any a priori model-based detection. On our first experiment on a single subject (results not shown) we observe that MF parameters are primarily influenced in brain regions involved in the experimental paradigm and also that these parameters remain quite unchanged in other regions when comparing the activation dataset to the resting state one. This has to be confirmed and reproduced on more datasets i.e. on more subjects.

Repeating such analyses to the whole brain and not specifically to activating clusters will also make the method more relevant for exploring long range task-related spatial interactions based upon the scaling properties, as suggested in [Maxim et coll. 2005]. Hence, the whole-brain extension of the proposed approach will offer the possibility to study functional connectivity based on resting-state functional fMRI data and to explain how such long range interactions are modulated by external stimulation. Therefore, this study shall fulfill several important objectives for the understanding of brain structure and function :

– Elucidate the structure of spontaneous activity seen in fMRI as a marker of functional connectivity : the correlation of the spontaneous activity between distant regions observed in fMRI reflects the functional relationship between these distant regions [Biswal et coll. 1995]. This is related to the concept of functional integration [Tononi et coll. 1998] and provides the structure of some fundamental subdivisions of the brain (see e.g. [Damoiseaux et coll. 2006, Golland et coll. 2007]).
– Disentangle the relationship with anatomical connectivity : it is not completely clear whether the spatial structure of spontaneous activity reflects the spatial structure of the anatomical connectivity or whether there are other components that could be identified. Such a comparison becomes possible because diffusion MRI provides an in vivo access on human brain anatomical connectivity structure [El Kouby et coll. 2005, Cathier et Mangin 2006].
– Understand the relationship between on-going and evoked activity : Some studies have proposed that evoked activity (i.e. the activity resulting from an input stimulus) would superimpose linearly with spontaneous activity [Fox et coll. 2006] while other studies insist that spontaneous activity could be predictive of the behavior [Fox et coll. 2007]. These two points of view require further assessment.
– Compare the spontaneous activity in different populations (patients and controls). By considering functional connectivity as a measure that characterizes the integration of brain structures into networks, one can use it for comparing different kinds of populations : infants versus adults, normal subjects versus patients [Fair et coll. 2007]. A particular interest in the study of resting-state networks is that the connectivity or integration of these networks are not confounded by different levels of the subjects in the performance of task.

Currently, many approaches have been proposed to study resting-state brain connectivity, mainly bivariate approaches, based on distant correlations of the signal [Cordes et coll. 2000, Lahaye et coll. 2003, Roebroeck et coll. 2005, Achard et coll. 2006], and multivariate ap-

proaches, mainly based on ICA [Damoiseaux et coll. 2006, Perlbarg et coll. 2007] and clustering [Cordes et coll. 2002, Voultsidou et coll. 2005, Thirion et coll. 2006]. Our feeling is that brain connectiviy analysis based upon multifractal attributes would help discriminate linear from non-linear inter-dependency and that the MF methodology defines a framework for studying how network activity is modulated between resting and stimulation periods.
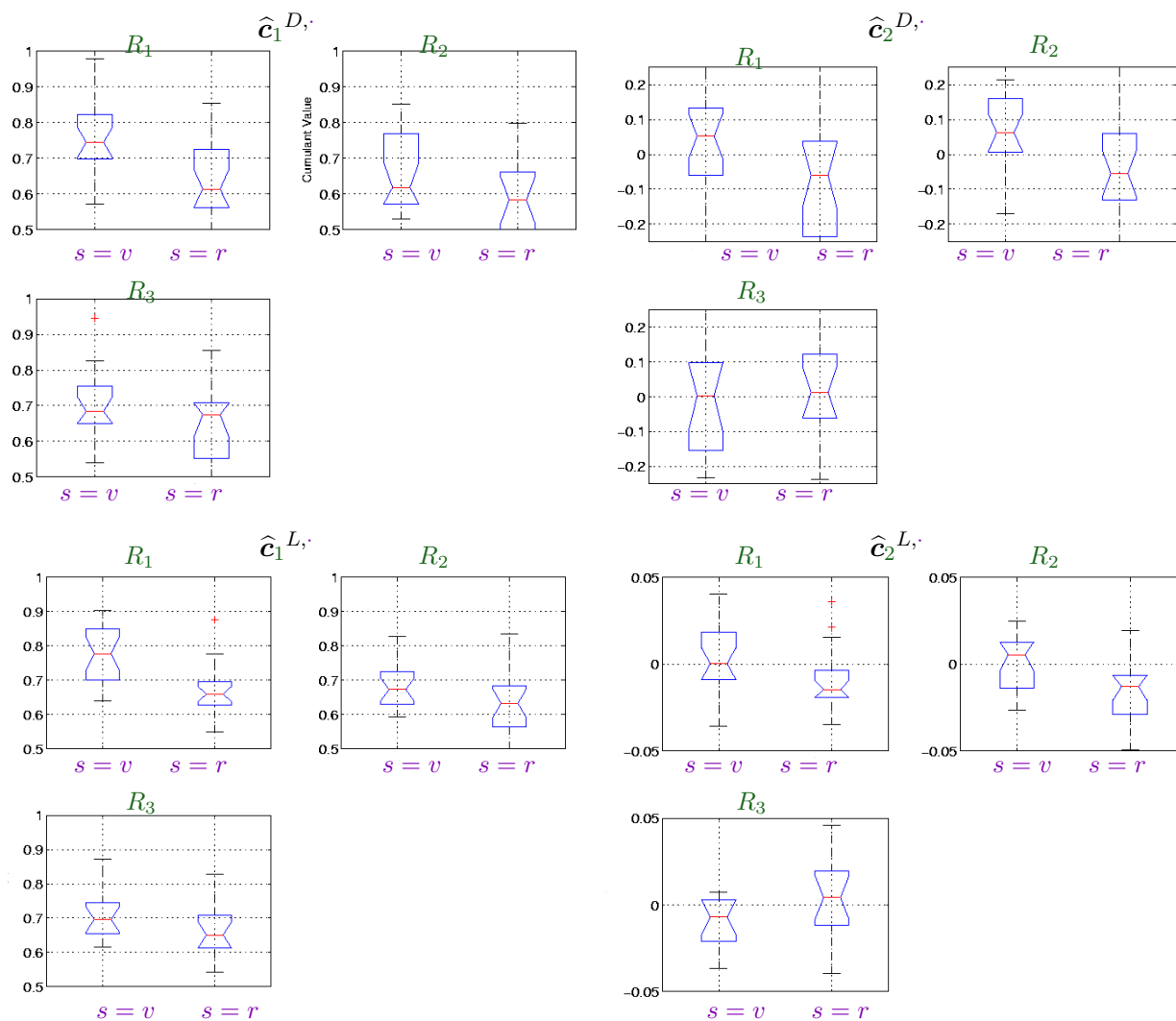
FIG. V.9 – **Statistical comparison of evoked v.s. ongoing activity** : estimation of c_1 and c_2 from the wavelet coefficients (top rows) and wavelet Leaders (bottom rows) for Subject 2. Each panel corresponds to a region (R_1 to R_3) identified as a significant SPM-cluster during the visual experiment. In each ROI, the box-plot shows the *median* of the voxel-dependent cumulant estimates based upon the wavelet coefficients ($\widehat{\boldsymbol{c}}_1^{d,s}, \widehat{\boldsymbol{c}}_2^{d,s}$) and Leaders ($\widehat{\boldsymbol{c}}_1^{L,s}, \widehat{\boldsymbol{c}}_2^{L,s}$) for visual ($s = v$) and rest ($s = r$) sessions, respectively.
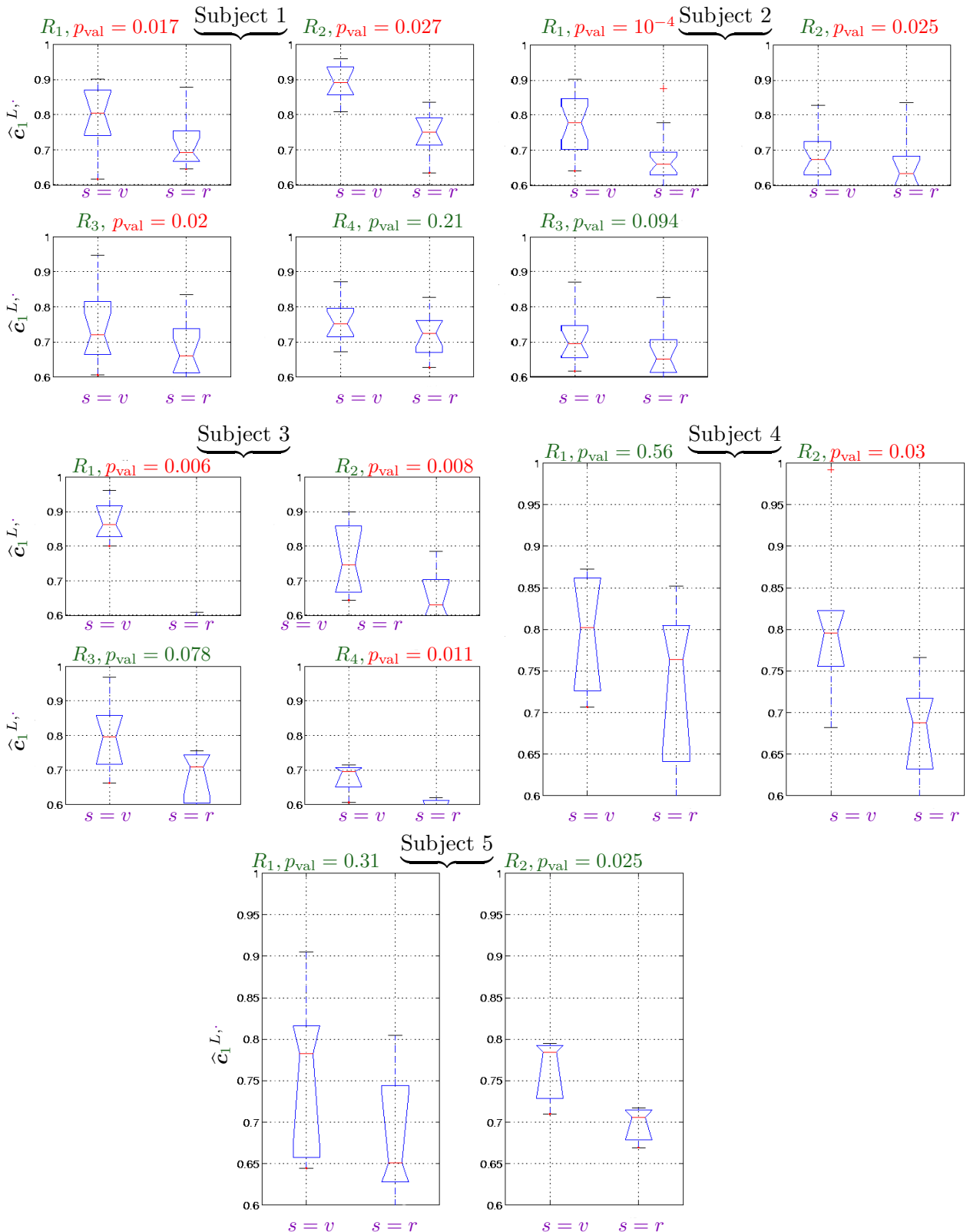
FIG. V.10 – **Statistical comparison of evoked v.s. ongoing activity** : estimation of c_1, *i.e.,* the linear component of $\zeta(q)$. Each panel corresponds to a region (R_1 to R_4) identified as a significant SPM-cluster during the visual experiment. This explains why the number of regions varies across subjects. In each ROI, the box-plot shows the *median* of the voxel-dependent WLMF estimates $\widehat{c}_1^s$ for visual ($s = v$) and rest ($s = r$) sessions, respectively. The WSR statistic p-values are displayed at the top of each panel *i.e.,* for each region. Red marks show significant changes between resting state and visual experiment at 5%. 20 juin 2008
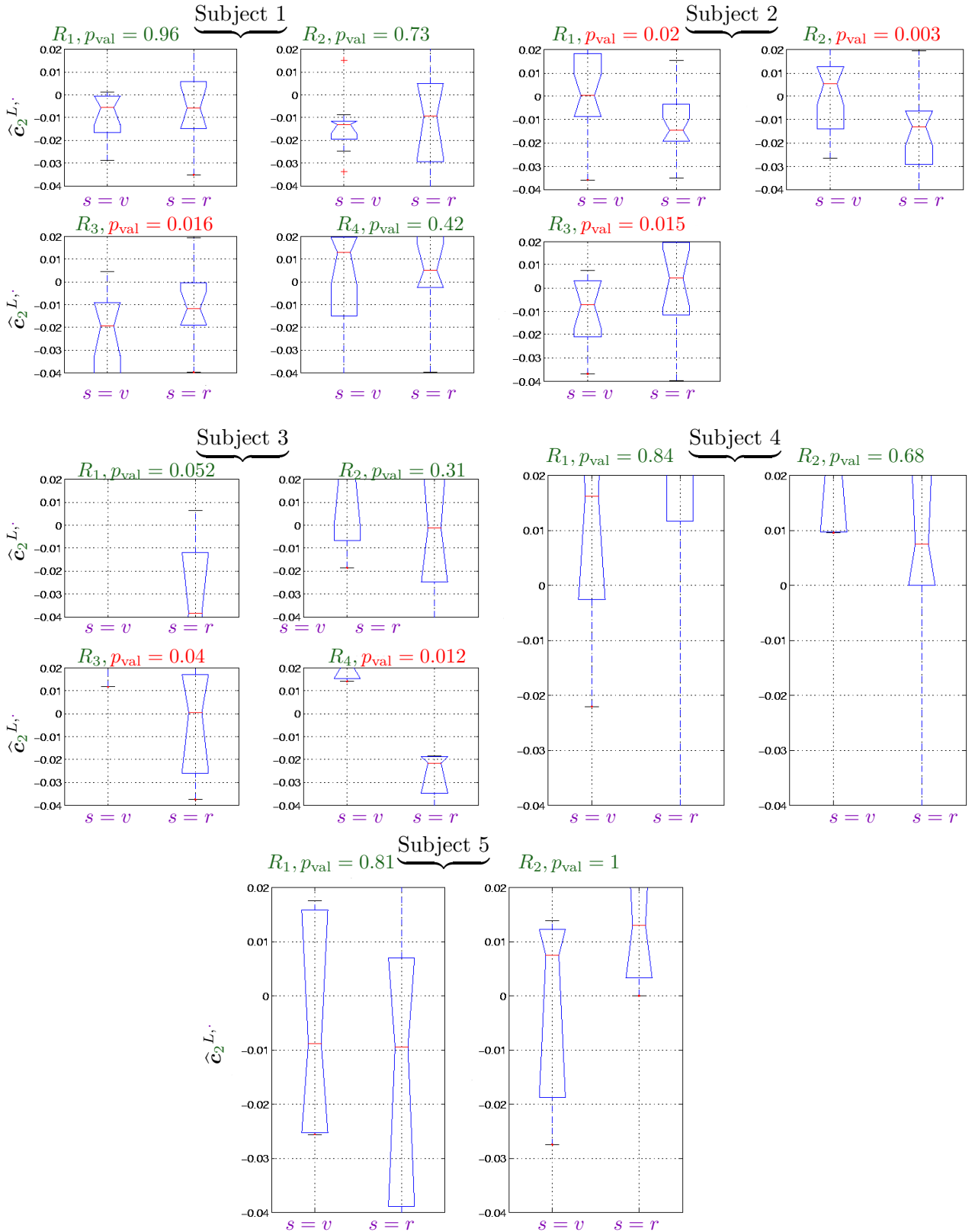
FIG. V.11 – **Statistical comparison of evoked v.s. ongoing activity** : estimation of c_2, *i.e.,* the quadratic component of $\zeta(q)$. Each panel corresponds to a region (R_1 to R_4) identified as a significant SPM-cluster during the visual experiment. This explains why the number of regions varies across subjects. In each ROI, the box-plot shows the *median* of the voxel-dependent WLMF estimates $\widehat{c}_2^s$ for visual ($s = v$) and rest ($s = r$) sessions, respectively. The WSR statistic p-values are displayed at the top of each panel *i.e.,* for each region. Red marks show significant changes between resting state and visual experiment at 5%.

20 juin 2008

Troisième partie

Projet de recherches

Chapitre VI

CONCLUSIONS ET PERSPECTIVES NEUROSCIENTIFIQUES

Dans ce chapitre, je synthétise d'abord mes différentes contributions avant d'évoquer les questions méthodologiques et neuro-scientifiques sur lesquelles je souhaite avancer ces prochaines années. Sur les aspects applicatifs en neurosciences, je distingue le versant cognitif des applications cliniques. Le dénominateur commun à ces préoccupations demeure la dynamique cérébrale.

VI.1 Conclusions

Dans ce mémoire, j'ai présenté un certain nombre de contributions originales relatives à l'estimation de la dynamique des activations cérébrales. Le fruit de ce travail a débouché sur la proposition d'un cadre unifié pour l'analyse intra-sujet des données d'IRMf dont on a pu commencer à mesure l'impact au niveau de l'analyse de groupe au chapitre IV. J'ai mis en évidence que les approches développées dans ce cadre permettaient de résoudre un certain nombre de problèmes de façon élégante, comme la détection-estimation conjointe de l'activité du cerveau évoquée par une stimulation expérimentale, ou l'analyse de la variabilité inter-individuelle de la dynamique cérébrale. Le formalisme développé étant vaste, les méthodes qui s'y intègrent sont encore perfectibles à bien des égards. Les pistes présentées *in extenso* à la fin du chapitre IV font mention d'un certain nombres d'améliorations envisageables à court ou moyen terme (parcellisation adaptative, régularisation spatiale non-supervisée, ...). J'indique aussi dans la suite quelques idées pour les rendre plus robustes dans certaines circonstances (voir section VI.3).

J'ai également proposé au chapitre V des approches originales pour analyser l'activité cérébrale de « fond » à l'aide de techniques exploratoires multifractales, disponibles depuis peu dans la littérature du domaine. J'ai étudié également la signature d'activations cérébrales sur ces attributs multifractaux et établi un lien démontrant l'accroissement du phénomène d'invariance d'échelle lié à cette activité évoquée. La méthodologie proposée s'appuyant sur des décompositions en ondelettes, elle a permis de s'abstraire d'un certain nombre d'artéfacts présents dans les données, comme la superposition de tendances non-stationnaires imputables à différentes sources de bruit. Ceci laisse entrevoir de nombreuses perspectives à ce travail, comme évoquées à la fin du chapitre V, notamment pour les analyses de connectivité fonctionnelle et l'étude de la modulation de réseaux fonctionnels par différents facteurs comme la perception, la conscience ...

Mon projet de recherches s'inscrit donc en partie dans la continuité des travaux déjà entrepris ; il vise à parfaire les méthodes développées pour l'analyse des données d'IRMf, notamment en intra-sujet, mais aussi au niveau d'une population de sujets (utilisation des analyses à effets mixtes). Toutefois, il ne se réduit pas à ces aspects. Il ouvre d'autres sillons tant sur le versant des neurosciences cognitives qu'expérimentales dans le but de prendre des risques susceptibles

d'être à la génèse de véritables découvertes. Ces axes sont plus largement évoqués dans la suite de ce chapitre, étant donné la prédominance laissée à la méthodologie jusqu'ici.

VI.2   Perspectives sur la compréhension du cerveau sain

Par rapport au travail présenté dans ce manuscrit, trois questions relatives à la compréhension du fonctionnement cérébral chez le sujet sain m'intéressent particulièrement.

VI.2.1   Effets hémodynamiques *vs* neuronaux

La première question a trait à la compréhension de l'effet BOLD puisqu'il s'agit de distinguer dans ce signal les effets d'origine hémodynamique (*i.e.,* la « circuiterie du cerveau ») de ceux plutôt cognitifs directement issus de la contribution des neurones. Pour parvenir à cet objectif, j'ai déjà investi beaucoup de temps dans les modèles, mais aussi une énergie significative dans la compréhension de l'imagerie fonctionnelle en tant que telle, qu'elle soit bi- ou tri-dimensionnelle, ... Pour approfondir ma compréhension de ce mécanisme de dissociation, je travaille par ailleurs depuis un an dans l'utilisation d'une nouvelle modalité d'imagerie fonctionnelle en IRM, l'*imagerie de diffusion fonctionnelle*, s'appuyant sur des séquences IRM à double écho de spin pondérée en diffusion. L'idée est d'utiliser une pondération relativement élevée (coefficient de diffusion $b \approx 2000$) de façon à accroître la sensibilité de la mesure au mouvement des molécules d'eau au voisnage des neurones, malgré la dégradation du rapport signal à bruit. Ce travail fait l'objet d'une collaboration active au sein du groupe **Diffusion fonctionnelle** piloté par Denis LE BIHAN à NEUROSPIN.

Des premiers travaux [LE BIHAN et coll. 2006, LE BIHAN 2007] ont montré en effet qu'à l'aide de cette modalité d'imagerie, il semblait possible de détecter de façon plus précoce des activations en réponse à un stimulus externe, confortant ainsi l'hypothèse d'une origine neuronale à ce signal. Plus précisément, lors de l'activation cérébrale, le gonflement des neurones impliqués dans cette activation perturberait le déplacement des molécules d'eau, imagé par IRM de diffusion, et raccourcirait ainsi leur libre parcours moyen. Ces résultats sont néanmoins soumis à une controverse. Une étude récente a en effet défendu la thèse d'un phénomène plutôt d'origine vasculaire [MILLER et coll. 2007] en s'appuyant sur une expérience démontrant une fluctuation similaire du signal de diffusion dans un contexte de repos (pas de stimulus délivré au sujet) et d'*hyper-capnia*, où les sujets inspirent du CO_2, modifiant ainsi le ratio de concentration en oxy -et déoxyhémoglobine, et donc l'effet BOLD vasculaire. Le stage post-doctoral de LAURENT RISSER, que j'encadre au sein du groupe de travail **Diffusion fonctionnelle** est consacré à la modélisation et à l'estimation de la dynamique perçue en imagerie de diffusion fonctionnelle. Ses premiers résultats ont tendance à établir un compromis entre les différents résultats précédemment rapportés dans la littérature [rr.5]. Les bases neuro-physiologiques sous-tendant cette nouvelle technique d'imagerie fonctionnelle pourraient combiner des effets neuronaux et des effets vasculaires. Nous travaillons actuellement à l'éclaircissement de cette piste de recherche dont les conclusions pourront nous aider à mieux discriminer les différentes contributions neuro- et hémodynamiques.

La dissociation de ces effets peut aussi être traitée de façon complémentaire en analysant le lien entre la mesure de l'activation cérébrale captée par les modalités d'imagerie électrique (EEG/MEG) et hémodynamique (IRMf, voire imagerie optique [TOYODA et coll. 2008]) ; pour ce faire, il est intéressant de mener des analyses conjointes sur ces différents types de données, aux résolutions spatio-temporelle complémentaires, acquises isolément ou conjointement. Les paradigmes événementiels nous permettent en effet de soumettre le sujet à la même

expérience dans ces différentes modalités. L'analyse de données simultanées (EEG acquis sous IRM) comporte de nombreux problèmes spécifiques liés aux artéfacts d'une modalité sur l'autre. Néanmois, sur la base de modèles génératifs, il est possible de combiner ces différentes sources d'information. Pour avancer dans cette voie, nous commençons dans le cadre de la thèse de Thomas VINCENT à développer des méthodes de fusion d'imagerie multi-modales (EEG/IRMf) sur la surface corticale, siège de la majorité des activations cérébrales[1]. Il s'agit notamment d'appréhender le lien existant entre l'activité évoquée mesurable d'une part en EEG et d'autre part en IRMf. Pour ce faire, il est clair que la prise en compte de la variabilité inter-essais demeure un point crucial pour l'avenir afin de mettre en correspondance les deux types d'activité essai à essai.

Enfin, l'utilisation de paradigmes expérimentaux évolués (*e.g.,* de type *repetition suppression*) permet partiellement de mettre en évidence cette différence entre des effets vasculaires et des effets neuronaux. Néanmoins, les non-linéarités des réponses BOLD mesurées sont mal capturées par les modèles standards. C'est la raison pour laquelle le développement de modèles parcimonieux — les plus simples possibles — permettant de capturer ces phénomènes non-linéaires continuera à faire partie de mes priorités. De même, étant donné le nombre de combinaisons possibles dans le choix des effets qu'on cherche à modéliser, il me paraît opportun de poursuivre l'effort qui vise à mettre en place des méthodes de comparaison ou de sélection de modèles, notamment dans un cadre bayésien.

VI.2.2 Analyse de la variabilité intra- et inter-individuelle

La seconde voie dans laquelle je souhaite m'investir concerne l'analyse de la variabilité inter-individuelle sur des systèmes cognitifs haut niveau tels que le langage. Cette étude s'appuie sur une collaboration fructueuse depuis plusieurs années avec les spécialistes du domaine *in situ* pour définir au mieux les protocoles expérimentaux (Ghislaine DEHAENE-LAMBERTZ et Christophe PALLIER pour le langage, ainsi que sur la possibilité d'analyser d'importantes cohortes de sujets à l'aide de systèmes de bases de données mis en place à NEUROSPIN. Sur le plan méthodologique, il s'agit de renforcer les liens avec Alexis ROCHE pour le développement d'analyses de groupe à effets mixtes s'appuyant sur des analyses intra-sujet de type détection-estimation conjointe.

Pour bien comprendre les différentes sources de cette variabilité, une possibilité est de mener des études développementales comme le fait Ghislaine DEHAENE-LAMBERTZ afin de mettre en perspective les évolutions liées à la maturation cérébrale. Il s'agit d'analyser de façon longitudinale dès les premières semaines de vie (à 3, 6, 12 mois ...), la modification de l'activation cérébrale chez chaque sujet en associant ces transformations aux phénomènes structurant l'anatomie cérébrale tels que le processus de sulco-gyrification (apparition des sillons et des gyri *i.e.,* des plissements sur le cortex), et en dégageant des éléments communs à tous les sujets. Cette question sera donc traitée en étroite collaboration avec les spécialistes d'anatomie cérébrale du nourrisson (François LEROY, Jean–François MANGIN). La méthodologie présentée au Chapitre IV doit nous permettre d'avancer efficacement dans cette voie.

VI.2.3 Emergence des liens entre activité spontanée et évoquée

La troisième voie qui attire mon attention porte sur la compréhension de l'activité de « fond » ou spontanée, observable sur des données de repos, et de son lien avec l'activité évoquée mesurable au cours d'un protocole d'activation. Les travaux préliminaires que j'ai menés dans cette direction ont été présentés au Chapitre V. Ils nous permettront d'approndir notre compréhension de certains phénomènes, notamment de mieux appréhender la modulation de l'activité cérébrale

---

[1]On néglige ici l'activité des noyaux gris centraux, structures sous-corticales.

par la perception consciente ou inconsciente des stimuli en lien avec le niveau d'activité de fond. Ces premières études pourront aussi servir de point d'accroche pour mener des analyses de connectivité fonctionnelle ou effective afin de mieux caractériser le rôle de certains réseaux fonctionnels et leur modulation entre les situations de repos et d'activation. Il est par exemple envisageable de construire des matrices de connectivité à partir des attributs multifractaux plutôt que sur les signaux originaux dans le but d'améliorer l'extraction d'informations intelligibles, c'est-à-dire en quelque sorte de se débarasser du bruit. De plus, les analyses présentées au Chapitre V sont exploratoires, elle sont donc bien adaptées à l'usage de stimuli sous-liminaire ou écologiques (*e.g.,* visionnage de films), pour lesquels l'*onset* de la réponse du sujet est mal défini. Ces aspects font l'objet d'un groupe de travail *Resting State* à Neurospin piloté par Bertrand Thirion, en collaboration avec et Hugues Berri de l'inria Futurs, Andreas Kleinschmidt et Sepideh Sadaghiani de l'unité inserm U562.

VI.3   Applications à la compréhension de certaines pathologies

Concernant les questions relatives au cerveau pathologique, la maladie neurologique sur laquelle je souhaite me concentrer est l'épilepsie, tant chez l'adulte que chez l'enfant. Tout d'abord, nous poursuivons avec Thomas Vincent une collaboration avec l'équipe de Politechnico de Milan (*i.e.,* Maria-Gabriella Tanna et Anna-Maria Bianchi) sur des acquisitions eeg/irmf simultanées chez des patients épileptiques adultes, notamment pour caractériser la dynamique et la localisation de l'activité inter-critique associée à des pointes-ondes perceptibles sur l'examen eeg. Dans ce but, nous développons à l'heure actuelle une extension multi-sessions des modèles présentés au chapitre IV. Il s'agit d'un modèle à *effets fixes*, c'est-à-dire qui suppose la reproductibilité du modèle bilinéaire du signal bold au sein de chaque parcelle à travers toutes les sessions d'acquisition et le caractère homoscédastique du bruit de mesure (variance constante à travers les sessions).

Grâce à l'avènenemt de nos nouveaux outils d'analyse de la dynamique cérébrale, et leur diffusion au sein de la plate-forme logicielle Brainvisa de l'IFR 49, nous sommes désormais en mesure de caractériser régionalement la différence de forme de la réponse hémodynamique chez ces patients par rapport à des sujets contrôles dans certains foyers épileptogènes, notamment lorsque cette activité inter-critique génère des déactivations c'est-à-dire une inversion de la forme de réponse hémodynamique. Cette inversion peut résulter de processus d'inhibitions médiés par des neurotransmetteurs tels que le GABA. En effet, le mécanisme d'inhibition requiert peu d'énergie et provoque une décroissance importante du taux de décharge des neurones, *i.e.,* de la quantité de potentiels d'action émis, qui se traduit par une diminution de la demande énergétique, et donc du flux sanguin régional. De même, une vaso-constriction peut induire des déactivations vasculaires en vertu de la réduction du volume et du flux dans les vaisseaux. Sur le plan méthodologique, cette caractérisation s'appuie donc sur un modèle de mélange *a priori* à trois classes, présenté en Section III.3.4 du chapitre III.

Je souhaite bien sûr participer à l'acquisition de telles données à Neurospin et me rapprocher des compétences présentes au sein du laboratoire de recherche biomédicale de Neurospin dirigé par Lucie Hertz-Pannier en collaborant activement avec Marion Hulli, neuropédiatre et doctorante au sein de ce laboratoire, et membre de l'équipe INSERM U663 *Épilepsies de l'enfant et plasticité cérébrale* dirigée par Catherine Chiron[2]. Il s'agira d'analyser des formes infantiles de l'épilepsie sur la base d'examens d'irmf.

Pour l'épilepsie, les signes cliniques précèdent les investigations d'imagerie cérébrale, qu'elles soient menées à partir de simples électro-encéphalogrammes ou d'examens d'irmf. Par

---

[2]neuropédiatre et chercheur à l'Hôpital Nécker Enfants malades.

conséquent, il ne s'agira pas dans ces collaborations d'apporter une méthodologie d'analyse à des fins diagnostiques, mais plutôt d'offrir des moyens pour :

– réaliser un suivi longitudinal des patients par des analyses objectives voire quantitatives ;
– étudier l'impact de certains médicaments et en déduire leur pharmacodynamie en fonction du stade de la maladie ;
– caractériser fonctionnellement certaines formes de la maladie peu étudiées.
– potentiellement transférer le savoir faire acquis à d'autres pathologies, orphelines ou non.

Quatrième partie

Références bibliographiques

[ABRY ET VEITCH 1998] P. ABRY ET D. VEITCH (1998), « Wavelet analysis of long-range dependent traffic », *IEEE Trans. on Info. Theory*, **44**, 1, pages 2–15.

[ABRY et coll. 1995] P. ABRY, P. GONÇALVÈS ET P. FLANDRIN (1995), « Wavelets, Spectrum Analysis and $1/f$ processes », dans A. ANTONIADIS ET G. OPPENHEIM, éditeurs, *Lecture Notes in Statistics : Wavelets and Statistics*, volume 103, pages 15–29, 1995.

[ABRY et coll. 2000] P. ABRY, P. FLANDRIN, M. TAQQU ET D. VEITCH (2000), « Wavelets for the analysis, estimation and synthesis of scaling data », dans *K. Park and W. Willinger Eds, Self-similar Network Traffic and Performance Evaluation*, Wiley, spring 2000.

[ABRY et coll. 2002] P. ABRY, R. BARANIUK, P. FLANDRIN, R. RIEDI ET D. VEITCH (2002), « Multiscale nature of network traffic », *IEEE Signal Processing Magazine*, **19**, pages 28–46.

[ACHARD et coll. 2006] S. ACHARD, R. SALVADOR, B. WHITCHER, J. SUCKLING ET E. BULLMORE (2006), « A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. », *J Neurosci*, **26**, 1, pages 63–72.

[AGUIRRE et coll. 1998] G. K. AGUIRRE, E. ZARAHN ET M. D'ESPOSITO (1998), « The variability of human BOLD hemodynamic responses », *Neuroimage*, **7**, pages 574.

[AKAIKE 1974] H. AKAIKE (1974), « A new look at the statistical model identification », *IEEE Trans. Automat. Contr.*, **AC-19**, 6, pages 716–723.

[ALLAIN et coll. 2006] M. ALLAIN, J. IDIER ET Y. GOUSSARD (2006), « On global and local convergence of half-quadratic algorithms », *IEEE Trans. Image Processing*, **15**, 5, pages 1131–1142.

[ANDRADE 2002] A. ANDRADE (2002), « low frequency oscillations, functional connectivity, physiological noise, coherence/phase analyses :bibliographic overview and tentative interpretation of data », rapport technique, MRC Cognition and Brain Science Unit.

[ANDRIEU ET DOUCET 1999] C. ANDRIEU ET A. DOUCET (1999), « Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC », *IEEE Trans. Signal Processing*, **47**, 10, pages 456–463.

[ANTONIAK 1974] C. ANTONIAK (1974), « Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems », *The annals of Statistics*, **2**, 6, pages 1152-1174.

[ARNEODO et coll. 1993] A. ARNEODO, E. BACRY ET J. MUZY (1993), « Solving the inverse fractal problem from wavelet analysis », *Europhys. Lett.*, **25**, pages 479–.

[ARNEODO et coll. 2002] A. ARNEODO, B. AUDIT, N. DECOSTER, J.-F. MUZY ET C. VAILLANT (2002), « Wavelet-based multifractal formalism : applications to DNA sequences, satellite images of the cloud structure and stock market data », *The Science of Disasters*, **A. Bunde, J. Kropp, H. J. Schellnhuber eds.**, pages 27–102.

[AUBERT ET COSTALAT 2002] A. AUBERT ET R. COSTALAT (2002), « A model of the coupling between brain electrical activity, metabolism, and hemodynamics : application to the interpretation of functional neuroimaging », *Neuroimage*, **17**, pages 1162–1181.

[AUBERT et coll. 2005] A. AUBERT, R. COSTALAT, M. P.J. ET P. L. (2005), « Free in PMC Brain lactate kinetics : Modeling evidence for neuronal lactate uptake upon activation », *Proc. Natl. Acad. Sci. USA*, **102**, 45, pages 16448–16453.

[BEAL 2003] M. BEAL (2003), *Variational algorithms for approximate Bayesian inference*, thèse de doctorat, University College of London, London, Royaume-Uni.

[BECKMANN ET SMITH 2004] C. BECKMANN ET S. SMITH (2004), « Probabilistic independent component analysis for functional magnetic resonance imaging », *IEEE Trans. Med. Imag.*, **23**, 2, pages 137–152.

[BECKMANN et coll. 2003] C. BECKMANN, M. JENKINSON ET S. SMITH (2003), « General multi-level linear modelling for group analysis in fMRI », *Neuroimage*, **20**, pages 1052–1063.

[BERTSEKAS 1995] D. P. BERTSEKAS (1995), *Nonlinear programming, Second Edition*, Athena Scientific, Belmont, MA, USA.

[BIRN et coll. 1998] R. M. BIRN, P. A. BANDETTINI, R. W. COX, A. JESMANOWICZ ET R. SHAKER (1998), « Magnetic field changes in the human brain due to swallowing or speaking », *Magn. Reson. Med.*, **40**, pages 55–60.

[BIRN et coll. 1999] R. M. BIRN, P. A. BANDETTINI, R. W. COX ET R. SHAKER (1999), « Event-Related fMRI of Tasks Involving Brief Motion », *Hum. Brain Mapp.*, **7**, pages 106–114.

[BISWAL et coll. 1995] B. BISWAL, F. Z. YETKIN, V. M. HAUGHTON ET J. S. HYDE (1995), « Functional connectivity in the motor cortex of resting human brain using echo-planar MRI », *Magn. Reson. Med.*, **34**, pages 537–541.

[BRÉMAUD 1999] P. BRÉMAUD (1999), *Markov Chains. Gibbs fields, Monte Carlo Simulation, and Queues*, Texts in Applied Mathematics 31, Spinger, New York, NY , USA.

[BROOKS et coll. 2003] S. BROOKS, P. GIUDICI ET G. ROBERTS (2003), « Efficient construction of reversible jump MCMC proposal distributions (with discussion) », *J. R. Statist. Soc. B*, **65**, pages 3–55.

[BROOKS et coll. 2006] S. P. BROOKS, Y. FAN ET J. S. ROSENTHAL (2006), « Perfect forward simulation via simulated tempering », *Communications in Statistics – Simulation and Computation*, **35**, pages 683–713.

[BUCKNER et coll. 1998] R. L. BUCKNER, J. GOODMAN, M. BUROCK, M. ROTTE, W. KOUTSTAAL, D. L. SCHACHTER, B. R. ROSEN ET A. M. DALE (1998), « Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI », *Neuron*, **20**, pages 285–296.

[BULLMORE et coll. 1999] E. BULLMORE, J. SUCKLING, S. OVERMEYER, S. RABE-HESKETH, E. TAYLOR ET M. BRAMMER (1999), « Global, voxel, and cluster tests, by theory and permutation, for difference between two groups of structural MR images of the brain », *IEEE Trans. Med. Imag.*, **18**, pages 32–42.

[BULLMORE et coll. 2001] E. BULLMORE, C. LONG, J. SUCKLING, J. FADILI, G. CALVERT, F. ZELAYA, T. CARPENTER ET M. BRAMMER (2001), « Colored noise and computational inference in neurophysiological (fMRI) time series analysis », *Hum. Brain Mapp.*, **12**, pages 61-78.

[BULLMORE et coll. 2004] E. BULLMORE, J. FADILI, V. MAXIM, L. ŞENDUR, B. WHITCHER, J. SUCKLING, M. BRAMMER ET M. BREAKSPEAR (2004), « Wavelets and functional magnetic resonance imaging of the human brain », *Neuroimage*, **23**, N Suppl.1, pages S234–S249.

[BUROCK et coll. 1998] M. A. BUROCK, R. L. BUCKNER, M. G. WOLDORFF, B. R. ROSEN ET A. M. DALE (1998), « Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. », *Neuroreport*, **9**, 16, pages 3735–3739.

[BUXTON et coll. 1998] R. B. BUXTON, E. C. WONG ET F. L. R. (1998), « Dynamics of blood flow and oxygenation changes during Brain activation : the balloon model », *Magn. Reson. Med.*, **39**, pages 855–864.

[BUXTON et coll. 2004] R. B. BUXTON, K. U. G, D. J. DUBOWITZ ET T. T. LIU (2004), « Modeling the hemodynamic response to brain activation », *Neuroimage*, **23, Supplement 1**, pages S220–S233.

[Bydder et coll. 2002] M. Bydder, D. J. Larkman et J. V. Hajnal (2002), « Generalized SMASH imaging. », *Magn. Reson. Med.*, **47**, 1, pages 160–170.

[Bï¿½ar et coll. 2003] C. Bï¿½ar, Y. Aghakhani, Y. Wang, A. Izenberg, A. A. A. et G. J. Dubeau, F. (2003), « Quality of EEG in simultaneous EEG-fMRI for epilepsy », *Clin. NeuroPhysiol.*, **114**, pages 569–580.

[Cachia et coll. 2003] A. Cachia, J.-F. Mangin, D. Rivière, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch et J. Régis (2003), « A generic framework for parcellation of the cortical surface into gyri using geodesic Voronoï diagrams », *Medical Image Analysis*, **7**, 4, pages 403–416.

[Calhoun et coll. 2001] V. Calhoun, T. Adali, G. Pearlson et J. Pekar (2001), « Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms », *Hum. Brain Mapp.*, **13**, pages 43–53.

[Castaing et coll. 1993] B. Castaing, Y. Gagne et M. Marchand (1993), « Log-similarity for turbulent flows », *Physica D*, **68**, pages 387-400.

[Cathier et Mangin 2006] P. Cathier et J.-F. Mangin (2006), « Registration of cortical connectivity matrices », dans *Proc. MMBIA2006*, New York, USA.

[Champagnat et coll. 1996] F. Champagnat, Y. Goussard et J. Idier (1996), « Unsupervised deconvolution of sparse spike trains using stochastic approximation », *IEEE Trans. Signal Processing*, **44**, 12, pages 2988-2998.

[Chang et coll. 2004] S.-C. Chang, J. L. Jacobsen, J. Salas et R. Shrock (2004), « Exact Potts Model Partition Functions for Strips of the Triangular Lattice », *J.STAT.PHYS.*, **114**, pages 763.

[Charbonnier et coll. 1997] P. Charbonnier, L. Blanc-Féraud, G. Aubert et M. Barlaud (1997), « Deterministic edge-preserving regularization in computed imaging », *IEEE Trans. Image Processing*, **6**, 2, pages 298–311.

[Chaux et coll. 2007] C. Chaux, P.-L. Combettes, J.-C. Pesquet et V. R. Wajs (2007), « A variational formulation for frame-based inverse problems », *Inverse Problems*, **23**, pages 1494–1518.

[Cheng et coll. 1996] Q. Cheng, R. Chen et T.-H. Li (1996), « Simultaneous Wavelet Estimation and Deconvolution of Reflection Seismic Signals », *IEEE Trans. Geosci. Remote Sensing*, **34**, 2, pages 377–384.

[Chib 1995] S. Chib (1995), « Marginal likelihood from the Gibbs output », *J. Amer. Statist. Assoc.*, **90**, pages 1313–1321.

[Chib et Jeliazkov 2001] S. Chib et I. Jeliazkov (2001), « Marginal likelihood from the Metropolis-Hastings output », *J. Amer. Statist. Assoc.*, **96**, 453, pages 270–281.

[Chipman et coll. 2001] H. Chipman, E. I. George et R. McCulloch (2001), « The practical implementation of Bayesian model selection », dans *Model selection*, volume 38 of *IMS Lectures notes – Monograph series*, pages 65–134, Institute of mathematical statistics.

[Clouchoux et coll. 2006] C. Clouchoux, O. Coulon, J.-L. Anton, , J.-F. Mangin et J. Régis (2006), « A new cortical surface parcellation model and its automatic implementation », dans *Proc. 9th MICCAI*, LNCS 4191, pages 193–200, Copenhaegen,, Springer Verlag.

[Cohen 1997] M. S. Cohen (1997), « Parametric analysis of MRI data using linear systems methods », *Neuroimage*, **6**, pages 93–103.

[Combettes et Wajs 2005] P.-L. Combettes et V. R. Wajs (2005), « Signal recovery by proximal forward-backward splitting », *Multiscale Model. Simul.*, **4**, pages 1168–1200.

[Cordes et coll. 2000] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, Q. M. A. et M. M.E. (2000), « Mapping functionally related regions of brain with functional connectivity MR imaging. », *AJNR Am. J. Neuroradiol.*, **21**, pages 1636-1644.

[Cordes et coll. 2002] D. Cordes, V. Haughton, J. Carew, K. Arfanakis et K. Maravilla (2002), « Hierarchical clustering to measure connectivity in fMRI resting-state data. », *Magn. Reson. Imaging*, **20**, 4, pages 305-17.

[Damoiseaux et coll. 2006] J. S. Damoiseaux, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith et C. F. Beckmann (2006), « Consistent resting-state networks across healthy subjects. », *Proc. Natl. Acad. Sci. USA*, **103**, 37, pages 13848–13853.

[Daubechies et coll. 2004] I. Daubechies, M. Defrise et C. Demol (2004), « An iterative thresholding algorithm for linear inverse problems with a sparsity constraint », *Comm. Pure Appl. Math.*, **57**, pages 1413-1457.

[Delyon et coll. 1999] B. Delyon, M. Lavielle et E. Moulines (1999), « Convergence of a stochastic approximation version of the EM algortihm », *The Annals of Statistics*, **27**, 1, pages 94–128.

[Dempster et coll. 1977] A. Dempster, A. Laird et D. Rubin (1977), « Maximum likelihood from incomplete data via the EM algorithm (with discussion) », *Journal of the Royal Statistical Society, Series B*, **39**, pages 1–38.

[Deneux et Faugeras 2006] T. Deneux et O. Faugeras (2006), « EEG-fMRI fusion of non-triggered data using Kalman filtering », dans *accepté pour publication Proc. 3th Proc. IEEE ISBI*, pages –, Arlington, VA, USA.

[D'Esposito et coll. 1999] M. D'Esposito, E. Zarahn, G. K. Aguirre et B. Rypma (1999), « The Effect of Normal Aging on the Coupling of Neural Activity to the Bold Hemodynamic Response », *Neuroimage*, **10**, 1, pages 6–14.

[Diedrichsen et Shadmehr 2005] J. Diedrichsen et R. Shadmehr (2005), « Detecting and adjusting for artifacts in fMRI time series data », *Neuroimage*, **27**, pages 624–634.

[Donnet et coll. 2006] S. Donnet, M. Lavielle et J.-B. Poline (2006), « Are fMRI event-related response constant in time ? A model selection answer. », *Neuroimage*, pages 1169–1176.

[El Kouby et coll. 2005] V. El Kouby, Y. Cointepas, C. Poupon, D. Rivière, N. Golestani, J.-B. Poline, D. L. Bihan et J.-F. Mangin (2005), « MR Diffusion-Based Inference of a Fiber Bundle Model from a Population of Subjects », dans *MICCAI'05*, pages 196–204, 2005.

[Escobar et West 1995] M. D. Escobar et M. West (1995), « Bayesian density estimation and inference using mixtures », *J. Amer. Statist. Assoc.*, **90**, 430, pages 577–588.

[Everitt et Bullmore 1999] B. S. Everitt et E. T. Bullmore (1999), « Mixture model mapping of brain activation in functional magnetic resonance images », *Hum. Brain Mapp.*, **7**, pages 1–14.

[Fadili et Bullmore 2002] J. Fadili et E. Bullmore (2002), « Wavelet-Generalized Least Squares : A New BLU Estimator of Linear Regression Models with 1/f Errors », *Neuroimage*, **15**, 1, pages 217–232.

[Fair et coll. 2007] D. A. Fair, N. U. F. Dosenbach, J. A. Church, A. L. Cohen, S. Brahmbhatt, F. M. Miezin, D. M. Barch, M. E. Raichle, S. E. Petersen et B. L. Schlaggar (2007), « Development of distinct control networks through segregation and integration. », *Proc Natl Acad Sci U S A*, **104**, 33, pages 13507–13512.

[Ferguson 1973] T. S. Ferguson (1973), « A Bayesian analysis of some nonparametric problems », *The annals of Statistics*, **1**, pages 209–230.

[Flandin 2004] G. Flandin (2004), *Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle*, thèse de doctorat, Université de Nice-Sophia Antipolis.

[Flandin et Penny 2007] G. Flandin et W. D. Penny (2007), « Bayesian fMRI data analysis with sparse spatial basis function priors », *Neuroimage*, **34**, 3, pages 1108–1125.

[Flandrin et P.GonÃ§alvÃ¨s 2004] P. Flandrin et P.GonÃ§alvÃ¨s (2004), « Empirical Mode Decomposition as Data-Driven Wavelet-Like Expansions », *Int. J. Wav. Multires. Info. Proc.*, **2**, 4, pages 477–496.

[Fox et coll. 2006] M. D. Fox, A. Z. Snyder, J. M. Zacks et M. E. Raichle (2006), « Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. », *Nat Neurosci*, **9**, 1, pages 23–25.

[Fox et coll. 2007] M. D. Fox, A. Z. Snyder, J. L. Vincent et M. E. Raichle (2007), « Intrinsic fluctuations within cortical systems account for intertrial variability in human evoked brain responses », *Proc. Natl. Acad. Sci. USA*, **56**, 1, pages 171–84.

[Frahm et coll. 1994] J. Frahm, K. Merboldt, W. Hanicke, A. Kleinschmidt et H. Boecker (1994), « Brain or vein – Oxygenation or flow ? on signal physiology in functional MRI of human brain activation », *NMR in Biomedicine*, **7**, pages 45–53.

[Frank et coll. 2001] L. Frank, R. Buxton et E. Wong (2001), « Estimation of respiration induced noise fluctuations from undersampled multislice fMRI data », *Magn. Reson. Med.*, **45**, pages 635–644.

[Friston 1994] K. Friston (1994), « Statistical Parametric Mapping », dans R. Thatcher, M. Hallet, T. Zeffiro, E. John et M. Huerta, éditeurs, *Functional Neuroimaging : Technical Foundations*, pages 79–93, 1994.

[Friston et coll. 2002] K. Friston, W. Penny, C. Phillips, S. Kiebel, G. Hinton et J. Ashburner (2002), « Classical and Bayesian Inference in Neuroimaging : Theory », *Neuroimage*, **16**, 2, pages 465–483.

[Friston et coll. 2005] K. Friston, K. Stephan, T. Lund, A. Morcom et S. Kiebel (2005), « Mixed-effects and fMRI studies », *Neuroimage*, **24**, pages 244–252.

[Friston et coll. 1995] K. J. Friston, A. P. Holmes, J.-B. Poline, P. Grasby, S. Williams, R. Frackowiak et R. Turner (1995), « Analysis of fMRI time-series revisited », *Neuroimage*, **2**, pages 45–53.

[Friston et coll. 1995] K. J. Friston, C. Frith, R. Turner et R. Frackowiak (1995), « Characterizing evoked hemodynamics with fMRI », *Neuroimage*, **2**, pages 157–165.

[Gelman et Meng 1998] A. Gelman et X.-L. Meng (1998), « Simulating normalizing constants : from importance sampling to bridge sampling to path sampling », *Statistical Science*, **13**, pages 163–185.

[Gelman et coll. 2004] A. Gelman, J. Carlin, H. Stern et D. Rubin (2004), *Bayesian data analysis*, Chapmann & Hall, London, UK, Second edition.

[Genovese 2000] C. Genovese (2000), « A Bayesian time-course model for functional magnetic resonance imaging data (with discussion) », *J. Amer. Statist. Assoc.*, **95**, pages 691–719.

[Geyer et Thompson 1995] C. Geyer et A. Thompson (1995), « Annealing Markov Chain Monte Carlo with applications to ancestral inference », *J. Amer. Statist. Assoc.*, **90**, pages 909-920.

[GLOVER 1999] G. H. GLOVER (1999), « Deconvolution of impulse response in event-related BOLD fMRI », *Neuroimage*, **9**, pages 416–429.

[GOLLAND et coll. 2007] P. GOLLAND, Y. GOLLAND et R. MALACH (2007), « Detection of spatial activation patterns as unsupervised segmentation of fMRI data. », *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, **10**, 1, pages 110–118.

[GOOD 2005] P. GOOD (2005), *Permutation, Parametric, and Bootstrap Tests of Hypotheses*, Springer, 3rd edition edition.

[GOTMAN et coll. 2004] J. GOTMAN, C. BENAR et F. DUBEAU (2004), « Combining EEG and fMRI in epilepsy : methodological challenges and clinical results », *Journal of Clin. NeuroPhysiol.*, **21**, pages 229–240.

[GOUTTE et coll. 2000] C. GOUTTE, F. A. NIELSEN et L. K. HANSEN (2000), « Modeling the haemodynamic response in fMRI using smooth filters », *IEEE Trans. Med. Imag.*, **19**, 12, pages 1188–1201.

[GREEN 1995] P. J. GREEN (1995), « Reversible jump MCMC computation and Bayesian model determination », *Biometrika*, **82**, pages 711–732.

[GREEN et RICHARDSON 2002] P. J. GREEN et S. RICHARDSON (2002), « Hidden Markov models and desease mapping », *J. Amer. Statist. Assoc.*, **97**, 460, pages 1–16.

[GRISWOLD et coll. 2000] M. A. GRISWOLD, P. M. JAKOB, M. NITTKA, J. W. GOLDFARB et A. HAASE (2000), « Partially parallel imaging with localized sensitivities (PILS). », *Magn. Reson. Med.*, **44**, 4, pages 602–609.

[GRISWOLD et coll. 2005] M. A. GRISWOLD, M. BLAIMER, F. BREUER, R. M. HEIDEMANN, M. MUELLER et P. M. JAKOB (2005), « Parallel magnetic resonance imaging using the GRAPPA operator formalism. », *Magn. Reson. Med.*, **54**, 6, pages 1553–1556.

[GÖSSL et coll. 2001] C. GÖSSL, L. FAHRMEIR et D. P. AUER (2001), « Bayesian modeling of the hemodynamic response function in BOLD fMRI », *Neuroimage*, **14**, pages 140–148.

[HANDWERKER et coll. 2004] D. A. HANDWERKER, J. M. OLLINGER, et M. D'ESPOSITO (2004), « Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses », *Neuroimage*, **21**, pages 1639–1651.

[HANSEN 1992] P. HANSEN (1992), « Analysis of discrete ill-posed problems by means of the L-curve », *SIAM Rev.*, **34**, pages 561–580.

[HAYASAKA et NICHOLS 2003] S. HAYASAKA et T. NICHOLS (2003), « Validating Cluster Size Inference : Random Field and Permutation Methods », *Neuroimage*, **20**, 4, pages 2343–2356.

[HEIDEMANN et coll. 2001] R. M. HEIDEMANN, M. A. GRISWOLD, A. HAASE et P. M. JAKOB (2001), « VD-AUTO-SMASH imaging. », *Magn. Reson. Med.*, **45**, 6, pages 1066–1074.

[HERNANDEZ et coll. 2002] L. HERNANDEZ, D. BADRE, D. NOLL et J. JONIDES (2002), « Temporal Sensitivity of Event-Related fMRI », *Neuroimage*, **17**, pages 1018–1026.

[HIGDON et coll. 1997] D. M. HIGDON, J. E. BOWSHER, V. E. JOHNSON, T. G. TURKINGTON, D. R. GILLAND et R. J. JASZCZAK (1997), « Fully Bayesian Estimation of Gibbs Hyperparameters for Emission Computed Tomography Data », *IEEE Trans. Med. Imag.*, **16**, 5, pages 516–526.

[HIMBERG et coll. 2004] J. HIMBERG, A. HYVARINEN et F. ESPOSITO (2004), « Validating the independent components of neuroimaging time series via clustering and visualization. », *Neuroimage*, **22**, 3, pages 1214-22.

[HOGE et coll. 2004] W. S. HOGE, D. H. BROOKS, B. MADORE et W. KYRIAKOS (2004), « On the Regularization of SENSE and Space-RIP in Parallel MR Imaging », dans *Proc. 2th Proc. IEEE ISBI*, pages –, Arlington, VA, USA.

[HOLMES et coll. 1996] A. HOLMES, R. BLAIR, J. WATSON et I. FORD (1996), « Nonparametric analysis of statistic images from functional mapping experiments », *J. Cereb. Blood Flow Metab.*, **16**, pages 7–22.

[HUAIEN et PUTHUSSERYPADY 2007] L. HUAIEN et S. PUTHUSSERYPADY (2007), « fMRI data analysis with nonstationary noise models : a Bayesian approach », *IEEE Trans. Biomed. Eng.*, **54**, 9, pages 1621–1630.

[HUANG et coll. 1998] N. HUANG, Z. SHEN, S. LONG, M. WU, H. SHIH, Q. ZHENG, N. YEN, C. TUNG et H. LIU (1998), « The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis », *Proc. R. Soc. Lond. A*, **454**, pages 903-995.

[HUBER 1981] P. J. HUBER (1981), *Robust Statistics*, John Wiley, New York, NY, USA.

[HUETTEL et MCCARTHY 2001] S. HUETTEL et G. MCCARTHY (2001), « Regional differences in the refractoy period of the hemodynamic response : an event-related fMRI study », *Neuroimage*, **14**, pages 967–976.

[IDIER 2001] J. IDIER (2001), *Approche bayésienne pour les problèmes inverses*, Traitrement du signal et de l'image, Hermès, Paris,.

[IDIER 2001] J. IDIER (2001), « Convex half-quadratic criteria and interacting auxiliary variables for image restoration », *IEEE Trans. Image Processing*, **10**, 7, pages 1001–1009.

[ISHWARAN et JAMES 2001] H. ISHWARAN et L. F. JAMES (2001), « Gibbs sampling methods for stick-breaking priors », *J. Amer. Statist. Assoc.*, **96**, 453, pages 161–173.

[JACOBS et coll. 2008] J. JACOBS, C. HAWCO, E. KOBAYASHI, R. BOOR, P. LEVAN, U. STEPHANI, M. SINIATCHKIN, et J. GOTMAN (2008), « Variability of the hemodynamic response as a function of age and frequency of epileptic discharge in children with epilepsy », *Neuroimage*.

[JACQUEMOT, C .AND PALLIER et coll. 2003] C. JACQUEMOT, C .AND PALLIER, D. LEBIHAN, S. DEHAENE et E. DUPOUX (2003), « Phonological grammar shapes the auditory cortex : a functional magnetic resonance imaging study », *Journal of Neurosciences*, **23**, 29, pages 9541–9546.

[JAFFARD 2004] S. JAFFARD (2004), « Wavelet techniques in multifractal analysis », dans *Fractal Geometry and Applications : A Jubilee of Benoît Mandelbrot, M. Lapidus et M. van Frankenhuijsen Eds., Proceedings of Symposia in Pure Mathematics*, volume 72(2), pages 91–152, AMS, 2004.

[JAFFARD et coll. 2006] S. JAFFARD, B. LASHERMES et P. ABRY (2006), « Wavelet leaders in multifractal analysis », dans *Wavelet Analysis and Applications,T Qian, M. I; Vai, X. Yuesheng, Eds.*, pages 219–264, Basel, Switzerland, Birkhäuser Verlag.

[JAKOB et coll. 1998] P. M. JAKOB, M. A. GRISWOLD, R. R. EDELMAN et D. K. SODICKSON (1998), « AUTO-SMASH : a self-calibrating technique for SMASH imaging. SiMultaneous Acquisition of Spatial Harmonics. », *MAGMA*, **7**, 1, pages 42–54.

[JOSEPH et HENSON 1999] O. JOSEPH et R. N. HENSON (1999), « Event-related functional magnetic resonance imaging : modelling, inference and optimization », *Philos. Trans. Roy. Soc. London Ser. B : Biol.*, **354**, 1387, pages 1215–1228.

[KAIPIO et SOMERSALO 2005] J. KAIPIO et E. SOMERSALO (2005), *Statistical and computational inverse problems*, Applied mathematical sciences (160), Springer, USA.

[KASS ET RAFTERY 1995] R. E. KASS ET A. E. RAFTERY (1995), « Bayes factors », *J. Amer. Statist. Assoc.*, **90**, pages 773–795.

[KELLER et coll. 2008] M. KELLER, A. ROCHE, S. MÉRIAUX, B. THIRION, PERCEPT. PSYCHOPHY., ANNALES DES TÉLÉCOMMUNICATIONS ET M. LAVIELLE (2008), « Dealing with normalization errors in fMRI group uinference using hierarchical Bayesian modeling », en revision à statistica sinica, CEA/NeuroSpin and INRIA Saclay, Gif-sur-Yvette,.

[KHERIF et coll. 2002] F. KHERIF, J.-B. POLINE, G. FLANDIN, H. BENALI, O. SIMON, S. DEHAENE ET K. WORSLEY (2002), « Multivariate model specification for fMRI data », *Neuroimage*, **16**, 4, pages 1068–1083.

[KING 2001] K. F. KING (2001), « SENSE image quality improvement using matrix regularization », dans *Proc. 9th ISMRM*, page 1771, Glasgow.

[KORMYLO ET MENDEL 1982] J. J. KORMYLO ET J. M. MENDEL (1982), « Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes », *IEEE Trans. Inf. Theory*, **28**, pages 482–488.

[KRUGER ET GLOVER 2001] G. KRUGER ET G. GLOVER (2001), « Physiological noise in oxygenationsensitive magnetic resonance imaging », *Magn. Reson. Med.*, **46**, pages 631–637.

[KUHN ET LAVIELLE 2004] E. KUHN ET M. LAVIELLE (2004), « Coupling a stochastic approximation of EM with a MCMC procedure », *ESAIM P&S*, **8**, pages 115–131.

[KYRIAKOS et coll. 2000] W. E. KYRIAKOS, L. P. PANYCH, D. F. KACHER, C. F. WESTIN, S. M. BAO, R. V. MULKERN ET F. A. JOLESZ (2000), « Sensitivity profiles from an array of coils for encoding and reconstruction in parallel (SPACE RIP). », *Magn. Reson. Med.*, **44**, 2, pages 301–308.

[LABAT ET IDIER 2007] C. LABAT ET J. IDIER (2007), « Convergence of conjugate gradient methods with a closed-form stepsize formula », *J. Optim. Theory Appl.*, **135**, 2.

[LAHAYE et coll. 2003] P.-J. LAHAYE, J.-B. POLINE, G. FLANDIN, S. DODEL ET L. GARNERO (2003), « Functional connectivity : studying non-linear, delayed interactions between BOLD signal », *Neuroimage*, **20**, pages 962–974.

[LANGE 1997] N. LANGE (1997), « Empirical and Substantive Models, the Bayesian Paradigm, and Meta-Analysis in Functional Brain Imaging », *Hum. Brain Mapp.*, **5**, pages 259–263.

[LE BIHAN 2007] D. LE BIHAN (2007), « The 'wet mind' : water and functional neuroimaging », *Phys Med Biol*, **52**, 7, pages 57–90.

[LE BIHAN et coll. 2006] D. LE BIHAN, T. ASO, S. URAYAMA, T. HAMAKAWA ET H. FUKUYAMA (2006), « Direct and fast detection of neuronal activation in the human brain with diffusion MRI », *Proc. Natl. Acad. Sci. USA*, **103**, 21, pages 8263–8268.

[LEE et coll. 2005] J. LEE, J. HU, J. GAO, K. WHITE, B. CROSSON, C. WIERENGA, K. MCGREGOR ET K. PECK (2005), « identification of brain activity by fractal scaling analysis of functional MRI data », dans *30th Proc. IEEE ICASSP*, volume II, pages 137–140, Philadelphia, , USA.

[LEOPOLD et coll. 2003] D. LEOPOLD, Y. MURAYAMA ET N. LOGOTHETIS (2003), « Very slow activity fluctuations in monkey visual cortex : implications for functional brain imaging. », *Cereb. Cortex*, **13**, 4, pages 422-33.

[LEROUX et coll. 1998] P. LEROUX, R. J. GILLES, G. C. MCKINNON ET P. G. CARLIER (1998), « Optimized outer volume suppression for single-shot fast spin-echo cardiac imaging », *Journal of Magnetic Resonance Imaging*, **8**, pages 1022-1032.

[LI 1995] S. Z. LI (1995), « On Discontinuity-Adaptive Smoothness Priors in Computer Vision », *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-17**, 6, pages 576–586.

[LIN et coll. 2004] F. LIN, K. KWONG, J. BELLIVEAU ET L. WALD (2004), « Parallel imaging reconstruction using automatic regularization », *Magn. Reson. Med.*, **51**, pages 559-567.

[LIN et coll. 2005] F. LIN, T.-Y. HUANG, N.-K. CHEN, F.-N. WANG, S. M. STUFFLEBEAM, J. BELLIVEAU, L. WALD ET K. KWONG (2005), « Functional MRI using regularized parallel imaging acquisition », *Magn. Reson. Med.*, **54**, pages 343-353.

[LINDSAY 1983] B. G. LINDSAY (1983), « The Geometry of Mixture Likelihoods : a General Theory », *The Annals of Statistics*, **11**, 1, pages 86–94.

[LINKENKAER-HANSEN et coll. 2005] K. LINKENKAER-HANSEN, S. MONTO, H. RYTSALA, K. SUOMINEN, E. ISOMETSA ET S. KAHKONEN (2005), « Breakdown of long-range temporal correlations in theta oscillations in patients with major depressive disorder. », *J Neurosci*, **25**, 44, pages 10131–10137.

[LONG et coll. 2005] C. LONG, E. N. BROWN, L. W. C. TRIANTAFYLLOU, I. AHARON ET V. SOLO (2005), « Nonstationary noise estimation in functional MRI », *Neuroimage*, **28**, 4, pages 890-903.

[LU et coll. 2002] H. LU, X. GOLAY ET P. VAN ZIJL (2002), « Intervoxel heterogeneity of event-related functional magnetic resonance imaging responses as a function of T1 weighting », *Neuroimage*, **17**, 17, pages 943–955.

[LUCZAK ET VIGODA 2005] T. LUCZAK ET E. VIGODA (2005), « Torpid mixing of the Wang-Swendsen-Kotecká$\frac{1}{2}$ algorithm for sampling colorings. », *J. Discrete Algorithms*, **3**, 1, pages 92-100.

[MACKAY 2003] D. J. MACKAY (2003), *Information theory, inference and learning algorithms*, Cambridge university press.

[MALLAT 1998] S. MALLAT (1998), *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA.

[MANSFIELD 1977] P. MANSFIELD (1977), « Multi-planar image formation using NMR spin-echoes », *J.Phys.C : Solid State Phys.*, **10**, pages L55-L58.

[MANSFIELD et coll. 1995] P. MANSFIELD, R. COXON ET J. HYKIN (1995), « Echo-volumar imaging (EVI) of the brain at 3.0T : first normal volunteer and functional imaging results », *J. Computer Assisted Tomography*, **19**, 6, pages 847–852.

[MAXIM et coll. 2005] V. MAXIM, L. SENDUR, J. FADILI, J. SUCKLING, R. GOULD, R. HOWARD ET E. BULLMORE (2005), « Fractional Gaussian noise, functional MRI and Alzheimer's disease. », *Neuroimage*, **25**, 1, pages 141–158.

[MCKENZIE et coll. 2001] C. A. MCKENZIE, E. N. YEH ET D. K. SODICKSON (2001), « Improved spatial harmonic selection for SMASH image reconstructions. », *Magn. Reson. Med.*, **46**, 4, pages 831–836.

[MCKEOWN 2000] M. MCKEOWN (2000), « Detection of consitently task-related activation in fMRI data with hybrid independent component analysis », *Neuroimage*, **11**, pages 24–35.

[MENDEL 1983] J. M. MENDEL (1983), *Optimal Seismic Deconvolution*, Academic Press, New York, NY, USA.

[MENG ET RUBIN 1993] X.-L. MENG ET D. B. RUBIN (1993), « Maximum likelihood estimation via the ECM algorithm : A general framework », *Biometrika*, **80**, pages 267–278.

[MÉRIAUX et coll. 2006] S. MÉRIAUX, A. ROCHE, B. THIRION ET G. DEHAENE-LAMBERTZ (2006), « Robust statistics for nonparametric group analysis in fMRI », dans *Proc. 3th Proc. IEEE ISBI*, pages 936–939, Arlington, VA, USA.

[MEYER 2003] F. G. MEYER (2003), « Wavelet-based estimation of a semi-parametric generalized linear model of fMRI time series », *IEEE Trans. Med. Imag.*, **22**, 3, pages 315–322.

[Miezin et coll. 2000] F. M. Miezin, L. Maccotta, J. M. Ollinger, S. E. Petersen et R. L. Buckner (2000), « Characterizing the hemodynamic response : effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing », *Neuroimage*, **11**, pages 735–759.

[Miller et coll. 2007] K. L. Miller, D. P. B. H. Devlin, M. D. Robson, R. G. Wise, M. W. Woolrich, P. Jezzard et T. E. Behrens (2007), « Evidence for a vascular contribution to diffusion FMRI at high b-value », *Proc. Natl. Acad. Sci. USA*, **104**, 52, pages 20967–20972.

[Muzy et coll. 1993] J.-F. Muzy, E. Bacry et A. Arneodo (1993), « Multifractal formalism for fractal signals. The structure function method versus the wavelet transform modulus maxima method », *Phys. Rev. E*, **47**, pages 875–884.

[Mériaux et coll. 2006] S. Mériaux, A. Roche, G. Dehaene-Lambertz, B. Thirion et J.-B. Poline (2006), « Combined permutation test and mixed-effect model for group average analysis in fMRI. », *Hum. Brain Mapp.*, **27**, 5, pages 402–410.

[Newton et Raftery 1994] M. Newton et E. Raftery (1994), « Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion) », *J. R. Statist. Soc. B*, **56**, pages 3–48.

[Nichols et Hayasaka 2003] T. Nichols et S. Hayasaka (2003), « Controlling the Familywise Error Rate in Functional Neuroimaging : A Comparative Review », *Statistical Methods in Medical Research*, **12**, 5, pages 419–446.

[Nichols et Holmes 2002] T. Nichols et A. Holmes (2002), « Nonparametric Permutation Tests for Functional Neuroimaging : A Primer with Examples », *Hum. Brain Mapp.*, **15**, pages 1-25.

[Nikolova et Ng 2005] M. Nikolova et M. Ng (2005), « Analysis of half-quadratic minimization methods for signal and image recovery », *SIAM J. Sc. Comput.*, **27**, pages 937–966.

[Operto et coll. 2006] G. Operto, R. Bulot, J.-L. Anton et O. Coulon (2006), « Anatomically informed convolution kernels for the projection of fMRI data on the cortical surface », dans *Proc. 9th MICCAI*, LNCS 4191, pages 300–307, Copenhaegen,, Springer Verlag.

[Orbanz et Buhmann 2008] P. Orbanz et J. M. Buhmann (2008), « Nonparametric Bayesian image segmentation », *International Journal of Computer Vision*, **77**, pages 25–45.

[Ou et Golland 2005] W. Ou et P. Golland (2005), « From Spatial Regularization to Anatomical Priors in fMRI Analysis », dans *IPMI, Glenwood Springs, Colorado*, juil. 2005.

[Owen 2001] A. B. Owen (2001), *Empirical Likelihood*, CRC Press, New York.

[Paschos 2005] V. Paschos (2005), *Optimisation combinatoire 1 – concepts fondamentaux*, Informatique et systÃ¨mes d'information, HermÃ¨s, Paris,.

[Penny et coll. 2003] W. D. Penny, S. Kiebel et K. J. Friston (2003), « Variational Bayesian inference for fMRI time series », *Neuroimage*, **19**, 3, pages 727–741.

[Penny et coll. 2004] W. D. Penny, K. E. Stephan, A. Mechelli et K. J. Friston (2004), « Comparing dynamic causal models. », *Neuroimage*, **22**, 3, pages 1157-72.

[Penny et coll. 2005] W. D. Penny, N. Trujillo-Barreto et K. J. Friston (2005), « Bayesian fMRI time series analysis with spatial priors », *Neuroimage*, **23**, 2, pages 350–362.

[Perlbarg et coll. 2007] V. Perlbarg, P. Bellec, J.-L. Anton, M. PÃ©lÃ©grini-Issac, J. Doyon et H. Benali (2007), « CORSICA : correction of structured noise

in fMRI by automatic identification of ICA components », *Magn. Reson. Imaging*, **25**, 1, pages 35–46.

[PRUESSMANN et coll. 2001] K. PRUESSMANN, M. WEIGER, M. SCHEIDEGGER ET P. BOE-SIGER (2001), « Advances in sensitivity encoding with arbitrary k-space trajectories », *Magn. Reson. Med.*, **46**, 4, pages 638–651.

[PRUESSMANN et coll. 1999] K. P. PRUESSMANN, M. WEIGER, M. SCHEIDEGGER ET P. BOE-SIGER (1999), « SENSE : Sensitivity encoding for fast MRI », *Magn. Reson. Med.*, **42**, pages 952–962.

[PURDON ET WEISSKOFF 1998] P. L. PURDON ET R. M. WEISSKOFF (1998), « Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. », *Hum. Brain Mapp.*, **6**, 4, pages 239–249.

[RAFTERY et coll. 2007] A. E. RAFTERY, M. A. NEWTON, J. M. SATAGOPAN ET P. N. KRI-VITSKY (2007), « Estimating the integrated likelihood via posterior simulation using the harmonic mean identity », dans J. BERNARDO, M. BAYARRI, O. BERGER, A. DAVID, D. HECKERMANN, A. SMITH ET M. WEST, éditeurs, *Bayesian statistics 8*, pages 1–45, Oxford University Press, 2007.

[RAJAPAKSE et coll. 1998] J. C. RAJAPAKSE, F. KRUGGEL, J. M. MAISOG ET D. VON CRA-MON (1998), « Modeling Hemodynamic Response for Analysis of Functional MRI Time-Series », *Hum. Brain Mapp.*, **6**, pages 283–300.

[RIBÉS et coll. 2007] A. RIBÉS, C. POUPON, C. RABRAIT, D. LE BIHAN ET F. LETHIMONNIER (2007), « Tikhonov regularisation optimisation for PreLearn : effects on the detection of activation in functional MRI », dans *Proc. 15th ISMRM*, page 1757, Berlin.

[RICHARDSON ET GREEN 1997] S. RICHARDSON ET P. J. GREEN (1997), « On Bayesian analysis of mixtures with an unknown number of components (with discussion) », *J. R. Statist. Soc. B*, **59**, 4, pages 731–792.

[RIEDI 2003] R. H. RIEDI (2003), « Multifractal Processes », *in : "Theory and applications of long range dependence ", eds. Doukhan, Oppenheim and Taqqu*, pages 625–716.

[RIERA et coll. 2004] J. RIERA, J. WATANABE, I. KAZUKI, M. NAOKI, E. AUBERT, T. OZAKI ET R. KAWASHIMA (2004), « A state-space model of the hemodynamic approach : nonlinear filtering of BOLD signal », *Neuroimage*, **21**, pages 547–567.

[RIVIÈRE et coll. 2002] D. RIVIÈRE, J.-F. MANGIN, D. PAPADOPOULOS-ORFANOS, J.-M. MARTINEZ, V. FROUIN ET J. RÉGIS (2002), « Automatic recognition of cortical sulci of the Human Brain using a congregation of neural networks », *Medical Image Analysis*, **6**, 2, pages 77–92.

[ROBERT et coll. 2000] C. ROBERT, T. RYDEN ET D. TITERRINGTON (2000), « Bayesian inference in hidden Markov models through reversible jump Markov chain Monte Carlo », *J. R. Statist. Soc. B*, **62**, pages 57–75.

[ROBERT 2001] C. P. ROBERT (2001), *The Bayesian Choice. Second Edition*, Springer Texts in Statistics, Springer Verlag, New York, NY, USA.

[ROCHE et coll. 2007] A. ROCHE, S. MÉRIAUX, ET B. THIRION (2007), « Mixed-effects statistics for group analysis in fMRI : A nonparametric maximum likelihood approach », *Neuroimage*, **38**, pages 501–510.

[ROEBROECK et coll. 2005] A. ROEBROECK, E. FORMISANO ET R. GOEBEL (2005), « Mapping directed influence over the brain using Granger causality and fMRI. », *Neuroimage*, **25**, 1, pages 230–242.

[Rosen et coll. 1998] B. R. Rosen, R. L. Buckner et A. M. Dale (1998), « Event-related functional MRI : past, present and future. », *Proc. Natl. Acad. Sci. USA*, **95**, pages 773–780.

[Rousseeuw et Leroy 1987] P. Rousseeuw et A. Leroy (1987), *Robust regression and outlier detection*, Wiley.

[Saint Pierre 2003] G. Saint Pierre (2003), *Identification du nombre de composants d'un mélange gaussien par chaînes de Markov à sauts réversibles dans le cas multivarié ou par maximum de vraisemblance dans le cas univarié*, thèse de doctorat, Université Paul Sabatier-Toulouse III, Toulouse.

[Samorodnitsky et Taqqu 1994] G. Samorodnitsky et M. Taqqu (1994), *Stable non-Gaussian random processes*, Chapman and Hall, New York ISBN 0-412-05171-0.

[Schölkopf et Smola 2002] B. Schölkopf et A. J. Smola (2002), *Learning with kernels. Support vector machines, regularization, optimzation and beyond*, Adaptive computation and machine learning, The MIT Press, Cambridge, Massachussets, USA.

[Schwartz 1978] G. Schwartz (1978), « Estimating the Dimension of a Model », *Ann. Statist.*, **6**, pages 461–464.

[Shimizu et coll. 2004] Y. Shimizu, M. Barth, C. Windischberger, E. Moser et S. Thurner (2004), « Wavelet-based multifractal analysis of fMRI time series. », *Neuroimage*, **22**, 3, pages 1195–1202.

[Smith et coll. 1999] A. M. Smith, B. K. Lewis, U. E. Ruttimann, F. Q. Ye, T. M. Sinnwell, Y. Yang, J. H. Duyn et J. A. Frank (1999), « Investigation of Low Frequency Drift in fMRI Signal », *Neuroimage*, **9**, 5, pages 526–533.

[Smith et coll. 2003] M. Smith, B. Pütz, D. Auer et L. Fahrmeir (2003), « Assessing brain activity through spatial Bayesian variable selection », *Neuroimage*, **20**, pages 802-815.

[Snoussi et Idier 2006] H. Snoussi et J. Idier (2006), « Bayesian blind separation of generalized hyperbolic processes in noisy and underdeterminate mixtures », *IEEE Trans. Signal Processing*, **54**, 9, pages 3257–3269.

[Sodickson 2000] D. K. Sodickson (2000), « Tailored SMASH image reconstructions for robust in vivo parallel MR imaging. », *Magn Reson Med*, **44**, 2, pages 243–251.

[Soltysik et coll. 2004] D. A. Soltysik, K. K. Peck, K. D. White, B. Crosson et R. W. Briggs (2004), « Comparison of hemodynamic response nonlinearity across primary cortical areas. », *Neuroimage*, **22**, 3, pages 1117–1127.

[Speigelhalter et coll. 2002] D. J. Speigelhalter, N. G. Best, B. P. Carlin et A. van der Linde (2002), « Bayesian measures of model complexity and fit (with discussion) », *J. R. Statist. Soc. B*, **64, Part 4**, pages 583–639.

[Swendsen et Wang 1987] R. H. Swendsen et J. S. Wang (1987), « Nonuniversal critical dynamics in Monte Carlo simulations », *Physical Review Letters*, **58**, pages 86–88.

[Tana et coll. 2007] M. G. Tana, A. M. Bianchi, P. Vitali, F. Villani et S. Cerutti (2007), « The haemodynamic response to the interictal epileptic spikes », dans *à paraître dans 29th Proc. IEEE EMBC*, pages –, Lyon,.

[Tanner 1993] M. Tanner (1993), *Tools for statistical inference – Methods for the exploration of posterior distributions and likelihood functions*, Springer series in Statistics, Springer–Verlag, New-York, USA, 3rd ed. edition.

[Thirion et coll. 2006] B. Thirion, S. Dodel et J.-B. Poline (2006), « Detection of signal synchronizations in resting-state fMRI datasets. », *Neuroimage*, **29**, 1, pages 321–327.

[THURNER et coll. 2003] S. THURNER, E. WINDISCHBERGER, C. MOSER, P. WALLA et M. BARTH (2003), « Scaling laws and persistence in human brain activity », *Physica A*, **326**, 3, pages 511-521.

[THYREAU et coll. 2006] B. THYREAU, B. THIRION, G. FLANDIN et J.-B. POLINE (2006), « Anatomo-functional description of the brain : a probabilistic approach », dans *Proc. 31th Proc. IEEE ICASSP*, volume V, pages 1109–1112, Toulouse, France.

[TIKHONOV ET ARSENIN 1977] A. TIKHONOV et V. ARSENIN (1977), *Solutions of Ill-Posed Problems*, Winston, Washington, DC, USA.

[TONONI et coll. 1998] G. TONONI, G. EDELMAN et O. SPORNS (1998), « Complexity and coherency : integrating information in the brain », *Trends in cognitive neuroscience*, **2**, 12, pages 474–484.

[TORRES ET ABRY 2003] M.-E. TORRES et P. ABRY (2003), « Comparison of different methods for computing scaling parameter in the presence of trends », dans A. CORDOBA, éditeur, *14th BioIngineering argentin congress*, 2003.

[TOYODA et coll. 2008] H. TOYODA, K. KASHIKURA, T. OKADA, S. NAKASHITA, M. HONDA, Y. YONEKURA, H. KAWAGUCHI, A. MAKI et N. SADATO (2008), « Source of nonlinearity of the BOLD response revealed by simultaneous fMRI and NIRS », *Neuroimage*, **39**, 3, pages 997-1013.

[VAEVER HARTVIG ET JENSEN 2000] N. VAEVER HARTVIG et J. JENSEN (2000), « Spatial mixture modeling of fMRI data », *Hum. Brain Mapp.*, **11**, 4, pages 233–248.

[ZWAAG et coll. 2006] W. VAN DER ZWAAG, S. FRANCIS et R. BOTWELL (2006), « Improved Echo Volumar Imaging (EVI) for functional MRI », *Magn. Reson. Med.*, **56**, pages 1320–1327.

[VEIT ET IDIER 2007] T. VEIT et J. IDIER (2007), « Rééchantillonnage de l'échelle dans les algorithmes MCMC pour les problèmes inverses bilinéaires », dans *Actes du 21ᵉ colloque GRETSI*, pages 1233–1236, Troyes, France.

[VEITCH ET ABRY 1999] D. VEITCH et P. ABRY (1999), « A wavelet-based joint estimator of the parameters of long-range dependence », *IEEE Trans. on Info. Theory*, **45**, pages 878–897.

[VEITCH ET ABRY 2001] D. VEITCH et P. ABRY (2001), « A statistical test for the time constancy of Scaling Exponents », *IEEE Trans. on Sig. Proc.*, **49**, 10, pages 2325–2334.

[VOULTSIDOU et coll. 2005] M. VOULTSIDOU, M. VOULTSIDOU, S. DODEL et J. HERRMANN (2005), « Neural Networks Approach to Clustering of Activity in fMRI Data », *IEEE Trans. Med. Imag.*, **24**, 8, pages 987–996.

[WEIGER et coll. 2002] M. WEIGER, K. P. PRUESSMANN, R. OSTERBAUER, P. BARNERT, P. BOESIGER et P. JEZZARD (2002), « Sensitivity-encoded single-shot spiral imaging for reduced susceptibility artifacts in BOLD fMRI. », *Magn. Reson. Med.*, **48**, 5, pages 860–866.

[WENDT et coll. 2006] H. WENDT, S. ROUX et P. ABRY (2006), « Bootstrap for Log Wavelet Cumulant Based Multifractal Analysis », dans *Proceedings of of EUSIPCO*, Firenze, Italy.

[WENDT et coll. 2007] H. WENDT, P. ABRY et S. JAFFARD (2007), « Bootstrap for Empirical Multifractal Analysis with Application to Hydrodynamic Turbulences », *IEEE Signal Processing Mag.*, **24**, 4, pages 38–48.

[WOOLRICH ET BEHRENS 2006] M. WOOLRICH et T. BEHRENS (2006), « Variational Bayes inference of spatial mixture models for segmentation », *IEEE Trans. Med. Imag.*, **25**, 10, pages 1380–1391.

[Woolrich et coll. 2001] M. Woolrich, B. Ripley, M. Brady et S. Smith (2001), « Temporal autocorrelation in univariate linear modelling of fMRI data », *Neuroimage*, **14**, 6, pages 1370–1386.

[Woolrich et coll. 2004] M. Woolrich, M. Jenkinson, J. Brady et S. Smith (2004), « Fully Bayesian spatio-temporal modelling of fMRI data », *IEEE Trans. Med. Imag.*, **23**, 2, pages 213–231.

[Woolrich et coll. 2004] M. Woolrich, T. Behrens, C. Beckmann, M. Jenkinson et S. Smith (2004), « Multi-level linear modelling for fMRI group analysis using Bayesian inference », *Neuroimage*, **21**, 4, pages 1732–1747.

[Woolrich et coll. 2005] M. Woolrich, T. Behrens, C. Beckmann et S. Smith (2005), « Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data », *IEEE Trans. Med. Imag.*, **24**, 1, pages 1–11.

[Worsley 1994] K. Worsley (1994), « Local maxima and the expected Euler characteristic of excursion sets of χ^2, f, and t fields », *Adv. Appl. Prob.*, **26**, pages 13–42.

[Worsley et coll. 2002] K. Worsley, C. Liao, J. Aston, V. Petre, G. Duncan, F. Morales et A. Evans (2002), « A general statistical analysis for fMRI data », *Neuroimage*, **15**, 1, pages 1–15.

[Wotawa 2006] N. Wotawa (2006), *Low-level visual cortex and motion perception : an MRI study*, Phd thesis, Université de Nice Sophia-Antipolis, Nice, France.

[Wu 1983] C. F. J. Wu (1983), « On the convergence of the em algorithm », *Ann. Statist.*, **11**, 1, pages 95–103.

[Yang et coll. 1997] Y. Yang, V. Mattay, D. Weinberger, J. Frank et J. Duyn (1997), « Localized echo-Volume imaging methods for functional MRI », *Journal of Magnetic Resonance Imaging*, **7**, pages 371-375.

[Zarahn et coll. 1997] E. Zarahn, G. K. Aguirre et M. D'Esposito (1997), « Empirical analysis of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions », *Neuroimage*, **5**, 3, pages 179–197.

Cinquième partie

Publications et documents annexés

20 juin 2008

ESTIMATION NON-PARAMÉTRIQUE DE LA FONCTION DE RÉPONSE HÉMODYNAMIQUE

VII.1 Première approche

[a.6] G. Marrelec, H. Benali, P. Ciuciu, M. Pélégrini-Issac and J.-B. Poline « Robust Bayesian Estimation of the Hemodynamic Response Function in Event- Related BOLD fMRI Using Basic Physiological Information », *Human Brain Mapping*, vol. 19, pp. 1–17, mai 2003.

Robust Bayesian Estimation of the Hemodynamic Response Function in Event-Related BOLD fMRI Using Basic Physiological Information

Guillaume Marrelec,[1,4] **Habib Benali,**[1,4] **Philippe Ciuciu,**[2,4]
Mélanie Pélégrini-Issac,[3,4] **and Jean-Baptiste Poline**[2,4]

[1]*Institut National de la Santé et de la Recherche Médicale U494, Paris, France*
[2]*CEA, Service Hospitalier Frédéric Joliot, Orsay, France*
[3]*Institut National de la Santé, et de la Recherche Médicale U483, Paris, France*
[4]*Institut Fédératif de Recherche 49, Paris, France*

Abstract: In BOLD fMRI data analysis, robust and accurate estimation of the Hemodynamic Response Function (HRF) is still under investigation. Parametric methods assume the shape of the HRF to be known and constant throughout the brain, whereas non-parametric methods mostly rely on artificially increasing the signal-to-noise ratio. We extend and develop a previously proposed method that makes use of basic yet relevant temporal information about the underlying physiological process of the brain BOLD response in order to infer the HRF in a Bayesian framework. A general hypothesis test is also proposed, allowing to take advantage of the knowledge gained regarding the HRF to perform activation detection. The performances of the method are then evaluated by simulation. Great improvement is shown compared to the Maximum-Likelihood estimate in terms of estimation error, variance, and bias. Robustness of the estimators with regard to the actual noise structure or level, as well as the stimulus sequence, is also proven. Lastly, fMRI data with an event-related paradigm are analyzed. As suspected, the regions selected from highly discriminating activation maps resulting from the method exhibit a certain inter-regional homogeneity in term of HRF shape, as well as noticeable inter-regional differences. *Hum. Brain Mapping 19:1–17, 2003.* © **2003 Wiley-Liss, Inc.**

Key words: BOLD fMRI; hemodynamic response function; Bayesian analysis

INTRODUCTION

Discovered in the early 1990s, functional MRI (fMRI) has quickly become the leading method to

study hemodynamic changes in the brain in response to cognitive and behavioral tasks [Chen and Ogawa, 1999]. The relation between neural activity and the Blood Oxygen Level Dependent (BOLD) response is not yet clearly understood and is still under investigation [Vazquez and Noll, 1996; Buxton and Frank, 1997; Buxton et al., 1998; Li et al., 2000; Logothetis et al., 2001]. It is, therefore, convenient to model the various processes intervening in the brain, from reception of the stimulus to measurement of the BOLD contrast signal, as a whole system characterized by its transfer response function, the so-called Hemodynamic Response Function (HRF) [Friston et al., 1994].

The HRF is the theoretical signal that BOLD fMRI would detect in response to a single, very short stimulus of unit intensity. The key assumptions related to this model are the stationarity and linearity of the underlying physiological process. Such hypotheses are good approximations of the actual properties of the system as long as the inter-stimulus interval does not decrease below about two seconds [Dale and Buckner, 1997; Buckner, 1998].

Estimation of the HRF is a recent concern. Knowledge about the response function is believed to be a key issue to a better understanding of the underlying dynamics of brain activation and the relationship between brain areas [Biswal et al., 2000; Miezin et al., 2000]. HRFs are increasingly suspected to widely vary from region to region, from task to task, and from subject to subject [Aguirre et al., 1998; Buckner et al., 1998a,b; Miezin et al., 2000]. Unfortunately, precise and robust estimation of the HRF is still the subject of ongoing research, since the problem is badly conditioned, and various methods have been devised so far.

On the one hand, parametric methods assume that the HRF is a generally non-linear function of certain parameters that are to be estimated. These parameters are often bestowed with some physiological meaning. Such approaches have been applied to block or event-related stimuli. Function shapes that are typically used include Gaussian [Kruggel and von Cramon, 1999a, b; Kruggel et al., 2000] or spline-like [Gössl et al., 2001b]. Gössl et al. [2001a] use a parametric model on the temporal scale, whereas a more general prior is used on the spatial extension of the signal. Integration of a physiological model as prior information has also been considered to constrain parametric estimation of the HRF [Friston, 2002]. But assuming the shape of the hemodynamic response to be known a priori and invariant throughout the brain is a very strong constraint, since it fluctuates greatly.

On the other hand, non-parametric methods have been developed in an attempt to infer the HRF at each time sample. Such methods make no prior hypothesis about the shape of the response function. Since the low signal-to-noise ratio of fMRI data precludes direct voxelwise analysis (e.g. with averaging over time), more complex schemes have been proposed. Methods include: averaging over regions [Kershaw et al., 2000], selective averaging [Dale and Buckner, 1997], introduction of non-diagonal models for the temporal covariance of the noise [Burock and Dale, 2000], or introduction of smooth FIR filters [Goutte et al., 2000]. In a similar fashion, we recently proposed a Bayesian, non-parametric estimation of the HRF [Marrelec and Benali, 2001; Marrelec et al., 2001]. Relevant physio-

logical information was introduced to temporally regularize the problem and derive estimates of the HRF. This approach had the advantage of introducing no bias into the estimation, since the constraints imposed were clearly derived from physiological requirements. In Marrelec et al. [2001], the estimation features were based on a few examples and the authors' experience of the model. Real data consisted of the mean signals of BOLD fMRI measurements in a few regions of interest. Robust voxelwise analysis had, therefore, yet to be assessed.

In this report we quantify the performances of the estimation introduced. Simulations are used to analyze the behavior of the HRF estimator. When compared to the ML estimator, dramatic performance increase is actually proven. With these evaluations, we also show that robustness is achieved regarding the actual noise sampling distribution and the stimulus sequence.

The outline of the article is as follows. In the next section, we recall the theoretical background necessary for the understanding of the model treatment. We also develop a statistical tool to deal with model testing, including activation detection. In the third section, the major features of the model are assessed: importance of the prior, relevance of the actual noise structure and influence of the stimulus sequence. The method is finally applied to real data, where both HRF estimation and activation detection are performed on the same time series.

THEORETICAL BACKGROUND

Notations

In the following, x denotes a real number, $\boldsymbol{x}$ a vector, and $\boldsymbol{X}$ a matrix. "t" is the regular matrix transposition. $\boldsymbol{I}_N$ stands for the N-by-N identity matrix. "$\propto$" relates two expressions that are proportional. For two variables x and y, "$x|y$" stands for "x conditioned on y", or equivalently "x given y", and p(x) for the probability of x.

Model

Let $\boldsymbol{x} = (x_n)_{1 \leq n \leq N}$ be the time series of stimuli describing an experimental paradigm, and $\boldsymbol{y} = (y_n)_{1 \leq n \leq N}$ the corresponding BOLD fMRI time course of a voxel. A discrete linear model is assumed to hold between the stimulus vector $\boldsymbol{x}$ and the data $\boldsymbol{y}$:

$$(H) \quad y_n = \sum_{k=0}^{K} h_k x_{n-k} + \sum_{m=1}^{M} \lambda_m d_{m,n} + e_n$$

$$n = K + 1, \ldots, N.$$

The $(K + 1)$-dimensional vector $\boldsymbol{h} = (h_k)^{\mathrm{t}}$ represents the unknown HRF to be estimated. K is the order of the convolution model, and $L = N - K$ is the actual amount of data used in the calculation. $\boldsymbol{X} = (x_{n-k})$ is the regular L-by-$(K + 1)$ design matrix, consisting of the lagged stimulus covariates. The L-by-M matrix $\boldsymbol{D} = (d_{m,n})$ is a basis of M functions that takes a potential drift and any other nuisance effect into account, and the $\boldsymbol{\lambda} = (\lambda_m)^{\mathrm{t}}$ are the corresponding coefficients. For the sake of simplicity, the basis is assumed to be orthonormal, i.e., $\frac{1}{L}\boldsymbol{D}^{\mathrm{t}}\boldsymbol{D} = \boldsymbol{I}_L$. $\boldsymbol{e} = (e_n)^{\mathrm{t}}$ accounts for noise and is supposed to consist of independent and identically distributed Gaussian variables of unknown variance σ^2, assumed to be independent from the HRF. As will be shown in the simulation section, this assumption by no way requires that the sampling frequencies of the noise actually corrupting the data be normally distributed. In matrix form, (H) boils down to

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{D}\boldsymbol{\lambda} + \boldsymbol{e},$$

also called General Linear Model.

Bayesian Analysis With Temporal Prior

What is sought is estimation of the HRF $\boldsymbol{h}$ given the data $\boldsymbol{y}$. To cope with this issue, a suitable theoretical framework is required for dealing with information coming from various origins. On the one hand, the data follow a known mode, (H). The noise is also supposed to follow a definite (yet general) model, since it is Gaussian. On the other hand, it should be possible to take available information into account, in order to optimize the estimation. The problem faced being ill-conditioned, a priori knowledge about the HRF needs to be incorporated into the model in order to constrain it and enable coherent estimates. For doing so, Bayesian analysis imposes itself, allowing for robust yet flexible integration of a wide range of information types in a probabilistic framework.

Prior information

Since the underlying physiological process of BOLD fMRI is as of yet only partially understood, setting "hard" constraints on the HRF is most likely to introduce unwanted bias into the estimate. For this reason, we investigate basic and soft constraints that do not

contradict current knowledge. More precisely, the following is assumed:

(P0)   the HRF starts and ends at 0;
(P1)   the HRF is smooth.

These priors reflect that the underlying process evolves rather slowly on the experimental time scale. Our goal is then to translate this prior knowledge into information that can be directly implemented into a Bayesian analysis. First, prior (P0) can easily be introduced into the model by setting the first and last sample points of the HRF to 0, so that only $K - 1$ parameters (instead of $K + 1$) of the HRF are now unknown. Quantification of prior (P1) is achieved by setting a Gaussian prior for the norm of the second derivative of the HRF, whose relative weight is adjusted by a hyperparameter ϵ:

$$\mathrm{p}(\boldsymbol{h}|\sigma^2, \epsilon) \propto \left(\frac{\epsilon^2}{\sigma^2}\right)^{(K-1)/2} \exp\left[-\frac{\epsilon^2 \boldsymbol{h}^{\mathrm{t}} \boldsymbol{Q}\boldsymbol{h}}{2\sigma^2}\right], \qquad (1)$$

where

$$\boldsymbol{Q} = \frac{1}{\mathrm{TR}^4} \begin{pmatrix} 5 & -4 & 1 & 0 & & \cdots & & 0 \\ -4 & 6 & -4 & 1 & 0 & & & \\ 1 & -4 & 6 & -4 & 1 & 0 & & \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ & & & 0 & 1 & -4 & 6 & -4 & 1 \\ & & & & 0 & 1 & -4 & 6 & -4 \\ 0 & & \cdots & & & 0 & 1 & -4 & 5 \end{pmatrix}$$

is the $(K - 1)$-by-$(K - 1)$ concentration matrix of the Gaussian prior, chosen as the discrete second-order differentiation matrix. ϵ represents the relative weight of the prior probability compared to the likelihood of the data in the calculation of the posterior probability density function (pdf). The higher ϵ, the more the prior constraint is taken into account. On the contrary, a vanishing ϵ expresses that the solution comparatively integrates much more information from the data. The limiting case $\epsilon = 0$ yields results that are similar to the Maximum-Likelihood treatment (i.e., Bayesian with no specific prior).

Bayes' Theorem

Once the model and the prior information have been defined, the first step is to use Bayes' theorem

stating that, for a set of data compatible with the model:

$$p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y) = \frac{p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon) \cdot p(y | h, \boldsymbol{\lambda}, \sigma^2, \epsilon)}{p(y)}.$$

(2)

Since $p(y)$ is independent of h, $\boldsymbol{\lambda}$, σ^2, and ϵ, it is only a normalization factor that can be discarded from Equation (2), yielding

$$p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y) \propto p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon) \cdot p(y | h, \boldsymbol{\lambda}, \sigma^2, \epsilon). \quad (3)$$

This equation relates the prior information $p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon)$, the information brought by the data or likelihood $p(y | h, \boldsymbol{\lambda}, \sigma^2, \epsilon)$, and the information inferred a posteriori about the unknown parameters h, $\boldsymbol{\lambda}$, σ^2 and ϵ, $p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y)$. This posterior distribution contains all the knowledge about the parameters that can possibly be inferred from the data and the a priori information we have at hand.

Posterior pdf

Using the chain rule[†] and assuming no prior dependence between $\boldsymbol{\lambda}$, σ^2 and ϵ, as well as between h and $\boldsymbol{\lambda}$, the prior can be further expanded as

$$p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon) = p(h | \sigma^2, \epsilon) \cdot p(\boldsymbol{\lambda}) \cdot p(\sigma^2) \cdot p(\epsilon), \quad (4)$$

where $p(h | \sigma^2, \epsilon)$ has been defined in Equation (1). $p(\boldsymbol{\lambda})$, $p(\sigma^2)$, and $p(\epsilon)$ are classically set to uninformative priors (flat prior for $\boldsymbol{\lambda}$, Jeffreys priors for σ^2 and ϵ:

$$p(\boldsymbol{\lambda}) \propto \text{const} \quad p(\sigma^2) \propto (\sigma^2)^{-1} \quad p(\epsilon) \propto \epsilon^{-1}.$$

Assuming Gaussian noise, the likelihood rereads

$$p(y | h, \boldsymbol{\lambda}, \sigma^2) \propto (\sigma^2)^{-L/2} \exp\left[-\frac{1}{2\sigma^2} \| y - Xh - D\boldsymbol{\lambda} \|^2 \right].$$

(5)

Bringing Equations (4) and (5) together into Equation (3) leads to the posterior pdf for h, $\boldsymbol{\lambda}$, σ^2 and ϵ:

$$p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y) \propto \epsilon^{K-2}(\sigma^2)^{-[L + (K-1)/2 + 1]}$$

$$\times \exp\left[-\frac{1}{2\sigma^2} (\| y - Xh - D\boldsymbol{\lambda} \|^2 + \epsilon^2 h^t Q h) \right]. \quad (6)$$

[†]$p(\theta_1, \theta_2) = p(\theta_1 | \theta_2) \cdot p(\theta_2)$.

This distribution is the core of our inference, since any question concerning the problem can be answered by its manipulation and processing.

Marginal posterior pdf for h

In HRF estimation, though, the parameter of interest is usually h. In this case, all other parameters are only nuisance parameters whose estimation is not required, and all information relative to h is contained in the marginal posterior distribution of h, $p(h | y)$. This pdf can in turn be obtained from Equation (6) by integrating it with respect to the other parameters, according to the marginalization formula[‡]:

$$p(h | y) = \int p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y) \, d\boldsymbol{\lambda} \, d\sigma^2 \, d\epsilon.$$

Integrating $\boldsymbol{\lambda}$ and σ^2 is straightforward, resulting in

$$p(h | y) = \int p(h, \epsilon | y) \, d\epsilon = \int p(h | y, \epsilon) \cdot p(\epsilon | y) \, d\epsilon.$$

However, this integral cannot be calculated in closed form. A common way to circumvent the problem, as in [Friston et al., 2002a], is to estimate ϵ by $\hat{\epsilon}$ and approximate the sought density by

$$p(h | y) \approx p(h | y, \epsilon = \hat{\epsilon}).$$

This approximation holds if $p(\epsilon | y)$ is peaked enough around $\hat{\epsilon}$. Practically, checking its validity can be performed by examination of $p(\epsilon | y)$ (see, e.g., Fig. 6 for results on real data). $p(\epsilon | y)$ can then be approximated by a Dirac function and $p(h | y)$ by

$$p(h | y) \approx \int p(h | y, \epsilon) \cdot \delta(\epsilon - \hat{\epsilon}) \, d\epsilon = p(h | y, \epsilon = \hat{\epsilon}).$$

The strategy applied here is to first calculate the posterior pdf for the hyperparameter ϵ as

$$p(\epsilon | y) = \int p(h, \boldsymbol{\lambda}, \sigma^2, \epsilon | y) \, dh \, d\boldsymbol{\lambda} \, d\sigma^2,$$

[‡]$p(\theta_1) = \int p(\theta_1, \theta_2) \, d\theta_2$.

find an estimator $\hat{\epsilon}$ of ϵ and, then, approximate $p(h|y)$ by $p(h|y, \epsilon = \hat{\epsilon})$, which is calculated from the posterior pdf:

$$p(h|y, \epsilon = \hat{\epsilon}) = \frac{1}{p(\hat{\epsilon}|y)} \int p(h, \lambda, \sigma^2, \hat{\epsilon}|y) \, d\lambda \, d\sigma^2.$$

An approximation for the marginal posterior for σ^2 can also be calculated along the same lines. Using this scheme, it was shown in Marrelec et al. [2001] that

- ϵ follows a pdf that does not belong to any known family, but whose distribution is given by:

$$p(\epsilon|y) \propto \frac{\epsilon^{K-2}}{\sqrt{\det(X^tJX + \epsilon^2Q)}}$$
$$\times [y^t(I_L - J^tX(X^tJX + \epsilon^2Q)^{-1}X^t)Jy]^{-(L-M)/2}, \tag{7}$$

where $J = I_L - (1/L)DD^t$ is the projection matrix estimating and removing the nuisance trend from the data. Numerical calculation of this 1-dimensional pdf is straightforward, and an estimate can be inferred, such as the Maximum a posteriori (MAP):

$$\hat{\epsilon} = \arg \max p(\epsilon|y).$$

Choosing the mean instead of the MAP leads to similar results, as shown in Bretthorst [1992].

- $(\sigma^2|y, \epsilon = \hat{\epsilon})$ is scaled inverse-chi-square distributed, with $\nu = L - M$ degrees of freedom and scale parameter $s^2 = [y^t(I_L - J^tX(X^tJX + \hat{\epsilon}^2Q)^{-1}X^t)Jy]/\nu$. An estimator of σ^2 is given by

$$\hat{\sigma}^2 = \frac{\nu}{\nu - 2} s^2. \tag{8}$$

- $(h|y, \epsilon = \hat{\epsilon})$ is Student-t distributed with ν degrees of freedom, location parameter $\hat{h} = (X^tJX + \hat{\epsilon}^2Q)^{-1}X^tJy$ and scale matrix $V = s^2(X^tJX + \hat{\epsilon}^2Q)^{-1}$. The expectation of $(h|y, \epsilon = \hat{\epsilon})$ can be taken as an estimator for the HRF:

$$E[h|y, \epsilon = \hat{\epsilon}] = \hat{h} = (X^tJX + \hat{\epsilon}^2Q)^{-1}X^tJy. \tag{9}$$

Equation (9) with $\hat{\epsilon} = 0$ corresponds to the well-known Maximum-Likelihood estimate (ML estimate) or Ordinary Least Squares estimate (OLS estimate) commonly found in the literature [Mardia et al., 1979; Draper and Smith, 1981]. For $\hat{\epsilon} \neq 0$, this is the form of

a regularized estimator, with $\hat{\epsilon}$ playing the role of the regularization parameter. In a typical regularization-optimization process, one has to minimize a quantity that is the sum of a likelihood function and a regularization/penalization factor (e.g., the norm of the second derivative for smooth variations):

$$\|Xh - y\|^2 + \epsilon^2\|\partial^2 h\|^2.$$

In a Bayesian framework, the value of ϵ can automatically be estimated and set to the most probable value $\hat{\epsilon}$.

Divergence Tests on HRFs

Bayesian analysis has recently been applied to activation detection in fMRI data analysis [Friston et al., 2002a, b]. Another approach is to take advantage of the non-parametric framework developed in this study.

Once the estimation has been carried out as previously explained, it might be of interest to test whether a given function h_0 qualifies as a HRF in a voxel. For instance, if h_0 originates from a biological or physiological model, adequacy of this model with the experimental results can be tested. In a frequentist framework, this corresponds to testing against the null-hypothesis $(h = h_0)$. In other words, we test whether h is significantly different from h_0. h being Student-t distributed, the *deviance* of h_0 from model (H), defined as

$$\rho(h_0) = (h_0 - \hat{h})^tV^{-1}(h_0 - \hat{h}),$$

should be the realization of a $(K - 1) \cdot F_{K-1,\nu}$-distributed variable. As proposed in Tanner [1994], we, therefore, define the *deviance significance* $1 - \alpha_0$ of $(h = h_0)$ as

$$1 - \alpha_0 = \phi_{K-1,\nu}[\rho(h_0)] \tag{10}$$

where $\phi_{K-1,\nu}$ is the cumulative distribution function (cdf) of the $F_{K-1,\nu}$ distribution.

An interesting case of hypothesis testing occurs when h_0 is set to 0. The estimated HRF is then compared to a flat function, reflecting a model where the stimulus has no influence on the voxel signal, which is then nothing more than a baseline signal (drift and noise). This is nothing else than activation detection.

In this setting, it is hence possible to estimate the HRF *and* use the knowledge so gained to perform activation detection on the same dataset. This is pos-

sible, since the data are only used once, namely to infer the value of the HRF at each time sample. This information, contained in $p(h|y, \epsilon = \hat{\epsilon})$ and stating that h is t-distributed with parameters $\hat{h}$ and V, is in turn used to answer questions relative to certain characteristics of the HRF, such as, "What is the shape of the HRF," or "Is the response function significant?"

RESULTS FROM SYNTHETIC DATA

Materials and Methods

This section deals with the performance of the above estimations and focuses on the three following topics: importance of the temporal prior, relevance of the actual noise sampling distribution and influence of the stimulus sequence. Each feature was analysed using synthetic data. One thousand 224-point samples were simulated from the same original HRF h_0 ("canonical" HRF used by the SPM99 software[§]), stimulus sequence (one given realization of a random event-related stimulus) and quadratic drift as illustrated in Figure 1. Repetition time TR was set to 1.25 sec. The variance of the Gaussian noise σ^2 was successively set in $\{0.001, 0.005, 0.01, 0.05\}$, corresponding to SNRs[¶] given in Table I. For the analysis, K was set to 20 and quadratic drift was considered ($M = 3$).

Investigation of HRF estimation performance was assessed using three complementary criteria. First, the quadratic error $\eta_1(h)$ described how close the chosen estimator is to the real HRF. Second, variance score $\eta_2(h)$ was a measure of the uncertainty associated with the given estimator. Now, variance reduction is a desired feature only if the accuracy of the estimator increases consequently. As a matter of fact, a poor estimator (i.e., with high quadratic error) with a low variance is misleading and introduces a bias into the estimation. For instance, introduction of prior information into model (H) has a direct and logical consequence of decreasing the variance of the posterior pdf. By construction, the higher ϵ, the higher the variance reduction. Setting $\epsilon \to \infty$ even implies a vanishing variance, $|V| \to 0$, whereas the corresponding estimator $\hat{h}$ tends towards a flat function, which is obviously a very bad estimator of the true HRF. Bias estimation was, therefore, quantified by $\eta_3(\hat{h}, V)$: the smaller the bias, the more conservative the estimate.

Figure 1.
Simulations: **(i)** HRF h_0, **(ii)** paradigm, and **(iii)** quadratic drift.

Quadratic error was defined in a similar fashion as in Dale [1999]:

$$\eta_1(\hat{h}) = \frac{1}{K-1} \|\hat{h} - h_0\|^2. \tag{11}$$

[§]www.fil.ion.ucl.ac.uk/spm/spm99.html.
[¶]Defined as SNR $= 20 \log_{10}(\|Xh\|/\sqrt{L\sigma^2})$.

TABLE I. Simulations*

σ^2	SNR
0.001	16.39
0.005	9.40
0.01	6.39
0.05	−0.60

*Noise variances and corresponding SNRs for the HRF defined in Figure 1.

It is the average square error per time sample of the estimator compared to the true HRF h_0. Variance score was quantified by

$$\eta_2(V) = \frac{1}{K-1} \log \det(V). \tag{12}$$

As pointed out in Ruanaidh and Fitzgerald [1996], the determinant of the variance of a distribution has a simple interpretation in terms of hypervolume in a Gaussian approximation. The logarithm of this measure can then be related to an entropic measure.[¶] Finally, the bias was measured using the deviance of the real HRF h_0 from the model and Equation (10):

$$\eta_3(h_0) = \phi_{K-1,\nu}[\rho(h_0)]. \tag{13}$$

For each series of 1,000 simulations, the corresponding performance estimator was calculated on all the samples.

Importance of the Prior

We first compared a model with no a priori information corresponding to a Maximum-Likelihood estimation,** called (H_L), and the model with the temporal prior, (H_B). For typical simulations, Figure 2a represents true and estimated HRFs. Performance estimators were calculated for the 1,000 noise realizations using Equations (11), (12), and (13). The results are summarized in Figures 2b and c.

Figure 2b(i) clearly indicates that, regardless of the noise level, estimates of σ^2 were accurate for both models, showing the robustness of this estimator. Fig-

ure 2b(ii) shows that the relative spread of $\hat{\epsilon}$ in the Bayesian model increased with decreasing SNR. As for HRF estimation, benefits resulting from the introduction of a temporal prior were threefold. First, both models exhibited increasing quadratic error with increasing noise (Figure 2c(i)), but estimator $\hat{h}_B$ (corresponding to model (H_B)) was much more robust to increasing noise than $\hat{h}_L$ (corresponding to model (H_L)). Second, a dramatic decrease of variance was achieved when the prior was considered and, again, the lower the SNR, the larger the difference (Figure 2c(ii)). But this variance reduction was not the source of a bias in the estimation, since the deviance significance of model (H_B) was also improved compared to initial model (H_L), as can be seen on Figure 2c(iii).

Relevance of the Noise Sampling Distribution

According to Bretthorst [1999], the Gaussian structure of the noise in the model is a consequence of the Maximum-Entropy principle, in which only the mean and the variance of the actual noise are assumed to be known and relevant to the analysis. As such, the estimation should not depend on the sampling frequencies of the noise. This was also observed in Marrelec et al. [2001]. To confirm this, we simulated noise samples from various sampling distributions. First, in accordance with the model hypothesis, Gaussian noise was used with mean 0 and variance 0.01. In order to measure the robustness of the model with regard to the presence of temporal correlation in the noise, AR(4) with exponentially decreasing factors was also simulated.[††] Finally, physiological noise was considered as the BOLD fMRI signal of the real data used in the following section, selected in regions where no activation was detected. After every sample, the resulting time series was normed to get the same mean 0 and variance 0.01. Typical results and estimator performances are represented in Figure 3a–c.

As evidenced by the results depicted in Figure 3b(ii), estimation of hyperparameter ϵ varied relatively little with respect to the noise distribution: the MAP estimates were consistent with each other. As for the estimate of the noise variance, it was essentially independent from the noise structure (Figure 3b(i)). HRF estimation itself exhibited the same property. From the simulations, it obviously appeared that the actual sampling distribution of the noise is indeed of little importance (Figure 3c(i)–(iii)).

[¶]The entropy of a $\mathcal{N}(\mu, \Sigma)$ distribution is given by $S = \frac{1}{2} \log[2\pi \exp(1)\det(\Sigma)]$.

**In this case, the order ν changes from $K - 1$ to $K + 1$, the number of degrees of freedom changes from $L - M$ to $L - M - (K + 1)$, $\hat{\epsilon}$ is set to 0, and all formulas are modified accordingly [Marrelec et al., 2001].

[††]With equation $e_n = 0.3679e_{n-1} + 0.1353e_{n-2} + 0.0498e_{n-3} + 0.0183e_{n-4} + \epsilon_n$ and $\epsilon_n \approx \mathcal{N}(0, 0.01)$.
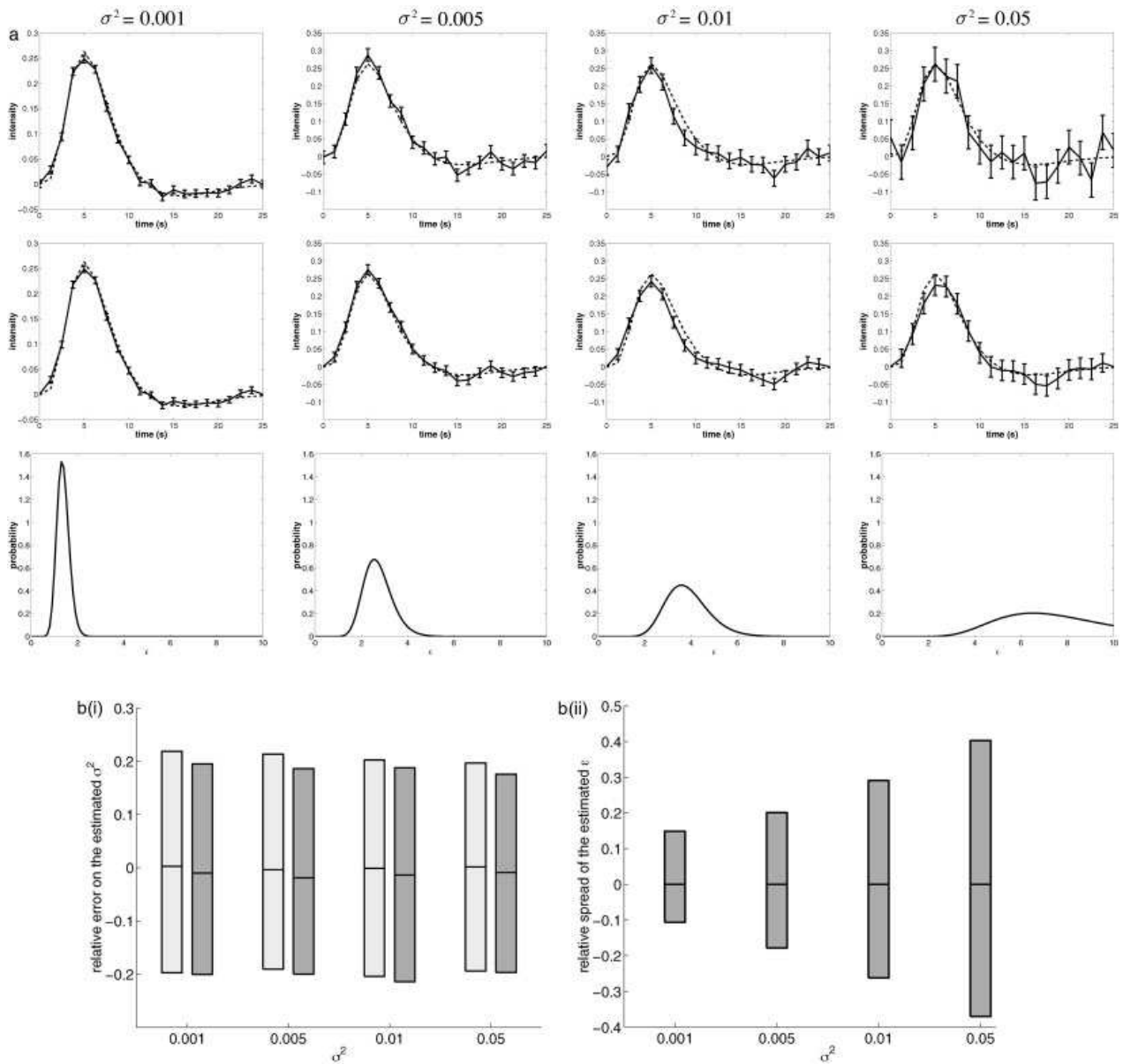
Figure 2.

Simulated data: importance of the prior. **a:** Typical results of simulations. Top and middle rows: simulated HRF (dotted line) and estimated HRF plus standard deviation (solid line) for the ML estimate (top) and the Bayesian estimate (middle). Bottom row: marginal pdf for ϵ. **b:** For each noise variance σ^2 in {0.001, 0.005, 0.01, 0.05}: **(i)** relative error of the noise estimator for the ML estimate (light gray) and the Bayesian estimate with prior (dark gray); **(ii)** relative spread of the estimated ϵ for the model with prior. The mean and the upper and lower 2.5% tails are represented, and the gray area represents the behavior of 95% of the data simulated. **c:** For each noise variance σ^2 in {0.001, 0.005, 0.01, 0.05}: **(i)** quadratic error η_1, **(ii)** variance score η_2, and **(iii)** deviance significance η_3 of the ML estimate (light gray) and the Bayesian estimate with prior (dark gray). The mean and the upper and lower 2.5% tails are represented, and the gray area represents the behavior of 95% of the data simulated.

Influence of the Stimulus Sequence

As pointed out in Buxton et al. [2000] and Worsley and Friston [1995], the choice of a stimulus sequence (periodic vs. no-periodic) is very important and can dramatically influence the power of an estimation method. To demonstrate the behavior of our technique and ensure that the method gives reliable results on
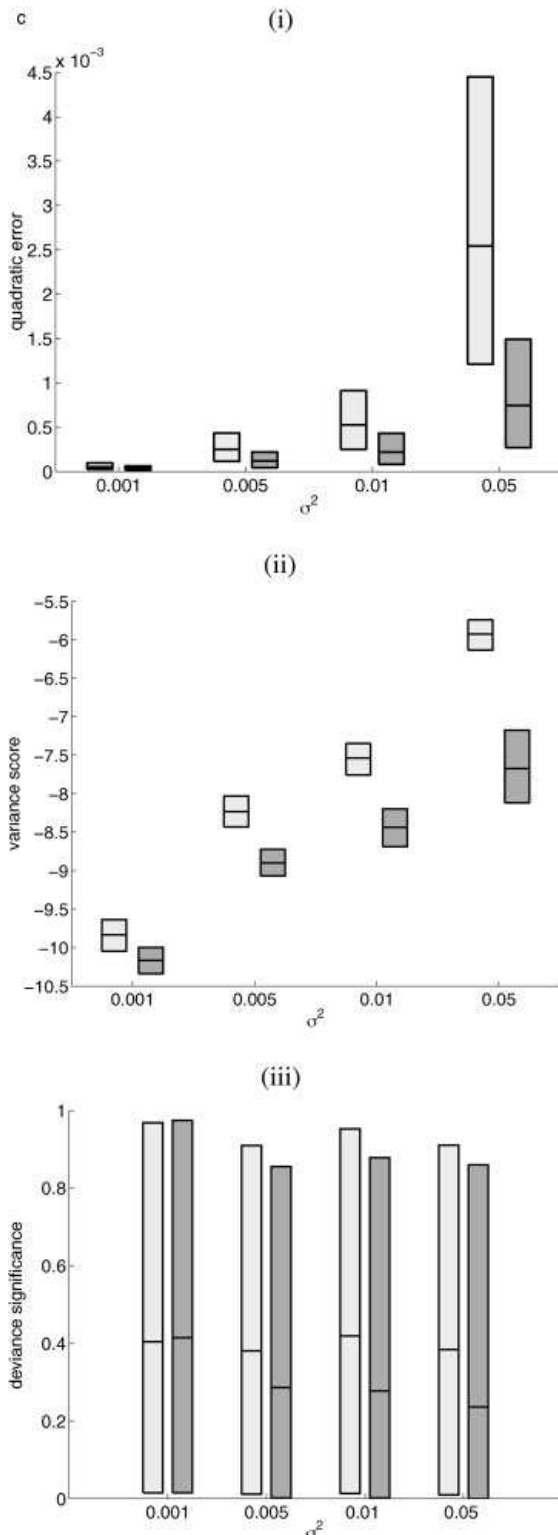
Figure 2.

the real data (see Results From Real Data), we compared estimates inferred from a simulation with a periodic vs. non-periodic stimulus. As in the data analyzed in this report, the periodic stimulus repeated itself every 10 s (corresponding to 8 TRs), and we estimated the HRF on 12.5 s (corresponding to *K* = 10). The results are summarized in Figure 4a–c and must be compared to the results in Figure 2a–c.

Our first conclusion is that Bayesian analysis is robust with regard to the stimulus sequence. Even though estimates were, as predicted, worse for a periodic stimulus sequence than in the case of non-periodic stimulus (Fig. 4c vs. Fig. 2c), they did not mislead us, since the variance increased consequently. The resulting bias is comparable to the case where the stimulus is non-periodic.

RESULTS FROM REAL DATA: VISUO-SPATIAL JUDGMENT TASK

Materials and Methods

Participants and task

Eleven healthy right-handed volunteers (age 24–35), with no neurological or psychiatric illness, gave written informed consent and were scanned, while performing the following visual task: they had to decide whether two visual dots flashed on the periphery of an 8-ray wheel projected on a screen were symmetrical with respect to the central fixation cross. The two dots were presented simultaneously for 150 ms every 10 seconds and their position had to be compared immediately. Subjects had to give a motor response by using a keypad (symmetrical: click with their right index finger; nonsymmetrical: click with their right middle finger). Participants were instructed to maintain eye fixation on the central cross throughout the whole experiment.

Data imaging and preprocessing

A 1.5 Tesla General Electric Signa imager (La Salpêtrière Hospital, Paris) with a standard head coil was used for the imaging. High resolution structural T1-weighted MPRAGE images were acquired from all participants for anatomical localization (0.9375 × 0.9375 × 1.5 mm). The functional images were produced by T2*-weighted echo-planar MRI at 8 contiguous 6-mm axial slices covering dorsal prefrontal and parietal regions (field of view: 24 cm, repetition time TR: 1.25 sec, echo time TE: 60 msec, flip angle: 90 degrees, 64 × 64 matrix of 3.75 × 3.75 mm voxels).
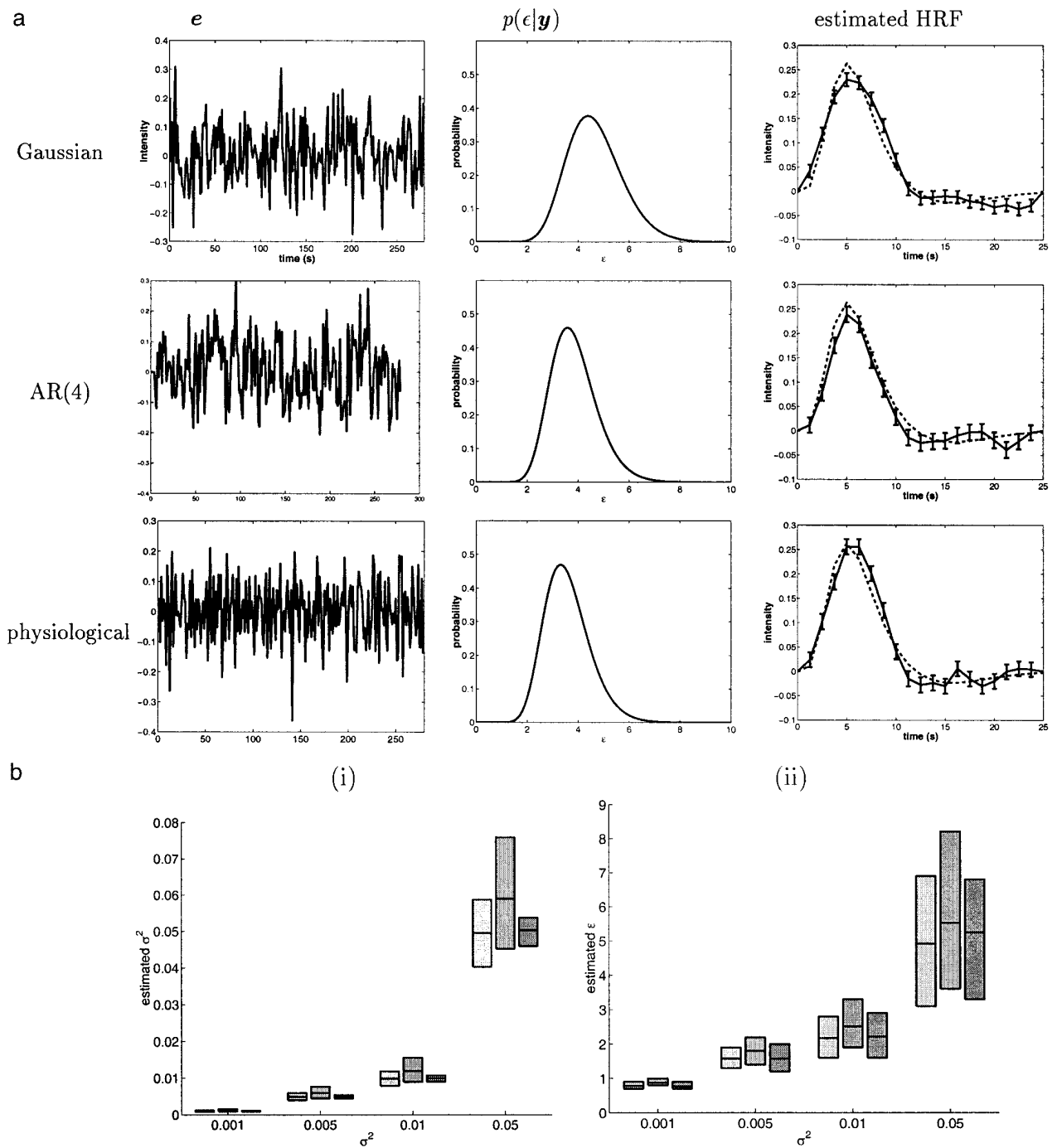
Figure 3.

Simulated data: relevance of the noise sampling distribution. **a:** Typical results of simulations with noise variance $\sigma^2 = 0.01$ and different noise sampling distributions. Right column: simulated HRF (dotted line) and estimated HRF with standard deviation (solid line). **b:** For each noise variance σ^2 in $\{0.001, 0.005, 0.01, 0.05\}$ (x-axis): **(i)** estimated noise and **(ii)** ϵ for Gaussian noise (light gray), AR(4) noise (middle gray) and physiological noise (dark gray). **c:** For each noise variance σ^2 in $\{0.001, 0.005, 0.01, 0.05\}$: **(i)** quadratic error, **(ii)** variance score, and **(iii)** deviance significance for Gaussian noise (light gray), AR(4) noise (middle gray) and physiological noise (dark gray). The mean and the upper and lower 2.5% tails are represented, and the gray area represents the behavior of 95% of the data simulated.

◆ Bayesian Estimation of **HRF** in **fMRI** ◆
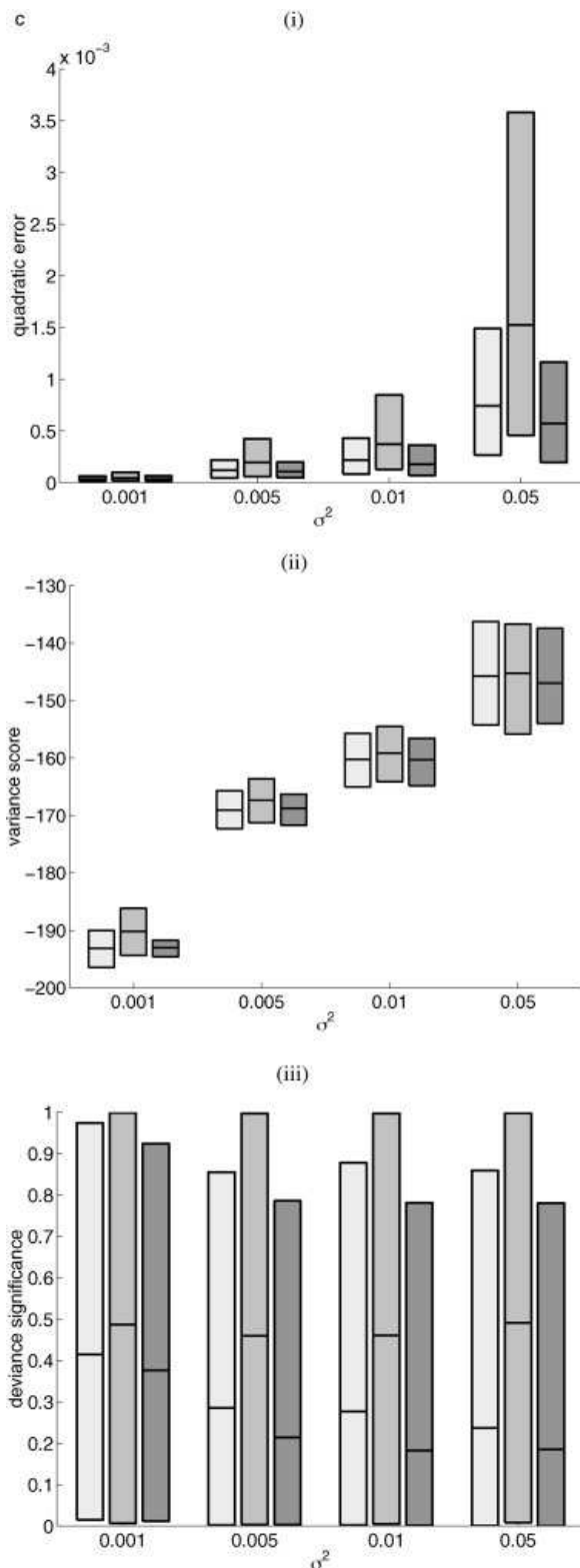
Figure 3.

Participants were studied in a single 224-scan session with a total duration of 4 min 40 sec. The scanner was in the acquisition mode for 20 sec before the experiment onset in order to achieve steady-state transverse magnetization. To compensate for subject motion, images were realigned to the middle image by using a rigid transformation and linear interpolation. The realigned images were filtered for low-frequency changes in BOLD signal over time by using high-pass filtering (namely, estimation of the baseline fluctuations using a moving average window and substraction of the estimated baseline from the input signal).

Data analysis

We estimated the HRF in each voxel. For the analysis, we set the order to $K = 10$, corresponding to a length of 12.5 sec. We also accounted for quadratic drift ($M = 3$). To handle significance levels, which were very high, we used the log-scale: for a significance level $1 - \alpha_0$, we therefore set $q_0 = -\log_{10}(1 - \alpha_0)$. We hereafter present the results from one subject.

Activation maps

Using the deviance test proposed in Equation (10), we defined voxel activation as the deviance of the zero HRF function ($h_0 = 0$) from the model.

The corresponding voxelwise activation map is represented in Figure 5(i). This map can be compared to Figure 5(ii), which was calculated by linear regression from a model with voxelwise adaptive Gaussian functions as proposed by Rajapakse et al. [1998]. It first appears that the two maps are comparable. On the other hand, the significance test developed in this study had a much higher discrimination level compared to the other method. As a matter of fact, regions where there should be no activation (such as the white matter) had a much lower significance in Figure 5(i) than in Figure 5(ii). Moreover, boundaries between activated and not-activated regions appeared much more clearly and sharply in Figure 5(i) than in Figure 5(ii). Potential activated regions can, therefore, be read off the map with ease. Whether such activated regions are indeed relevant is an issue that cannot be answered here, but high discrimination power is clearly a desired feature.

Regional stability

The activation map corresponding was thresholded to $1 - \alpha_0 = 1 - 10^{-11.5}$ (i.e., $q_0 = 11.5$). From this map, six

◆ **Marrelec et al.** ◆
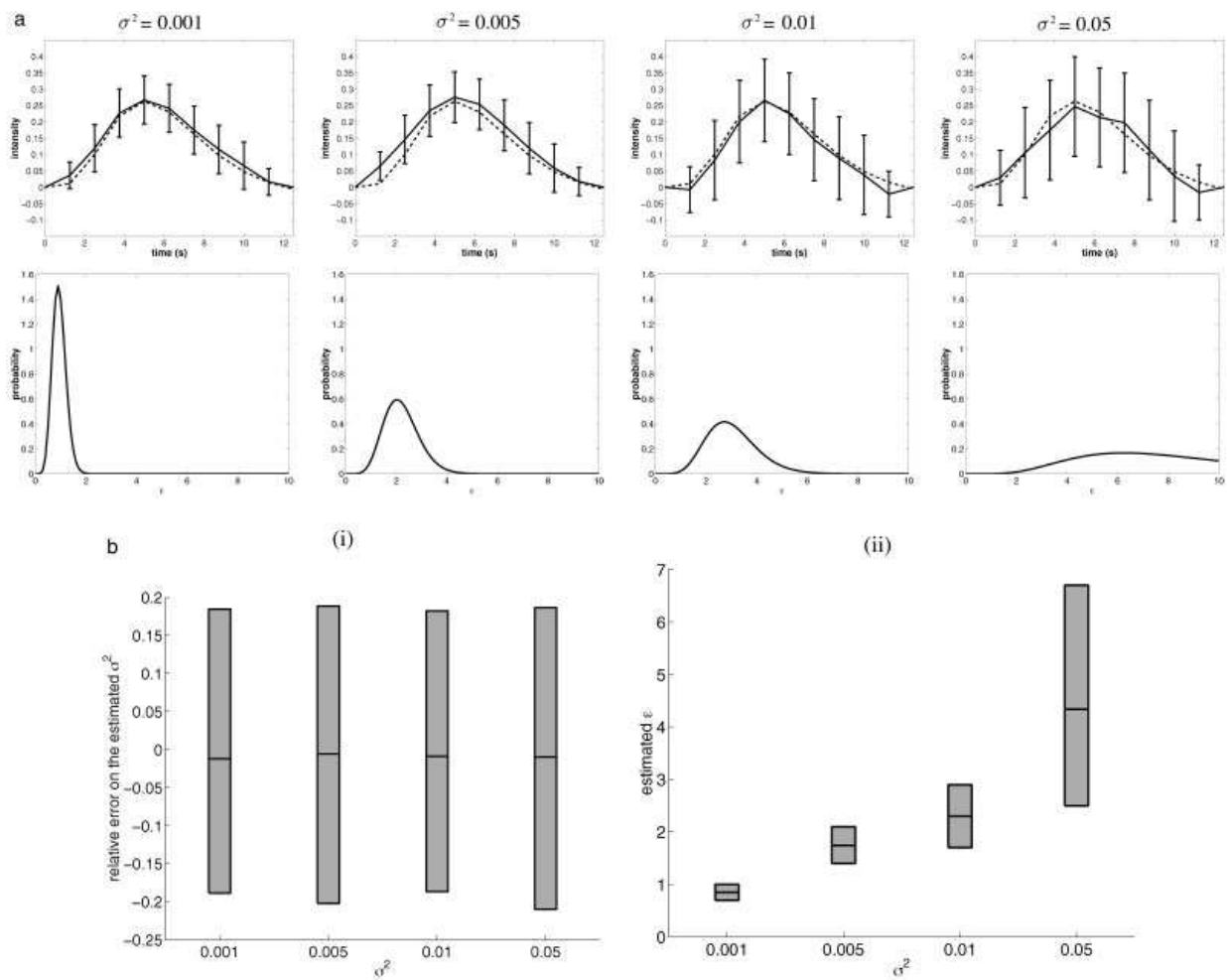
Figure 4.

Simulated data: influence of the stimulus. **a:** Typical results of simulations. Top row: simulated HRF (dotted line) and estimated HRF plus standard deviation (solid line). Bottom row: marginal pdf for ϵ. **b:** For each noise variance σ^2 in {0.001, 0.005, 0.01, 0.05}: **(i)** relative error of the noise estimator; **(ii)** estimated ϵ. The mean and the upper and lower 2.5% tails are represented, and the gray area represents the behavior of 95% of the data simulated. **c:** For each noise variance σ^2 in {0.001, 0.005, 0.01, 0.05}: **(i)** quadratic error η_1, **(ii)** variance score η_2, and **(iii)** deviance significance η_3. The mean and the upper and lower 2.5% tails are represented, and the gray area represents the behavior of 95% of the data simulated.

clusters were selected as shown in Figure 5(iii). For each cluster, Figure 6 represents the HRFs estimated in each voxel of the regions, the corresponding marginals for ϵ as well as the most peaked HRF.

Two main conclusions emerge from there. First, there is a clear idea of intracluster homogeneity. Indeed, the shape of the HRF seemed to be roughly constant within a region, "shape" meaning features of the curve such as increase/decrease, maximum/minimum or time-to-peak. However, it is not clear if this similarity is the consequence of physiological homogeneities, since parts of the resemblances may be due to non-physiological, intrinsic correlation of the signal, originating from the acquisition process or subject movement. On the other hand, the intensity of the response varied greatly in a given region, even though only highly significant voxels were taken into account. Second, HRFs did differ from region to region. They even seemed to be characteristic of the region involved. The differences concern the presence or absence of a post-stimulus undershoot, the presence or absence of a plateau, the pre- and post-maximum steepness, as well as the time-to-peak and the time-to-onset.
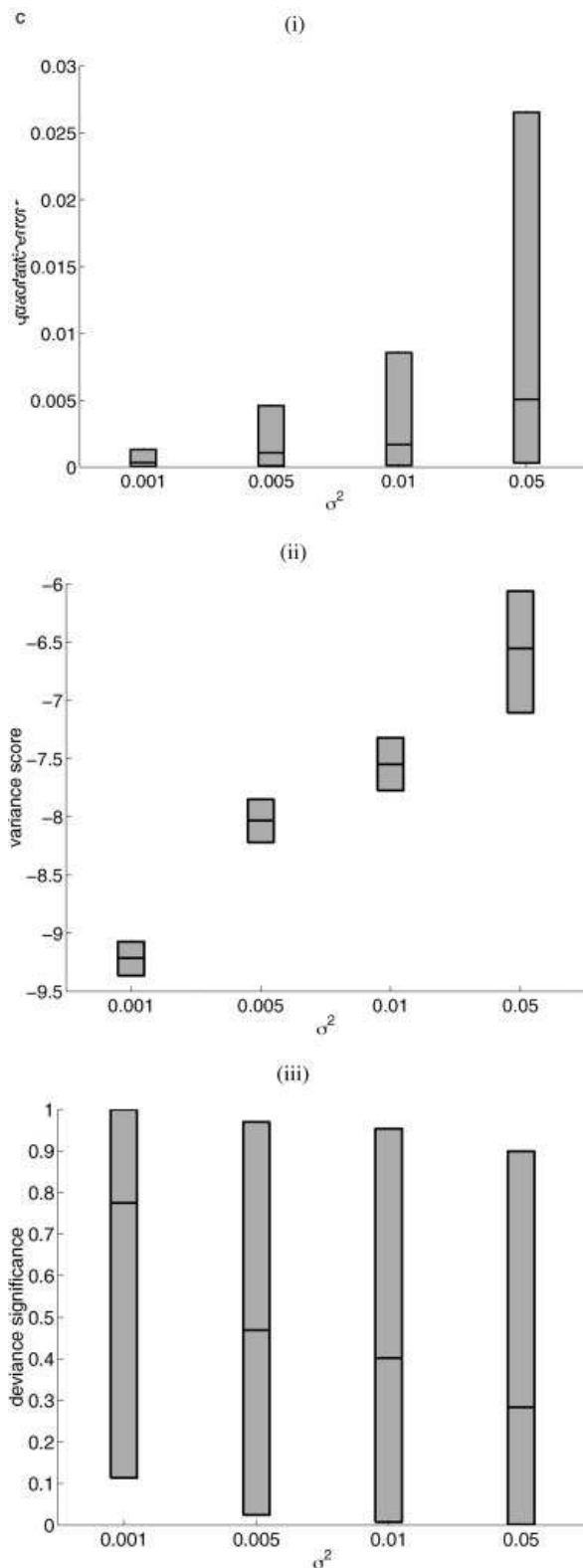
Figure 4.

DISCUSSION

The voxelwise HRF estimation technique that we proposed makes use of basic but relevant a priori information concerning the physiological process underlying the response. It proved to be reliable and robust regarding the actual noise level and structure, as well as the stimulus sequence.

Prior information and bias

Simulations comparing models with and without prior information clearly contradict common belief, which expects that introduction of prior information into analysis necessarily implies an increase of bias. In our case, introduction of a prior actually improved efficiency, variance, and bias at the same time. This is of course true given that the prior knowledge introduced into the model is respected. Estimation of peaked HRFs with this model would certainly give worse results.

Noise structure and estimation

The estimators introduced in this study were shown to be essentially insensitive to the true noise structure. This can be interpreted as follows. Two models were set: one for the HRF, and one for the noise. The latter was based on the sole hypothesis that the noise has given (yet unknown) mean and variance, and the Gaussian structure imposed itself as the least biased under this hypothesis. From there, two situations can happen. If the model for the HRF is sufficiently well defined (i.e., the prior information and the data are sufficient to lead correct inference), then the actual noise structure is mostly irrelevant to the estimation. In this case, introduction of more refined information (e.g., temporal correlation) would only slightly improve the estimation. On the other hand, if the model for the HRF is badly specified, then any additional information will greatly improve the results.

Noise level and smoothness

With decreasing SNR, $\hat{\epsilon}$ tends to be set to increasing values, giving more and more importance to the smoothing prior. Slow changes on long scales are then overfavored, and steep variations of the HRF seem to be smoothed out or rounded off (e.g., between 1 and 3 sec and around the peak in the simulations). However, the same simulations showed that our inference is still meaningful, since the mean ± standard deviation estimate stays accurate.
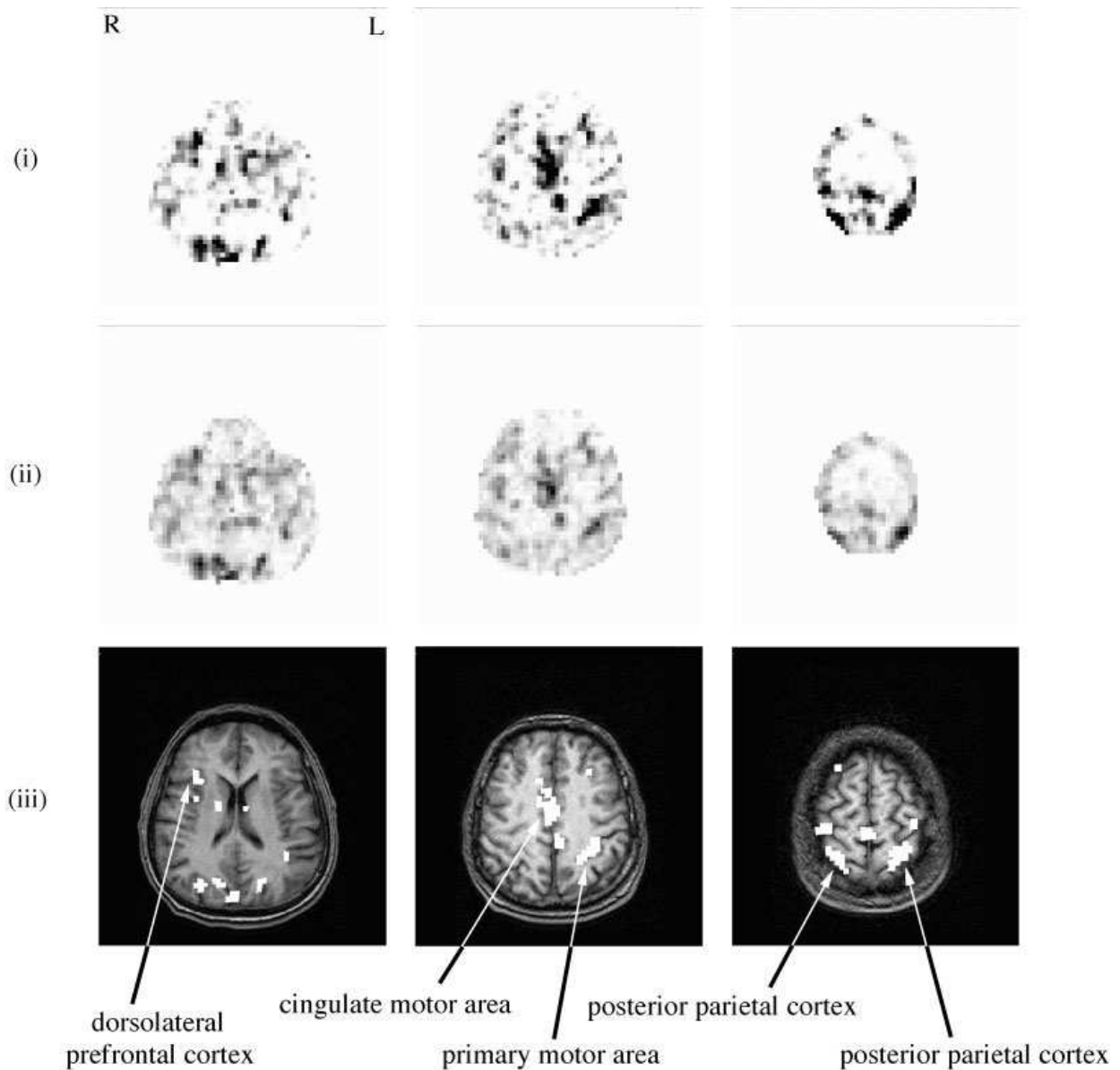
Figure 5.
Real data: activation maps. Activation maps from (**i**) the significance of the divergence test devised in this paper and (**ii**) the significance test by linear regression on adaptive Gaussian functions. Both maps have the same scale, between 0 (white) and 12 (black); (**iii**) anatomy and thresholded activation map from (i) ($q_0 > 11.5$).

Still with decreasing SNR, $p(\epsilon|y)$ becomes more and more diffuse around its peak: the model receives less and less information about the real value of ϵ from the data. Nonetheless, the simulations showed again that the MAP estimator for ϵ still makes sense for our purpose, since the resulting HRF estimates remain accurate. In this case, though, since the variance of ϵ is not considered, it is possible that the variance of h becomes more and more underestimated. This effect could possibly be taken into account (e.g., as proposed by Kaas and Steffey [1989]), at the cost of a more complicated and computationally time-consuming model. Whether this would significantly improve the inference is not
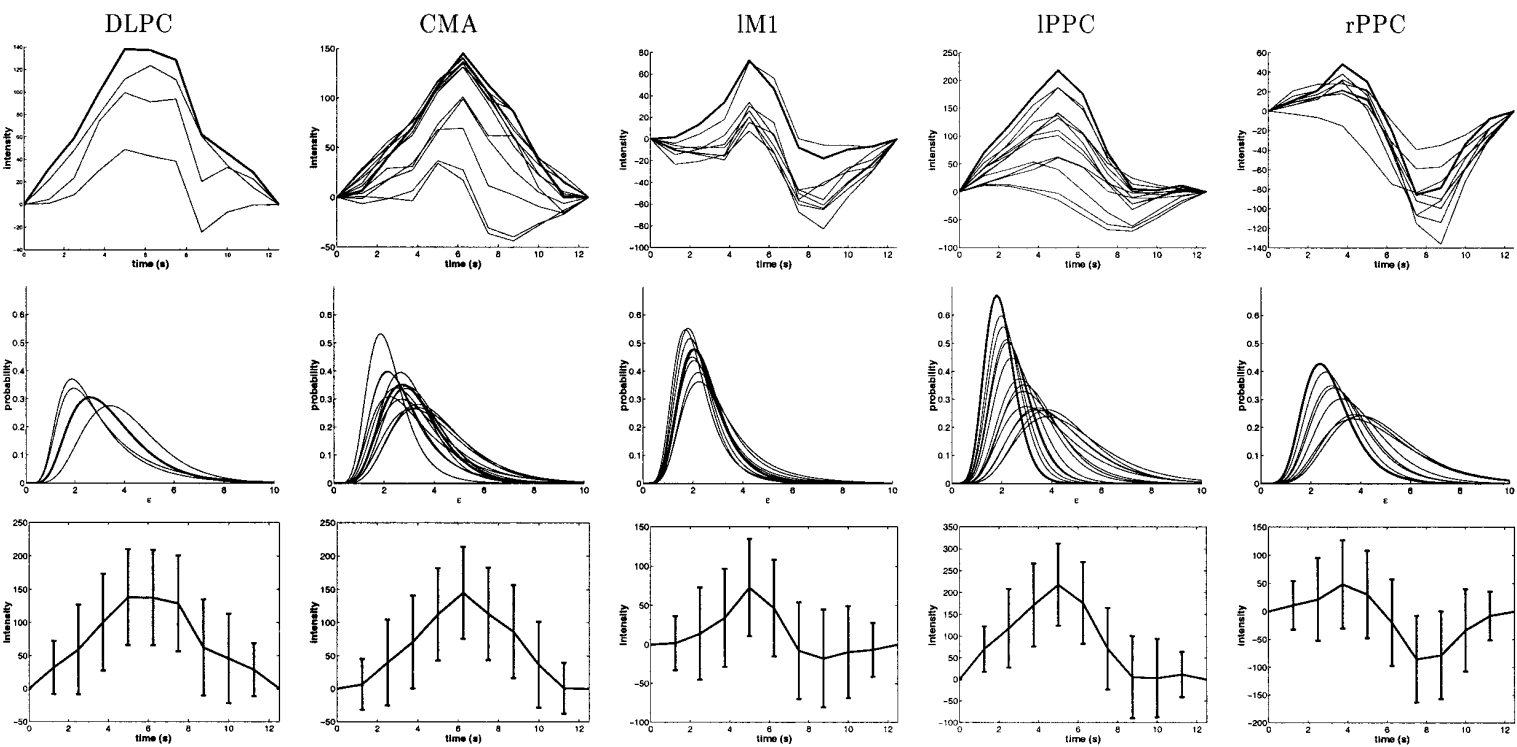
Figure 6.

Real data: HRFs (top) and marginal pdfs for ε (middle) for each voxel of the clusters shown in Figure 5 (iii). The HRF with highest peak among each region is also represented with standard deviation (bottom). DLPC: dorsolateral prefrontal cortex; CMA: cingulate motor area; lM1: left primary motor area; lPPC: left posterior parietal cortex; rPPC: right posterior parietal cortex.

quite clear yet, considered the good behavior of the estimators.

Convolution order

K was not estimated in our analysis but set to a certain value ($K = 20$ or $K = 10$). How did the choice of this parameter affect the analysis? Very little, indeed, if the stimulus sequence is of period higher than K or not periodic. In this case, setting K to a value that ensures a small HRF value gives satisfactory results. On the other hand, when the stimulus sequence is periodic, great care has to be taken. Giving K a value higher than the stimulus period implies that the model is not well determined. For this reason, ML estimators cannot be calculated. As we showed, the prior introduced regarding the smoothness of the HRF can somehow make up for this undeterminacy, but there are limits to this. In the simulation example we developed earlier, setting $K = 10$ is about as far as we could go without getting spurious effects.

CONCLUSION

This report provides an efficient and robust method to estimate the HRF and perform activation detection on the same dataset. The model integrates basic but relevant temporal information about the underlying physiological process of brain activation. Prior knowledge has proven to improve the accuracy and the robustness of the estimators. The actual structure of the noise and its level were shown to have little influence on the performance of the estimation. Simulations also showed that the estimators were robust to the stimulus sequence.

Highly discriminant activation maps were produced from the real data analyzed, as well as a wide variety of HRF shapes. The differences concerned the presence or absence of a post-stimulus undershoot, the presence or absence of a plateau, the pre- and post-maximum steepness, as well as the time-to-peak and the time-to-onset. Extra care has, therefore, to be taken when a fixed HRF is chosen and activation detection is performed, since no single function, whatever its characteristics, can account for activation throughout all the brain.

Ongoing research includes the search for new prior information and their translation in terms of constraints. It is also hypothesized that a more general resolution framework (e.g., integration of several stimuli, several sessions) is possible and would greatly improve the estimation.

REFERENCES

Aguirre GK, Zarahn E, D'Esposito M (1998): The variability of human BOLD hemodynamic responses. Neuroimage 7:S574.

Biswal B, Pathak A, Ward BD, Ulmer JL, Donahue KM, Hudetz AG (2000): Decoupling of the hemodynamic delay from the task-induced delay in fMRI. Neuroimage 11:S663.

Bretthorst GL (1992): Bayesian interpolation and deconvolution. Tech Rep CR-RD-AS-92-4. The U. S. Army Missile Command.

Bretthorst GL (1999): The near-irrelevance of sampling frequency distributions. In: Wolfgang von der Linden, et al., editors. Maximum entropy and Bayesian methods. Dordrecht: Kluwer. p 21–46.

Buckner RL (1998): Event-related fMRI and the hemodynamic response. Hum Brain Mapp 6:373–377.

Buckner RL, Koutstaal W, Schacter DL, Dale AM, Rotte M, Rosen BR (1998a): Functional-anatomic study of episodic retrieval using fMRI (II). Neuroimage 7:163–175.

Buckner RL, Koutstaal W, Schacter DL, Wagner AD, Rosen BR (1998b): Functional-anatomic study of episodic retrieval using fMRI (I). Neuroimage 7:151–162.

Burock MA, Dale AM (2000): Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. Hum Brain Mapp 11:249–260.

Buxton RB, Frank LR (1997): A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. J Cerebr Blood Flow Metab 17:64–72.

Buxton RB, Liu TT, Martinez A, Frank LR, Luh WM, Wong EC (2000): Sorting out event-related paradigms in fMRI: the distinction between detecting an activation and estimating the hemodynamic response. Neuroimage 11:S457.

Buxton RB, Wong EC, Frank LR (1998): Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn Reson Med 39:855–864.

Chen W, Ogawa S (1999): Principles of BOLD functional MRI. In: Moonen C, Bandettini P, editors. Functional MRI. Berlin: Springer. p 103–113.

Dale AM (1999): Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114.

Dale AM, Buckner RL (1997): Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 5:329–340.

Draper N, Smith H (1981): Applied regression analysis. Wiley Series in probability and mathematical statistics, 2nd ed. New York: John Wiley & Sons.

Friston KJ (2002): Bayesian estimation of dynamical systems: an application to fMRI. Neuroimage 16:513–530.

Friston KJ, Jezzard P, Turner R (1994): Analysis of functional MRI time-series. Hum Brain Mapp 1:153–171

Friston KJ, Glaser DE, Henson RNA, Kiebel S, Phillips C, Ashburner J (2002a): Classical and Bayesian inference in neuroimaging: applications. Neuroimage 16:484–512.

◆ **B**ayesian **E**stimation of **HRF** in **fMRI** ◆

Friston KJ, Penny W, Phillips C, Kiebel S, Hinton G, Ashburner J (2002b): Classical and Bayesian inference in neuroimaging: theory. Neuroimage 16:465–483.

Goutte C, Årup Nielsen F, Hansen LK (2000): Modeling the haemodynamic response in fMRI using smooth FIR filters. IEEE Trans Med Imag 19:1188–1201.

Gössl C, Auer D, Fahrmeir L (2001a): Bayesian spatiotemporal inference in functional magnetic resonance imaging. Biometrics 57:554–562.

Gössl C, Fahrmeir L, Auer DP (2001b): Bayesian modeling of the hemodynamic response function in BOLD fMRI. Neuroimage 14:140–148.

Kaas RE, Steffey D (1989): Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). J Am Stat Assoc 84:717–726.

Kershaw J, Abe S, Kashikura K, Zhang X, Kanno I (2000): A Bayesian approach to estimating the haemodynamic response function in event-related fMRI. Neuroimage 11:S474.

Kruggel F, von Cramon DY (1999a): Modeling the hemodynamic response in single-trial functional MRI experiments. Magn Reson Med 42:787–797.

Kruggel F, von Cramon DY (1999b): Temporal properties of the hemodynamic response in functional MRI. Hum Brain Mapp 8:259–271.

Kruggel F, Zysset S, von Cramon DY (2000): Nonlinear regression of functional MRI data: an item recognition task study. Neuroimage 12:173–183.

Li TQ, Haefelin TN, Chan B, Kastrup A, Jonsson T, Glover GH, Moseley ME (2000): Assessment of hemodynamic response during focal neural activity in human using bolus tracking, arterial spin labeling and BOLD techniques. Neuroimage 12:442–451.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001): Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150–157.

Mardia KV, Kent JT, Bibby JM (1979): Multivariate analysis. Probability and mathematical statistics. London: Academic Press.

Marrelec G, Benali H (2001): Non-parametric Bayesian deconvolution of fMRI hemodynamic response function using smoothing prior. Neuroimage 13:S194.

Marrelec G, Benali H, Ciuciu P, Poline JB (2001): Bayesian estimation of the hemodynamic response function in functional MRI. In: Fry R, editor. Bayesian inference and maximum entropy methods in science and engineering: 21st International Workshop. Melville: AIP. p 229–247.

Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000): Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11:735–759.

Rajapakse JC, Kruggel F, Maisog JM, von Cramon DY (1998): Modeling hemodynamic response for analysis of functional MRI time-series. Hum Brain Mapp 6:283–300.

Ruanaidh JJKO, Fitzgerald WJ (1996): Numerical Bayesian methods applied to signal processing. Statistics and computing. New York: Springer.

Tanner MA (1994): Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions. Springer Series in statistics, 2nd ed. New York: Springer.

Vazquez AL, Noll DC (1996): Non-linear temporal aspects of the BOLD response in fMRI. Proc ISMRM 1st Annual Meeting. p 1765.

Worsley KJ, Friston KJ (1995): Analysis of fMRI time-series revisited-again. Neuroimage 2:173–181.

VII.2 Extension multi-sessions multi-conditions

[a.3] P. Ciuciu, J.-B. Poline, G. Marrelec, J. Idier, C. Pallier et H. Benali, « Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment », *IEEE Trans. Med. Imag.*, vol. 22, 10, pp. 1235–1251, oct. 2003.

Unsupervised Robust Nonparametric Estimation of the Hemodynamic Response Function for Any fMRI Experiment

Philippe Ciuciu*, *Member, IEEE*, Jean-Baptiste Poline, Guillaume Marrelec, Jérôme Idier, Christophe Pallier, and Habib Benali

Abstract—**This paper deals with the estimation of the blood oxygen level-dependent response to a stimulus, as measured in functional magnetic resonance imaging (fMRI) data. A precise estimation is essential for a better understanding of cerebral activations. The most recent works have used a nonparametric framework for this estimation, considering each brain region as a system characterized by its impulse response, the so-called *hemodynamic response function* (HRF). However, the use of these techniques has remained limited since they are not well-adapted to real fMRI data. Here, we develop a threefold extension to previous works. We consider *asynchronous* event-related paradigms, account for different trial types and integrate several fMRI sessions into the estimation.**

These generalizations are simultaneously addressed through a badly conditioned observation model. Bayesian formalism is used to model temporal prior information of the underlying physiological process of the brain hemodynamic response. By this way, the HRF estimate results from a tradeoff between information brought by the data and by our prior knowledge. This tradeoff is modeled with hyperparameters that are set to the maximum-likelihood estimate using an expectation conditional maximization algorithm. The proposed unsupervised approach is validated on both synthetic and real fMRI data, the latter originating from a speech perception experiment.

Index Terms—**Bayesian estimation, ECM algorithm, event-related fMRI paradigm, HRF modeling.**

I. INTRODUCTION

DYNAMIC brain functional imaging was born in the last decade with functional magnetic resonance imaging (fMRI) [1]. For one subject, an fMRI experiment consists of the acquisition of a large number (100–1500) of three-dimensional (3-D) volumes (for instance, $64 \times 64 \times 32$ voxels, i.e., volume element) measuring in each voxel the BOLD contrast [2], which

is related to the total amount of deoxygenated hemoglobin present in *the* voxel. The subject is submitted to an experimental paradigm consisting of different conditions designed to study a particular brain system (e.g. memory, language, vision), while a continuous acquisition of brain volumes is performed. This is called a *run* or *session* and lasts for approximately 5–10 min. A session is generally repeated several times—typically between three and seven—for a given subject.

This technique has allowed to detect and localize dynamic brain processes for various stimulations or tasks [3] with a high spatial resolution (of the order of a millimeter), but a poor time resolution and a low signal or contrast-to-noise ratio (CNR) so far. This makes the use of well-designed data acquisition protocols necessary. Two classes of protocols can be distinguished: block-designed and event-related experiments. The latter has emerged as a means of observing the fMRI time course in response to a single, very short stimulus (a trial) [1], [4], while the former has better CNR and may require averaging over many trials presented in close succession. Despite their lower CNR, event-related paradigms are often inevitable, for instance, to avoid habituation effects. In such paradigms, random intermixing of trial types is used to eliminate habituation, anticipation or other strategy effects [1] that might occur in *deterministic* paradigms. Actual experiments consist of either *synchronous* or *asynchronous* paradigms, depending whether the onsets of the conditions are synchronized with the data acquisition rate or not. Paradigms are often asynchronous because the onset of the response can be given by the subject himself (response after a variable reaction time).

The end goal of activation detection in brain functional imaging experiments is to retrieve as much as possible of the neuronal activity in response to cognitive or behavioral tasks [3]. However, the relation between this activity and the BOLD response [2] is not completely understood and still under study [5]–[8]. The partially known mechanisms of coupling neuronal (synaptic) activity to the vascular system produces significant blurring and delay to the original neuronal response over time, indicating that the BOLD sequence is heavily low-pass filtered [5], [8]. As a consequence, hemodynamic events have time scales of a few seconds, whereas neuronal events have time constants of milliseconds. An accurate and robust estimate of the brain hemodynamic response to a stimulus may be a first step toward a better quantification of the brain neuronal activity. In each region, the brain hemodynamic response can be characterized in the first instance by the response function

to a very short stimulus (the transfer function if the system is linear), denoted as the *hemodynamic response function* (HRF).

Modeling the HRF has become an intensive topic of research for many reasons. First, a precise modeling should lead to a better understanding of cerebral activations. Second, within a region, the signal variations between conditions or stimuli (such as the magnitude of the response, but also its delay or width) can only be studied with an accurate estimate of the response and of its variability. Third, estimation of the HRF can be done at every position in the brain in order to investigate its spatial variability. Finally, some recent technological progress gives access to si-multaneous recordings of electrical (electroencephalography) and metabolic (fMRI) activities. For these reasons, the HRF has been the subject of many studies that usually assume that the brain system is linear and time invariant (LTI) [9]–[16]. Al-though the question whether the brain response can be consid-ered linear is not yet fully answered, it has been shown that this assumption is a tenable and useful approximation [4], [10], [14] and, thus, holds in the present work.

Parametric methods for estimating the HRF as a transfer function of a LTI system appeared first in the literature [9], [11]–[13]. These approaches impose the shape of the HRF by choosing a particular function (e.g.,Gamma or Gaussian density). The *nonlinear* parameters of this function are fitted to the data to take variations of the delay and blurring effects of the HRF into account. Parametric models may introduce some bias on the HRF, since it is unlikely that they capture the shape variations of the HRF within the brain.

By contrast, recent works have introduced temporal prior information on the underlying physiological process of the brain hemodynamic response to accurately estimate the HRF in a Bayesian framework. Such priors compensate for the lack of information provided by the data [15], [17], [18]. These techniques only apply to periodic or *synchronous* event-related paradigms and are devoted to the estimation of *one* HRF in response to one condition or stimulus. They also deal with each session separately and average the HRF estimates a posteriori without taking fluctuations of physiological factors across sessions into account. Hence, their use has remained limited since they are not well-adapted to actual fMRI data.

The aim of this paper is to propose a threefold efficient gener-alization of [15], [18]. First, we derive a temporally regularized estimator of the HRF when shorter and jittered interstimulus intervals (ISIs) are used, such as in asynchronous paradigms. Second, we propose a simple extension that is able to cope with *mixed task* paradigms, in which mixed trial stimuli are presented in a random order and in a rapid succession to one another. This extension consists in estimating a HRF per trial type. Third, we develop an estimation procedure that is able to simultaneoulsy process all fMRI time series (all sessions or runs) recorded for a subject in a given region. The specific treatment of each session is important because noise characteristics (low frequencies) may be different between sessions. We will show that this leads to more accurate estimates and relevant error bars provided that the drift terms are modeled with few parameters per session (typi-cally three or four). Because these extensions require a signifi-cant computational effort, we develop a powerful optimization scheme that makes the computation fast enough (typically one

second for a single time series of 200 samples) for the analysis of real data in an imaging center environment.

The rest of the paper is organized as follows. Section II starts with the introduction of the LTI model of the HRF for *single trial* asynchronous paradigms. Successive generalizations for *asynchronous multitrial multisession* paradigms are then taken into account in a more complex badly conditioned observation model. In Section III, we motivate our modeling choices and derive the selected estimator for the HRF within the Bayesian formalism. Since such an estimate depends on a few hyperpa-rameters, Section IV adresses the problem of their tuning ac-cording to the maximum-likelihood (ML) estimator. Section V illustrates the performances of our *unsupervised* approach with synthetic data. The method is applied to a language comprehen-sion fMRI study in Section VI. In Section VII, we discuss the limits of applicability and possible extensions of the proposed method.

II. MODELING THE CEREBRAL HEMODYNAMIC RESPONSE

A. Notations

Throughout the paper, random variables and realizations of thereof are respectively denoted by uppercase (e.g., $\boldsymbol{H}$) and corresponding lowercase (e.g., $\boldsymbol{h}$) symbols; in addition, notations such as $p(\boldsymbol{y}|\boldsymbol{h})$ are employed as shorthands for $p_{\boldsymbol{Y}|\boldsymbol{H}}(\boldsymbol{y}|\boldsymbol{h})$, whenever unambiguous. Furthermore, $p(\boldsymbol{y}|\boldsymbol{h};\boldsymbol{\theta})$ means that the parameter vector $\boldsymbol{\theta}$ is assumed unknown but deterministic. The main notations used in the following are summarized in Table I.

B. LTI System for Asynchronous Paradigms

In event-related protocols with synchronous ISI, the BOLD fMRI time course $(y_n)_{1\leq n\leq N}$ is measured in any voxel of the brain at times $(t_n = n\text{TR})_{1\leq n\leq N}$, TR being the time of repe-tition, while stimuli occur with a fixed-delayed impulse signal $(x_n)_{1\leq n\leq N}$. This means that the sampling period is equal to TR when the stimuli occur only at times of acquisition. The HRF is then modeled as the convolution kernel of a LTI system [9], [15], [18].

In *asynchronous* experiments, the presented stimuli occur at any time during scanning. In such cases, we propose to put the data and the trials on a finer temporal grid, which has to be defined such that *(i)* the time occurrences of the stimuli $(x_t)_{t\geq t_0}$ are defined on this grid, and that *(ii)* two stimuli do not occur at the same time. Let Δt be the sampling period of this grid. Our strategy consists in approximating the true onsets by their closer neighbor on this grid. This very simple procedure can be seen as a zero-order interpolation and generates what is called in the following an *instant-matching* error. Let $(x_t)_{t\geq t_0}$ the time occurrences of the stimuli on this finer grid. Accordingly, the HRF has to be estimated with the same temporal resolution:

$$y_{t_n} = \sum_{k=0}^{K} h_{k\Delta t} x_{t_n - k\Delta t} + b_{t_n}$$
$$= \boldsymbol{x}_{t_n}^{\text{t}} \boldsymbol{h} + b_{t_n}, \quad \text{for } t_n = t_1, \ldots, t_N, \quad (1)$$

with $\boldsymbol{x}_{t_n} = [x_{t_n}, x_{t_n - \Delta t}, \ldots, x_{t_n - K\Delta t}]^t \in \mathbb{R}^{K+1}$ and $\boldsymbol{h} = [h_0, h_{\Delta t}, \ldots, h_{K\Delta t}]^t$. Note first that the number of unknowns,

TABLE I
LIST OF NOTATIONS

$\boldsymbol{y}_i \in \mathbb{R}^{N_i}$	Data of session i for a given voxel
$\boldsymbol{b}_i \in \mathbb{R}^{N_i}$	Gaussian noise vector of session i
$K + 1$	Number of HRF coefficients for each condition
$\boldsymbol{X}_i^{(m)} \in \mathbb{R}^{N_i} \times \mathbb{R}^{K+1}$	Binary "onsets matrix" of stimulus m for session i
$\boldsymbol{h}_m \in \mathbb{R}^{K+1}$	HRF for the mth condition
M	Number of different conditions in the experiment
$\boldsymbol{h} = [\boldsymbol{h}_1^{\mathrm{t}}, \ldots, \boldsymbol{h}_M^{\mathrm{t}}]^{\mathrm{t}} \in \mathbb{R}^{M(K+1)}$	Set of HRFs
$\boldsymbol{X}_i = [\boldsymbol{X}_i^{(1)} \ldots | \boldsymbol{X}_i^{(M)}] \in \mathbb{R}^{N_i} \times \mathbb{R}^{M(K+1)}$	Binary "onsets matrix" of all conditions for session
Q_i	Number of drift parameters for session i
$\boldsymbol{P}_i = [P_{i,1}, \ldots, P_{i,Q_i}] \in \mathbb{R}^{N_i} \times \mathbb{R}^{Q_i}$	Low frequency orthogonal matrix for session i
$\boldsymbol{l}_i \in \mathbb{R}^{Q_i}$	nuisance parameters for session i
I	Sessions number
$N = \sum_{i=1}^{I} N_i$	Global number of data for all sessions
$\underline{\boldsymbol{y}} = [\boldsymbol{y}_1^{\mathrm{t}}, \ldots, \boldsymbol{y}_I^{\mathrm{t}}]^{\mathrm{t}} \in \mathbb{R}^N$	Complete set of data for all sessions
$\underline{\boldsymbol{X}} = [\boldsymbol{X}_1^{\mathrm{t}} | \ldots | \boldsymbol{X}_I^{\mathrm{t}}]^{\mathrm{t}}$	Binary "onsets matrix" for all sessions
$\underline{\boldsymbol{P}} = \mathrm{diag}\,[\boldsymbol{P}_1, \ldots, \boldsymbol{P}_I]$	Low frequency orthogonal matrix for all sessions
$\underline{\boldsymbol{l}} = [\boldsymbol{l}_1^{\mathrm{t}}, \ldots, \boldsymbol{l}_I^{\mathrm{t}}]^{\mathrm{t}} \in \mathbb{R}^{\sum_i Q_i}$	nuisance parameters for all sessions
$\boldsymbol{\theta_H} = [\tau_1, \ldots, \tau_M] \in \mathbb{R}^M$	Hyperparmters of the prior pdf $p(\boldsymbol{h}\,;\,\boldsymbol{R}, \boldsymbol{\theta_H})$
$\boldsymbol{\theta} = [r_b, \boldsymbol{\theta_H}^{\mathrm{t}}]^{\mathrm{t}}$	Complete set of hyperparameters
$\boldsymbol{R}$	Prior covariance matrix of $\boldsymbol{h}_m$
$\boldsymbol{R_H} = \mathrm{diag}\,[\tau_1 \boldsymbol{R}, \tau_2 \boldsymbol{R}, \ldots, \tau_M \boldsymbol{R}]$	Prior covariance matrix of $\boldsymbol{h}$
$\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$	MAP HRFs estimate
$\boldsymbol{\Sigma}$	Posterior covariance matrix of $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$
$\boldsymbol{\Upsilon}$	Covariance matrix of the data $\underline{\boldsymbol{y}}$
$\mathcal{E}(\widehat{\boldsymbol{h}}_m\,;\,\boldsymbol{h}_m)$	Quadratic error between $\boldsymbol{h}_m$ and its estimate $\widehat{\boldsymbol{h}}_m$
$\mathcal{S}(\widehat{\boldsymbol{h}}_m)$	Mean standard deviation error measure for $\widehat{\boldsymbol{h}}_m$

i.e. $K + 1$, may be dramatically larger than its counterpart in the synchronous case and second, that oversampling of the data is not required. b_{t_n} is the nth sample of a zero-mean Gaussian white noise process $\boldsymbol{b}$ of unknown variance $r_b > 0$, independent of $\boldsymbol{h}$. Such a hypothesis may seem restrictive since it is well-known that fMRI time series are correlated in time [19]. Nonetheless, as shown in [18] various noise correlation structures have little influence on the performances of the estimation. The same result has also been emphasized in [20], where the authors analyze the influence of the colored nature of fMRI noise on the *average bias* of the HRF estimate. Alternatively, one could estimate the temporal covariance structure of the noise with an autoregressive model, as done in [21].

In real neuroimaging experiments, the fMRI raw data are contaminated by a low-frequency drift mainly due to physiological artifacts [22]: breathing and cardiac pulses are aliased since the sampling frequency of the data is below Nyquist's bound. Thus, these physiological factors introduce some low frequency fluctuations. A high-pass filter is generally used to remove those trends before estimating the HRF. In this study, we simultaneously estimate the HRF and the trend with the following model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{h} + \boldsymbol{P}\boldsymbol{l} + \boldsymbol{b} \tag{2}$$

where $\boldsymbol{X} = [\boldsymbol{x}_{t_1}, \ldots, \boldsymbol{x}_{t_N}]^{\mathrm{t}}$ defines the binary onsets matrix. Matrix $\boldsymbol{P} = [P_1, \ldots, P_Q]$ consists of an orthonormal basis of functions $P_q = [P_q(t_1), \ldots, P_q(t_N)]^{\mathrm{t}}$ modeling the low frequencies (e.g. a one dimensional discrete cosine transform). The number Q of basis functions depends on the lowest frequency $f_{\min}$ attributable to the drift term and can be defined as $Q = [2N\,\mathrm{TR}f_{\min}] + 1$, where "+1" stands for the mean (constant term) and $[\cdot]$ is the integer part operator. Matrix $\boldsymbol{P}$ can also take any covariate of no interest into account, supposed to influence the signal intensity in a linear way. Vector $\boldsymbol{l} \in \mathbb{R}^Q$ defines the unknown weighting coefficients of the basis functions, called *nuisance variables* in the following.

C. Asynchronous Multitask Paradigms

We further extend (2) to allow for a different HRF estimate for different trial types (e.g.,different stimuli or conditions). Let $(\boldsymbol{X}^{(m)})_{1 \leqslant m \leqslant M}$ be the different trial-dependent matrices, each of them being defined as the previous $\boldsymbol{X}$ matrix, and then suppose that the HRFs $\boldsymbol{h}_m$ add in a *linear* way. Such an extension requires to correctly define the oversampling period Δt as the smallest sampling interval that allows to separate the two closest

events, whatever their type. For the sake of simplicity, let us define

$$\boldsymbol{X} = \left[\boldsymbol{X}^{(1)} | \dots | \boldsymbol{X}^{(M)} \right] \in \mathbb{R}^N \times \mathbb{R}^{M(K+1)}$$

$$\boldsymbol{h} = \left[\boldsymbol{h}_1^{\mathrm{t}}, \dots, \boldsymbol{h}_M^{\mathrm{t}} \right]^{\mathrm{t}} \in \mathbb{R}^{M(K+1)}$$

from which model (2) is able to cope with *asynchronous multi-task* paradigms.

D. Multisession Likelihood

As previously mentioned, the experimental paradigm is repeated several times for a given subject, leading to a few sessions of about 200 to 1000 data each.

It is generally assumed that the HRF remains approximately constant provided that all the exogenous parameters (voxel, task, subject) are fixed. Accordingly, the same vector $\boldsymbol{h}$ is sought from the available fMRI times series, say $\boldsymbol{y}_1, \dots, \boldsymbol{y}_I$, of respective length N_i, so that a specific matrix $\boldsymbol{X}_i$ is involved for each session. In addition, it seems relevant to select a session-dependent value of $\boldsymbol{l}_i$ and possibly Q_i for the definition of $\boldsymbol{P}_i \in \mathbb{R}^{N_i} \times \mathbb{R}^{Q_i}$ since the physiological factors (breathing and cardiac rates) fluctuate throughout the sessions. Hence, the multisession extension of (2) is given by

$$(\mathcal{M}): \quad \boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{h} + \boldsymbol{P}_i \boldsymbol{l}_i + \boldsymbol{b}_i, \quad \text{for } i \in \mathbb{N}_I^*$$

Model $(\mathcal{M})$ calls for two comments.

First, it relies on the following assumptions about noises $\boldsymbol{b}_i$:

- the mean of $\boldsymbol{b}_i$ (i.e.,the baseline of $\boldsymbol{h}$) may vary across sessions. This variation is captured by the constant column of matrix $\boldsymbol{P}_i$.
- the variance of $\boldsymbol{b}_i$ is supposed constant across the sessions for the sake of simplicity. In Section V, it will be shown that the proposed HRF estimate remains robust with regard to departures from this hypothesis. There is no theoretical limitation preventing us from introducing a specific variance for each session. Nonetheless, we advocate the use of the same unknown variance for all sessions on the same subject since we observed on our data that the major fluctuations occuring in real data are rather due to physiological variations (modeled by session-dependent trends) than to some modification of the noise scaling.

Second, model $(\mathcal{M})$ allows to introduce more information than model (2), through the introduction of new independent data.

To estimate and make inferences about the hemodynamic response, we first need to take model $(\mathcal{M})$ into account through the definition of the likelihood function. For the sake of conciseness, let us introduce the following notations:

$$\underline{y} = \left[\boldsymbol{y}_1^{\mathrm{t}}, \dots, \boldsymbol{y}_I^{\mathrm{t}} \right]^{\mathrm{t}}, \qquad \underline{X} = \left[\boldsymbol{X}_1^{\mathrm{t}} | \dots | \boldsymbol{X}_I^{\mathrm{t}} \right]^{\mathrm{t}}$$

$$\underline{P} = \mathrm{diag}[\boldsymbol{P}_1, \dots, \boldsymbol{P}_I], \qquad \underline{l} = \left[\boldsymbol{l}_1^{\mathrm{t}} | \dots | \boldsymbol{l}_I^{\mathrm{t}} \right]^{\mathrm{t}}$$

where the diag operator is used to define $\underline{P}$ as a block-diagonal matrix. From these assumptions, the likelihood of the asynchronous multitrials multisession model $(\mathcal{M})$ is given by

$$p\left(\underline{y}|\boldsymbol{h}, \mathcal{M}; r_b, BBl\right) = \prod_{i=1}^I p\left(\boldsymbol{y}_i|\boldsymbol{h}, \mathcal{M}; r_b, \boldsymbol{l}_i\right)$$

$$= (2\pi r_b)^{-\frac{N}{2}}$$

$$\exp\left(\frac{-\sum_{i=1}^I \|\boldsymbol{y}_i - \boldsymbol{X}_i \boldsymbol{h} - \boldsymbol{P}_i \boldsymbol{l}_i\|^2}{2r_b}\right)$$

$$= (2\pi r_b)^{-\frac{N}{2}} \exp\left(\frac{-\|\underline{y} - \underline{X}\boldsymbol{h} - \underline{Pl}\|^2}{2r_b}\right) \quad (3)$$

where $N = \sum_{i=1}^I N_i$.

The number of parameters still remains large so that least squares estimation is unreliable when $[\underline{X}|\underline{P}]^{\mathrm{t}}[\underline{X}|\underline{P}]$ is ill-conditioned (the variance of the fitted parameters is too large). A straight application of the inversion lemma for block matrices provides a necessary and sufficient condition: this block-matrix is invertible and well-conditioned if and only if the inverse of $\underline{X}^t(\boldsymbol{I}_M - \underline{P}(\underline{P}^t\underline{P})^{-1}\underline{P}^t)\underline{X} = \underline{X}^t(\boldsymbol{I}_M - \underline{P}\underline{P}^t)\underline{X}$ exists and is well-conditioned.[1] Note that such a matrix can be ill-conditioned even if $\underline{X}^t\underline{X}$ is well-conditioned. The limiting case, corresponding to the underdetermination of model $(\mathcal{M})$, can be reached if Δt is too low or M too large.

III. IDENTIFICATION OF THE HRFs

Since the information provided by the data may not be sufficient to derive a robust HRF estimate, we introduce some constraints on the temporal structure of the HRFs that correspond to some available physiological prior knowledge. The proposed HRF estimate will result from an appropriate tradeoff of both types of information (data-driven and prior).

A. Prior Information

Following [15], [18], we introduce temporal prior information within the Bayesian framework. As physiologically advocated in [5], each HRF is characterized by the following features:

1) Its amplitude is close to zero at the first and end points. As a matter of fact, the HRF is causal, hence, $h_{m,k\Delta t}$ should be zero for $k \le 0$ and $m \in \mathbb{N}_M^*$. This means that the stimulus at time k should only have influence for $k > 0$. Similarly, the influence of an activation should vanish in the past, implying that the filter parameters should tend to zero for large delays $(h_{m,K\Delta t} = 0)$.
2) Its variations are smooth.
3) Prior statistical independence is supposed to hold between stimuli and, thus, between HRFs. In addition, each HRF may be related to an underlying physiological process having its own dynamics (specific time-to-peak and dispersion of response).

---

[1] Let $\boldsymbol{A} \in \mathbb{R}^K \times \mathbb{R}^K$ be an invertible matrix, then $(\boldsymbol{A}) = (\boldsymbol{A}^{-1})$ if Cond stands for the condition number [23].

Condition 1) is easily introduced by redefining vectors $\boldsymbol{h}_m$ and $\boldsymbol{x}_{i,tn}^{(m)}$ for the ith session and all $n \in \mathbb{N}_N^*$, setting the first and last parameters of $\boldsymbol{h}_m$ to zero

$$\boldsymbol{h}_m = \left[h_{m,\Delta t}, \ldots, h_{m,(K-1)\Delta t} \right]^{\mathrm{t}} \in \mathbb{R}^{K-1}$$

$$\boldsymbol{x}_{i,t_n}^{(m)} = \left[x_{i,t_n-\Delta t}^{(m)}, \ldots, x_{i,t_n-(K-1)\Delta t}^{(m)} \right]^{\mathrm{t}} \in \mathbb{R}^{K-1}$$

$$\boldsymbol{X}_i^{(m)} = \left[\boldsymbol{x}_{i,t_1}^{(m)}, \ldots, \boldsymbol{x}_{i,t_1}^{(m)} \right]^{\mathrm{t}} \in \mathbb{R}^{N_i} \times \mathbb{R}^{K-1}.$$

The likelihood function (3) remains unchanged with $\boldsymbol{h} = [\boldsymbol{h}_1^{\mathrm{t}}, \ldots, \boldsymbol{h}_M^{\mathrm{t}}]^{\mathrm{t}} \in \mathbb{R}^{M(K-1)}$ and $\boldsymbol{X}_i = [\boldsymbol{X}_i^{(1)} | \ldots | \boldsymbol{X}_i^{(M)}] \in \mathbb{R}^{N_i \times M(K-1)}$ for the ith session.

Quantification of condition 2) is achieved by setting a Gaussian probability density function (pdf) $\mathcal{N}(0, \tau_m \boldsymbol{R})$ for $p(\boldsymbol{h}_m; \boldsymbol{R}, \tau_m)$. We have chosen $\boldsymbol{R} = (\boldsymbol{D}_2^{\mathrm{t}} \boldsymbol{D}_2)^{-1}$ for the prior covariance where $\boldsymbol{D}_2$ is the truncated second-order finite difference matrix in order to fullfil constraint 1)

$$\boldsymbol{D}_2 = \begin{bmatrix} -2 & 1 & 0 & \cdots & \cdots & 0 \\ 1 & -2 & 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 0 & 1 & -2 & 1 \\ 0 & \cdots & \cdots & 0 & 1 & -2 \end{bmatrix}.$$

Since $\boldsymbol{R}^{-1}$ is of full rank, $p(\boldsymbol{h}_m; \boldsymbol{R}, \tau_m)$ defines a proper prior.[2] For model $(\mathcal{M})$, condition 3) may be taken into account by the following pdf:

$$p(\boldsymbol{h}; \boldsymbol{R}, \boldsymbol{\theta}_H) = \prod_{m=1}^{M} p(\boldsymbol{h}_m; \boldsymbol{R}, \tau_m)$$

$$\propto \frac{\prod_{m=1}^{M} \tau_m^{-\frac{(K-1)}{2}}}{\det(\boldsymbol{R})^{+\frac{M}{2}}} \exp\left(-\frac{\boldsymbol{h}^{\mathrm{t}} \boldsymbol{R}_H^{-1} \boldsymbol{h}}{2}\right) \quad (4)$$

where $\boldsymbol{\theta}_H = [\tau_1, \ldots, \tau_M] \in \mathbb{R}^M$ stands for the hyperparameters of the prior model and

$$\boldsymbol{R}_H = \mathrm{diag}[\tau_1 \boldsymbol{R}, \tau_2 \boldsymbol{R}, \ldots, \tau_M \boldsymbol{R}] \in \mathbb{R}^{M(K-1)} \times \mathbb{R}^{M(K-1)}.$$

This prior model clearly favors *smooth* responses, since for each condition it amounts to minimizing $\|\boldsymbol{D}_2 \boldsymbol{h}_m\|^2$, i.e., the discrete approximation of the second-order derivative of $\boldsymbol{h}_m$. We note also that the sampling period Δt is integrated in the definition of the prior variances τ_m and that introduction of different parameters τ_m for different HRFs allows to model specific dynamics for each condition.

B. Modeling Choices and Error Bar Computation

Our HRF estimation technique is a two-steps operation, which first consists in jointly estimating $\underline{l}$ and $\boldsymbol{\theta}$ in the ML sense (see details in Section IV) before substituting the estimates for the true values in $p(\boldsymbol{h}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}, \underline{l})$ in order to compute its maximizer. In the first step, we assume that the drift parameters $\underline{l}$ as well as hyperparameters $\boldsymbol{\theta} = [r_b, \boldsymbol{\theta}_H^{\mathrm{t}}]^{\mathrm{t}}$ are *deterministic*. We motivate this choice below.

It does not seem to us that considering the drift terms as random variables would allow the addition of significant information because we have rather poor prior knowledge at hand.

[2]The pdf $p(\boldsymbol{h}; \boldsymbol{R}, \tau_m)$ is said to be proper if $\int_{\boldsymbol{h}} p(\boldsymbol{h}; \boldsymbol{R}, \tau_m) d\boldsymbol{h} < \infty$.

Therefore, a reasonable prior on those would have been uninformative or even improper. Moreover, in Section V-E we argue that with a small number of nuisance variables (i.e.,when Q/N is small, typically $Q/N \leqslant 0.02$) and a standard noise level, we are able to accurately estimate these parameters and be reasonably confident in the variability of our HRF estimate. In other words, when we substitute $\widehat{\underline{l}}^{\mathrm{ML}}$ for $\underline{l}$, the extra source of error on $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$ still is small. The same holds for hyperparameters $\boldsymbol{\theta}$. It follows that the most important quantity $\boldsymbol{h}$ is estimated from the maximum *a posteriori* (MAP) the maximum of the posterior distribution $p(\boldsymbol{h}|\underline{y}, \mathcal{M}, \boldsymbol{\theta}, \underline{l})$. The Bayes rule ensures the fusion of the likelihood (3) and the prior (4) into the Gaussian posterior distribution of $\boldsymbol{h}$ given $(\underline{y}, \boldsymbol{\theta}, \underline{l})$, from which we can derive the MAP estimate

$$p\left(\boldsymbol{h}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}, \underline{l}\right) \sim \mathcal{N}(\widehat{\boldsymbol{h}}^{\mathrm{MAP}}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{r_b} \sum_{i=1}^{I} \boldsymbol{X}_i^{\mathrm{t}} \boldsymbol{X}_i + \boldsymbol{R}_H^{-1}$$

$$\widehat{\boldsymbol{h}}^{\mathrm{MAP}} = \frac{1}{r_b} \boldsymbol{\Sigma} \sum_{i=1}^{I} \boldsymbol{X}_i^{\mathrm{t}} (\boldsymbol{y}_i - \boldsymbol{P}_i \boldsymbol{l}_i). \quad (5)$$

The influence of this choice of model should have limited impact on the estimation of $\boldsymbol{h}$. However, the influence on the error bars of $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$ might not be negligible. As we would like to assess the error made on $\widehat{\boldsymbol{h}}_m^{\mathrm{MAP}}$, we focus on the marginal posterior pdf $p(\boldsymbol{h}_m|\underline{y}, \mathcal{M}; \boldsymbol{\theta}, \underline{l})$. As expected, this pdf is $\mathcal{N}\widehat{\boldsymbol{h}}_m^{\mathrm{MAP}}, \boldsymbol{\Sigma}_m$—distributed with $\boldsymbol{\Sigma}_m$, the mth diagonal block of $\boldsymbol{\Sigma}$. The marginal error bars can then be derived from the standard deviations $\boldsymbol{\sigma}_m = [\sigma_{m,1}, \ldots, \sigma_{m,K-1}]$, that are given by the square roots of the main diagonal of $\boldsymbol{\Sigma}_m$. Although slightly underestimated, such error bars provide a good approximation of the range of variation of $\boldsymbol{h}_m$. Note that these error bars are created for pointwise inference and appropriate only for a single preselected time sample since the search over all time samples requires to analyze the posterior covariance matrix $\boldsymbol{\Sigma}_m$, which cannot be easily represented on the same graph as $\widehat{\boldsymbol{h}}_m^{\mathrm{MAP}}$.

C. Equivalence With the Marginal MAP Estimate

Under technical conditions, we show that our HRF estimate $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$ is a valid approximation of another Bayesian estimate that takes all fluctuations of $\underline{l}$ into account. This solution corresponds to the maximizer $\widehat{\boldsymbol{h}}^{\mathrm{MMAP}}$ of the *marginal* pdf $p(\boldsymbol{h}|\underline{y}, \mathcal{M}; \boldsymbol{\theta})$

$$p(\mathrm{h}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) = \int p(\boldsymbol{h}, \underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) d\underline{l}$$

$$= \int p(\boldsymbol{h}|\underline{y}, \underline{l}, \mathcal{M}; \boldsymbol{\theta}) p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) d\underline{l} \quad (6)$$

where $\underline{l}$ is assumed to be random. Using Bayes rule, the marginal posterior pdf $p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta})$ rereads

$$p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) \propto p(\underline{y}|\underline{l}, \mathcal{M}; r_b); p(\underline{l}; \boldsymbol{\theta}_L). \quad (7)$$

In the absence of relevant prior information about $\underline{l}$, a flat distribution is considered for $p(\underline{l}; \boldsymbol{\theta}_L)$ leading in (7) to $p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) \propto p(\underline{y}|\underline{l}, \mathcal{M}; \boldsymbol{\theta})$. Assuming that $p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta})$ is peaked enough, we may write $p(\underline{l}|\underline{y}, \mathcal{M}; \boldsymbol{\theta}) \approx \delta(\underline{l} - \widehat{\underline{l}}^{\mathrm{MAP}})$, where $\widehat{\underline{l}}^{\mathrm{MAP}} = \widehat{\underline{l}}^{\mathrm{ML}}$ is

the ML solution computed by our approach (see details in Section IV). If this hypothesis is fullfiled, the marginal distribution (6) is, thus, approximated by

$$p(\boldsymbol{h}|\underline{y},\mathcal{M};\boldsymbol{\theta}) \approx \int p(\boldsymbol{h}|\underline{y},\underline{l},\mathcal{M};\boldsymbol{\theta})\delta(\underline{l}-\widehat{\underline{l}}^{\mathrm{ML}})d\underline{l}$$

$$\approx p(\boldsymbol{h}|\underline{y},\underline{l}=\widehat{\underline{l}}^{\mathrm{ML}},\mathcal{M};\boldsymbol{\theta}). \qquad (8)$$

Therefore, our approach amounts to computing $\widehat{\boldsymbol{h}}^{\mathrm{MMAP}}$ provided that $p(\underline{l}|\underline{y},\mathcal{M};\boldsymbol{\theta}) \approx \delta(\underline{l}-\widehat{\underline{l}}^{\mathrm{ML}})$. In Section V-E, we discuss the validity of this assumption.

IV. HYPERPARAMETERS AND NUISANCE VARIABLE ESTIMATION

This section focuses on the automatic tuning of hyperparameters $\boldsymbol{\theta}$ and nuisance variables $\underline{l}$. This part starts with an introduction to the ML principle and then examines a well-adapted way to tackle the underlying optimization problem.

A. ML Principle

ML estimation for hyperparameters is a very common procedure, which is currently used in various fields of signal and image processing when dealing with a small number of unknown but deterministic hyperparameters $\boldsymbol{\theta}$ (see, for instance, [24] and [25]). The underlying reason that makes this approach feasible and attractive is that there is a large number of data (N) to estimate $(\boldsymbol{\theta}, \underline{l})$ accurately. Similarly to [15], we select the hyperparameters and nuisance variables by maximization of their likelihood $p(\underline{l}|\mathcal{M};\boldsymbol{\theta},\underline{l})$, obtained from the joint pdf of $(\boldsymbol{H}, \underline{Y})$ after integration over the parameters $\boldsymbol{H}$

$$p(\underline{y}|\mathcal{M};\boldsymbol{\theta},\underline{l}) = \int_{\boldsymbol{h}} p(\underline{y}|\boldsymbol{h},\mathcal{M};r_b,\underline{l})p(\boldsymbol{h};\boldsymbol{R},\boldsymbol{\theta_H})d\boldsymbol{h}$$

$$= \frac{(2\pi)^{-\frac{N}{2}}}{\det(\boldsymbol{\Upsilon})^{\frac{1}{2}}}\exp\left(\frac{(\underline{y}-\underline{Pl})^{\mathrm{t}}\boldsymbol{\Upsilon}^{-1}(\underline{y}-\underline{Pl})}{2}\right) \quad (9)$$

with[3]

$$\boldsymbol{\Upsilon} = r_b\boldsymbol{I}_N + \underline{X}\boldsymbol{R}_{\boldsymbol{H}}^{-1}\underline{X}^{\mathrm{t}}.$$

The ML estimator $(\widehat{\boldsymbol{\theta}}^{\mathrm{ML}}, \widehat{\underline{l}}^{\mathrm{ML}})$ of $(\boldsymbol{\theta}, \underline{l})$ minimizes

$$\log\det(\boldsymbol{\Upsilon}) + (\underline{y} - \underline{Pl})^{\mathrm{t}}\boldsymbol{\Upsilon}^{-1}(\underline{y} - \underline{Pl}). \qquad (10)$$

Computation of $(\widehat{\boldsymbol{\theta}}^{\mathrm{ML}}, \widehat{\underline{l}}^{\mathrm{ML}})$ is a complicated nonlinear optimization problem of several variables. First, we must resort to an iterative optimization scheme because of the nonquadratic behavior of $(\log\det\boldsymbol{\Upsilon})$. Second, since $\log p(\underline{y}|\mathcal{M};\boldsymbol{\theta},\underline{l})$ is not concave with respect to $(\boldsymbol{\theta}, \underline{l})$, there may exist several local minima. Therefore, the hyperparameters and nuisance variables given by any deterministic algorithm depend on the initial values. In the next subsection, we propose to resort to a variant of the expectation-maximization (EM) algorithm to avoid direct optimization of (10).

B. EM-Based Strategy

For notation compactness, let us denote $\widetilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}^{\mathrm{t}}, \underline{l}^{\mathrm{t}}]^{\mathrm{t}}$. The EM algorithm, introduced by [26], is a general iterative method which ensures the increasing of the likelihood function $p(\underline{y}|\mathcal{M};\widetilde{\boldsymbol{\theta}})$ of a parameter vector $\widetilde{\boldsymbol{\theta}}$ given observations $\underline{y}$ at each iteration. Starting from an initial value $\widetilde{\boldsymbol{\theta}}^0$, a series of succesive estimates $\widetilde{\boldsymbol{\theta}}^k$ is generated by alternating the following two steps:

$$\text{Expectation(E)}: \text{Evaluate}\mathcal{Q}(\widetilde{\boldsymbol{\theta}},\widetilde{\boldsymbol{\theta}}^k;\underline{y}) \qquad (11a)$$

$$\text{Maximization}(\mathcal{M}): \widetilde{\boldsymbol{\theta}}^{k+1} = \arg\max_{\widetilde{\boldsymbol{\theta}}} \mathcal{Q}(\widetilde{\boldsymbol{\theta}},\widetilde{\boldsymbol{\theta}}^k;\underline{y}) \quad (11b)$$

where function $\mathcal{Q}$ is defined as

$$\mathcal{Q}(\widetilde{\boldsymbol{\theta}},\widetilde{\boldsymbol{\theta}}^k;\underline{y}) = \int_{\boldsymbol{h}} \log p(\underline{y},\boldsymbol{h}|\mathcal{M};\widetilde{\boldsymbol{\theta}})p(\boldsymbol{h}|\underline{y},\mathcal{M};\widetilde{\boldsymbol{\theta}}^k)d\boldsymbol{h}$$

$$= \mathrm{E}_{\boldsymbol{H}}\left[\log p(\underline{y},\boldsymbol{h}|\mathcal{M};\widetilde{\boldsymbol{\theta}})|\underline{y},\mathcal{M};\widetilde{\boldsymbol{\theta}}^k\right] \qquad (12)$$

$\boldsymbol{H}$ being an auxiliary random variable whose practical role is to make the *complete* likelihood $p(\underline{y},\boldsymbol{h}|\mathcal{M};\widetilde{\boldsymbol{\theta}})$ easier to compute than the original one $p(\underline{y}|\mathcal{M};\widetilde{\boldsymbol{\theta}})$. Following [27], parameter vector $\widetilde{\boldsymbol{\theta}}$ can be partitioned into two subvectors $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}} = [r_b, l^{\mathrm{t}}]^t$ and $\boldsymbol{\theta}_{\boldsymbol{H}} = [\tau_1,\ldots,\tau_M]^t$ which respectively control the conditional pdfs $p(\underline{y}|\boldsymbol{h},\mathcal{M};\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}})$ and $p(\boldsymbol{h};\boldsymbol{R},\boldsymbol{\theta}_{\boldsymbol{H}})$. Then, as shown in Appendix I, the M-step (11b) can be divided into two operations: maximization of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$ and maximization of $\mathcal{Q}_{\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}_{\boldsymbol{H}}$ (see (19) and (20) for the definitions of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ and $\mathcal{Q}_{\boldsymbol{H}}$, respectively).

In the present case, detailed examination of the maximization of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$ shows a tricky problem: the *joint* maximization with respect to r_b and $\underline{l}$, required by the definition (11b), cannot be performed since the updating step of r_b strongly depends on $\underline{l}$. Nonetheless, the M-step (11b) can be replaced by a *conditional* optimization strategy that reads

$$\forall \beta_l \in \beta = \boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}, \beta_l^{k+1}$$

$$= \arg\max_{\beta_l} \mathcal{Q}_{\underline{Y}|\boldsymbol{H}}\left(\beta_{l-1}^{k+1},\widetilde{\boldsymbol{\theta}}^k;\underline{y}\right)$$

$$\text{with} \quad \beta_{l-1}^{k+1} = \{\beta_1^{k+1},\ldots,\beta_{l-1}^{k+1},\beta_l,\beta_{l+1}^k,\ldots,\beta_{Q+1}^k\}. \quad (13)$$

The variant (13) is known as the *expectation conditional maximization*, (ECM) [28]. It is actually a subclass of generalized EM algorithms that are more broadly applicable than EM but share its desirable convergence properties.[4] Therefore, we resort to ECM for function $\mathcal{Q}_{\underline{y}|\boldsymbol{H}}$. Hereafter, we provide the updating equations of ECM for the estimation of $\boldsymbol{\theta}_{\underline{y}|\boldsymbol{H}}$ and $\boldsymbol{\theta}_{\boldsymbol{H}}$.

C. Re-Estimation Formulas for $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$

The present M step of ECM is performed by searching an optimal parameter of $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$ at a time, say β_l^{k+1}, given by (13) or equivalently by the solution of $\partial\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}/\partial\beta_l^{k+1} = 0$. This procedure is repeated for all parameters belonging to $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$. As

---

[3]In these definitions, every matrix $\boldsymbol{P}_i$ is supposed to have $Q_i = Q$ columns $\forall i \in \mathbb{N}_I^*$. If such a choice turns out to be inappropriate, the definitions of matrix $\underline{P}$ and vector $\underline{l}$ should be revisited, by zero-padding matrices $\boldsymbol{P}_i$ and vectors $\boldsymbol{l}_i$ that do not have the maximal number of components, i.e., $\max_i Q_i$.

[4]The EM algorithm converges monotonically to a stationary point of $p(\underline{y}|\mathcal{M};\widetilde{\boldsymbol{\theta}})$ if $\mathcal{Q}(\widetilde{\boldsymbol{\theta}},\widetilde{\boldsymbol{\theta}}^k;\underline{y},\mathcal{M})$ is continuous in both $\widetilde{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}^k$ [26], [29].

shown in Appendix II, the computation of the optimal nuisance variables $\widehat{l}$ can be parallelized since

$$\widehat{l}_i = P_i^t(y_i - X_i\widehat{h}^{\text{MAP}}), \quad \text{for} \quad i \in \mathbb{N}_I^*. \tag{14}$$

Taking the new estimates $\widehat{l}_i$ and all the available datasets into account,[5] we get for the noise variance

$$\widehat{r}_b = \frac{\sum_{i=1}^{l} \mathcal{E}_i(\widehat{l}_i)}{N} \tag{15}$$

where functions $\mathcal{E}_i(\cdot)$ are defined by (24) (see Appendix II). As it appears in (24), updating both l_i and r_b requires the computation of $\widehat{h}^{\text{MAP}}$. Strictly speaking, this estimate is given by the solution of a linear system without explicitly computing matrix Σ. Nonetheless, the latter matrix is needed for setting $\widehat{r}_b$ and is, thus, computed before updating l_i.

D. Re-Estimation Formulas for θ_H

The M step of the ECM algorithm with respect to θ_H is actually identical to a M step of the standard version of EM since all the parameters belonging to θ_H can be simultaneously updated. As previously derived, the optimal parameters θ'_H are given by the solution of $\partial Q_H / \partial \theta'_H = 0$ that is (see Appendix III for details)

$$\tau'_m = \begin{cases} \dfrac{\text{tr}\left\{\left(\widehat{h}^{\text{MAP}}(\widehat{h}^{\text{MAP}})^t + \Sigma\right) R_H^{-1}\right\}}{M(K-1)}, & \text{if } \tau'_m = \tau', \forall m \\[3ex] \dfrac{\text{tr}\left\{\left(\widehat{h}_m^{\text{MAP}}\left(\widehat{h}_m^{\text{MAP}}\right)^t + \Sigma_m\right) R^{-1}\right\}}{K-1}, & \text{otherwise.} \end{cases} \tag{16}$$

Finally, the successive steps of ECM are summarized in Table II.

As to numerical implementation of ECM, the following conjunction has been selected as stopping criterion:

$$\frac{\left| Q(\widetilde{\theta}^k, \widetilde{\theta}^{k-1}; \underline{y}, \mathcal{M}) - Q(\widetilde{\theta}^{k-1}, \widetilde{\theta}^{k-1}; \underline{y}, \mathcal{M}) \right|}{\left| Q(\widetilde{\theta}^k, \widetilde{\theta}^{k-1}; \underline{y}, \mathcal{M}) \right|} < \gamma_1$$

$$\sup_i \left[\frac{\left\| \widetilde{\theta}_i^k - \widetilde{\theta}_i^{k-1} \right\|_2}{\left\| \widetilde{\theta}_i^k \right\|_2} \right] < \gamma_2$$

where $\widetilde{\theta}^k$ denotes the solution at the kth iteration of the minimization stage and the thresholds have been set to $(\gamma_1, \gamma_2) = (10^{-4}, 10^{-5})$.

Sections II–IV have allowed us to design *unsupervised nonparametric estimates* of the HRF at a given voxel of the brain, depending on the trial type and the subject under study. Moreover, we have provided uncertainty measures on this estimation to be able to quantify the confidence we may have in the results. Hereafter, we demonstrate the accuracy and robustness of the proposed estimates, first on simulated datasets, and finally through an experimental fMRI paradigm.

TABLE II
ECM ALGORITHM FOR HYPERPARAMETERS AND NUISANCE
VARIABLE ESTIMATION

1) Initialize $\widetilde{\theta}^0$.
2) Iteration k ($k \geqslant 1$)
   - Estimate $\theta_{\underline{Y}|H}^k = [l_1^k, \ldots, l_I^k, r_b^k]^t$:
     – Compute l_i^k using (14) for $i \in \mathbb{N}_I^*$;
     – Compute r_b^k using (15);
   - Estimate $\theta_H^k = [\tau_1^k, \ldots, \tau_M^k]^t$ using (16);
3) Compute $Q(\widetilde{\theta}^k, \widetilde{\theta}^{k-1}; \underline{y}, \mathcal{M})$ using (21)–(25);
4) Iterate $k = k + 1$ and $\widetilde{\theta}^k = [\theta_{\underline{Y}|H}^k, \theta_H^k]^t$ until stop rules are satisfied;

V. SIMULATIONS RESULTS

The first part of this section is a Monte Carlo study that compares the statistical properties of the ML and MAP HRF estimates (summarized in Table III).

In the following simulations, the CNR is defined as the ratio between the l_1-norm of the HRF and the standard deviation $r_b^{1/2}$ of the noise[6]:

$$\text{CNR} \triangleq \frac{\sum_{k=0}^{K} |h_{m,k\Delta t}|}{(K+1) r_b^{1/2}}$$

To provide a single CNR value for time series containing several HRFs with different shapes, we simply average the CNR of each response. The CNR is taken in the range of observed CNR in fMRI data.

The rest of this section emphasizes the performances of the MAP estimator and focuses on the following topics, specific to our extensions:

- Effect of *oversampling*, whether it is relevant to choose a HRF sampling period Δt for h lower than TR, the sampling period of the data (see Section V-C). For this topic, we have chosen a high CNR to rigorously quantify the instant-matching error.
- Successive improvements when allowing for trends and when processing several sessions at the estimation stage (see Section V-D). For this investigation, we have considered a higher noise level to better highlight the improvements brought by the multisession method on the average bias and variance of estimation of the HRF. In addition, a session-dependent drift term, similar to the one observed in fMRI data, has been added to quantify the gain that we can achieve when modeling these fluctuations.
- Validity of the error bars on the HRF estimates (see Section V-E).
- Robustness of the HRF estimate to departure from the hypothesis of equal noise variance across sessions (see Section V-F).
- Influence of *overparametrization*, i.e., when too many HRFs have been modeled and estimated compared with

---

[5]Since a constant noise variance has been assumed throughout the sessions.

[6]The CNR is sometimes defined as the ratio between the magnitude of the peak signal change and the standard deviation of the noise. However, the peak value is no longer a good descriptor of the entire signal when different HRF shapes are considered.

TABLE III
SIMULATION PARAMETERS FOR FOUR DIFFERENT DATASETS

Subsect.	Drift		r_b	CNR	# Sess.
	Flag	Cut-Off-Period (s)			
§ V-B	No		[0.7 0.028]	[0.3 1.53]	1
§ V-C	No		0.008	1.46	1
§ V-D	Yes	[180 170 160 180]	0.03	0.73	4
§ V-E	Yes	[180 170 160 180]	[0.03 0.2]	[0.73 0.29]	4
		[40 50 50 40]			
§ V-F	No		[0.08 0.3]	[0.46 0.24]	2
§ V-G	Yes	160	0.02	0.92	1

the number that are actually present in the voxel under study (see Section V-G). For this purpose, favorable experimental conditions have been chosen.

Each of these topics is analyzed using synthetic data. The parameters of the simulation have been chosen to be compatible with experimental conditions usually encountered, as explained now.

A. Common Procedure for Generating Simulated Datasets

For each session, we have first simulated a *random-intermixed* sequence of indexes coding for two different event types ($M = 2$). Each index corresponds to a specific stimulus. Unless otherwise specified, the timing of the trials is random, since the ISIs between successive trials follow a uniform distribution on [2.5, 3.5]. This might not be optimal for ML estimate. However, an optimal design for the estimation of the HRF will not be optimal for signal detection [20] and we have, therefore, chosen a tradeoff that is generally considered in neuro-imaging experiments. To investigate the influence of the experimental design on the behavior of the ML and MAP HRF estimates, we compare the statistical properties of the solutions computed for 1) an event-related paradigm (low detection efficiency and good HRF estimation [20], [30]) and 2) a block design (high detection efficiency and poor HRF shape estimation [20], [31]) at the same noise level.

While the optimization of the design parameters is out of the scope of this work, these two settings contrast two opposite situations from the HRF estimation/signal detection point of view. The reader interested by the optimization of experimental design may refer to an excellent survey [32] (and references quoted therein), which also introduces a stochastic framework based on genetic algorithms[7] to optimize a fitness measure of the experimental design wrt several parameters (detection efficiency, HRF shape estimation, counterbalancing of events, ...).

The onsets of the trials are put together on the same temporal grid using $\delta t = 0.5$ s for sampling period. This step simply requires to move the onsets to the nearer time points on the $n\delta t$ grid.

Each binary time series coupled to a stimulus is then convolved with a specific HRF, $\boldsymbol{h}_1$ or $\boldsymbol{h}_2$, whose exact shapes are

plotted in Fig. 1 for instance. $\boldsymbol{h}_1$ is the canonical HRF used by the SPM99 software,[8] whereas $\boldsymbol{h}_2$ is chosen as an extreme example of a very peaky HRF. The true HRFs used in the following subsections are similar up to a scaling factor, leading to different values of the CNR at a given noise level.

A white Gaussian noise of variance r_b has been added to the data, as well as a session-dependent low-frequency drift, which was generated from a cosine transform basis which coefficients $\boldsymbol{l}_i$ were drawn from a normal distribution. The amount of low frequency signal was tuned to be significant: we have checked for each session that the ratio between the quadratic norm of the drift components $\boldsymbol{P}_i\boldsymbol{l}_i$ and the quadratic norm of the drift-free data $\boldsymbol{X}_i\boldsymbol{h} + \boldsymbol{b}_i$ was no less than 50%. We stress here that the events are well distributed over time such that collinearity with the low frequency signal is unlikely.

Number Q_i depends on the chosen cutoff-period for the drift term present in session i. For instance, the selected parameters in row number 3 of Table III leads to $Q_i = 4$. The data are then obtained after undersampling the sequences at a $[\mathrm{TR}/\delta t]$ rate, the inter-scan interval being $\mathrm{TR} = 2$ s. Note also that the length of the datasets varies across sessions.

B. Statistical Properties of the Map Estimate

We assess the statistical properties of the ML and MAP estimates through a Monte Carlo study in the finite and asymptotic[9] data cases. The ML solution can be derived from (5) by setting $\lambda = r_b/\tau = 0$. In particular, we first outline the bias-variance tradeoff, which is intrinsic to Bayesian estimation in the finite data case and illustrate the asymptotic convergence of the MAP estimates to the true HRFs.

To obtain an approximation of the bias for the ML and MAP estimates, we have first computed the *average solution* over $J = 100$ realizations of the noise process

$$\widetilde{\boldsymbol{h}}_J^{\mathrm{MAP}} = \frac{1}{J}\sum_{j=1}^{J}\widehat{\boldsymbol{h}}_j^{\mathrm{MAP}} \quad \widetilde{\boldsymbol{h}}_J^{\mathrm{ML}} = \frac{1}{J}\sum_{j=1}^{J}\widehat{\boldsymbol{h}}_j^{\mathrm{ML}}.$$

[8]www.fil.ion.ucl.ac.uk/spm/spm99.html.

[9]There are generally two asymptotic situations. The first one occurs when the number of data $N \rightarrow +\infty$ whereas the second one amounts to increasing the CNR in order to characterize an estimator with noise-free data. Here, we have investigated the second case. Nonetheless, both situations are equivalent in our approach since hyperparameter $\lambda = r_b/\tau$ is automatically tuned from the data. This means that when $\mathrm{CNR} \rightarrow +\infty$, $\lambda \rightarrow 0$, which amounts to overweighting the likelihood term.

[7]These tools are available at http://www.lsa.umich.edu/psych/research\&labs/jjonides/download.html.

Fig. 1.   Average HRF estimates over $J = 100$ drawings of the noise distribution in the case of synchronous paradigm ($\Delta t = \mathrm{TR} = 1$ s) and for a low CNR value. (a)-(d): ML HRF estimates. (b)-(c) and (e)-(f): MAP HRF estimates computed with *constant* (b)-(e) and *adaptative* prior models (c)-(f). Solid and dashed lines represent $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, respectively. Fine and thick lines code for true and HRF estimates, respectively. In addition, the time samples of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are marked with $\star$ and $\square$, respectively.

The estimation variance of the MAP and ML solutions is then approximated using

$$\mathcal{V}\left(\widetilde{\boldsymbol{h}}_J{}^{\mathrm{MAP}}\right) = \frac{1}{J}\sum_{j=1}^{J}\left(\widehat{\boldsymbol{h}}_j{}^{\mathrm{MAP}} - \widetilde{\boldsymbol{h}}_J{}^{\mathrm{MAP}}\right)^2$$

$$\mathcal{V}\left(\widetilde{\boldsymbol{h}}_J{}^{\mathrm{ML}}\right) = \frac{1}{J}\sum_{j=1}^{J}\left(\widehat{\boldsymbol{h}}_j{}^{\mathrm{ML}} - \widetilde{\boldsymbol{h}}_J{}^{\mathrm{ML}}\right)^2$$

It follows that the quantities of interest, i.e., the mean square error for any HRF $\boldsymbol{h}_m$ can be computed using

$$\mathrm{MSE}(\widehat{\boldsymbol{h}};\boldsymbol{h}) \triangleq \mathrm{E}\left[\left\|\widehat{\boldsymbol{h}} - \mathrm{E}[\widehat{\boldsymbol{h}}]\right\|^2\right] \approx \mathcal{V}(\widetilde{\boldsymbol{h}}_J) + (\boldsymbol{h} - \widetilde{\boldsymbol{h}}_J)^2.$$

The global MSE (gMSE) is then obtained after averaging over all time points

$$\mathrm{gMSE}(\widehat{\boldsymbol{h}};\boldsymbol{h}) = \sum_{k=1}^{K}\frac{\mathrm{MSE}(\widehat{h}_k;h_k)}{K}.$$

To remove the *instant-matching error*, we have considered synchronous paradigms. The onsets of the stimuli and the data are defined on the same grid with $\Delta t = \mathrm{TR} = 1$ s. The ML and MAP estimates are now compared for low and high CNR values (see Table III).

1) Low CNR or "Finite Data" Case: Fig. 1 shows the average HRF estimates computed both for the ML and MAP strategies. As it clearly appears on Fig. 1(a), the ML solution is unbiased. The pointwise error bars that are depicted on Fig. 1(a) have been computed as the square roots of the variance of estimation $\mathcal{V}(\widetilde{\boldsymbol{h}}_j{}^{\mathrm{ML}})$. The average solutions depicted in Fig. 1(b), (c) correspond to the MAP estimates computed either for a *constant* prior model ($\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ have the same prior variance: $\tau_1 = \tau_2$) or for an *adaptative* prior model ($\tau_1 \neq \tau_2$). These average time courses illustrate the well-known intrinsic bias-variance tradeoff that appears in Bayesian solutions in the finite data case: the MAP solutions are biased but less variable than the ML estimate (their error bars, computed from $\mathcal{V}(\widetilde{\boldsymbol{h}}_J{}^{\mathrm{MAP}})$, are lower). To measure the gain brought by Bayesian methodology, we use the MSE and the summarizing index gMSE. We also note that $\boldsymbol{h}_2$ has been taken as an extreme example of HRF since it is much more peaked than the usual response, better represented by $\boldsymbol{h}_1$.

To emphasize the role played by the experimental paradigm in terms of bias-variance tradeoff, the same quantities have been computed for the block design and the results are depicted in Fig. 1(d)–(f). For the ML estimate plotted in Fig. 1(d), choosing a block design rather than an event-related one generates an increase of the variance of estimation and a slight bias. For the MAP estimates [Fig. 1(e)-(f)], the variance remains roughly

Event-related design

Block design
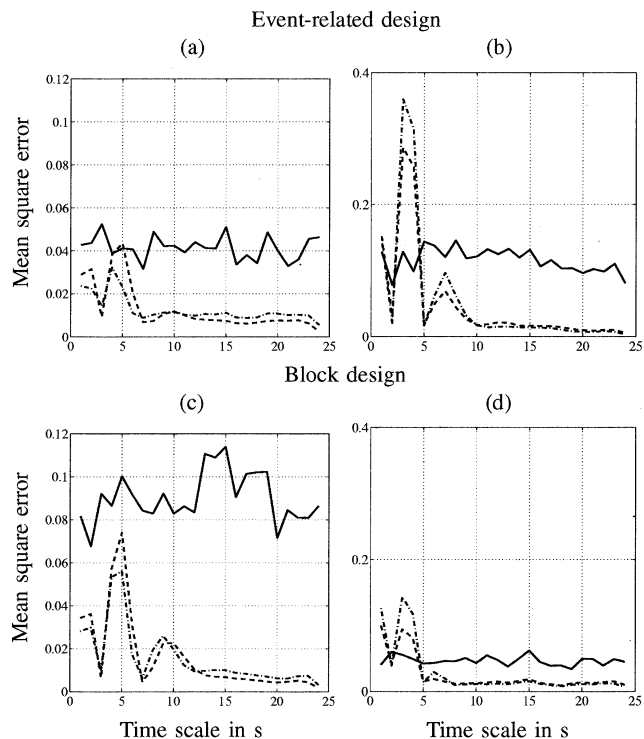
Fig. 2. Mean square error computed on h_1 in (a), (c) and on h_2 in (b), (d). Solid lines depict the MSE value for the ML estimates. Dash-dotted and dashed lines give the MSE values for the MAP estimates computed with constant and adaptative prior models, respectively.

TABLE IV
GLOBAL MSE OF THE HRF ESTIMATES COMPUTED BY ML AND MAP METHODOLOGIES

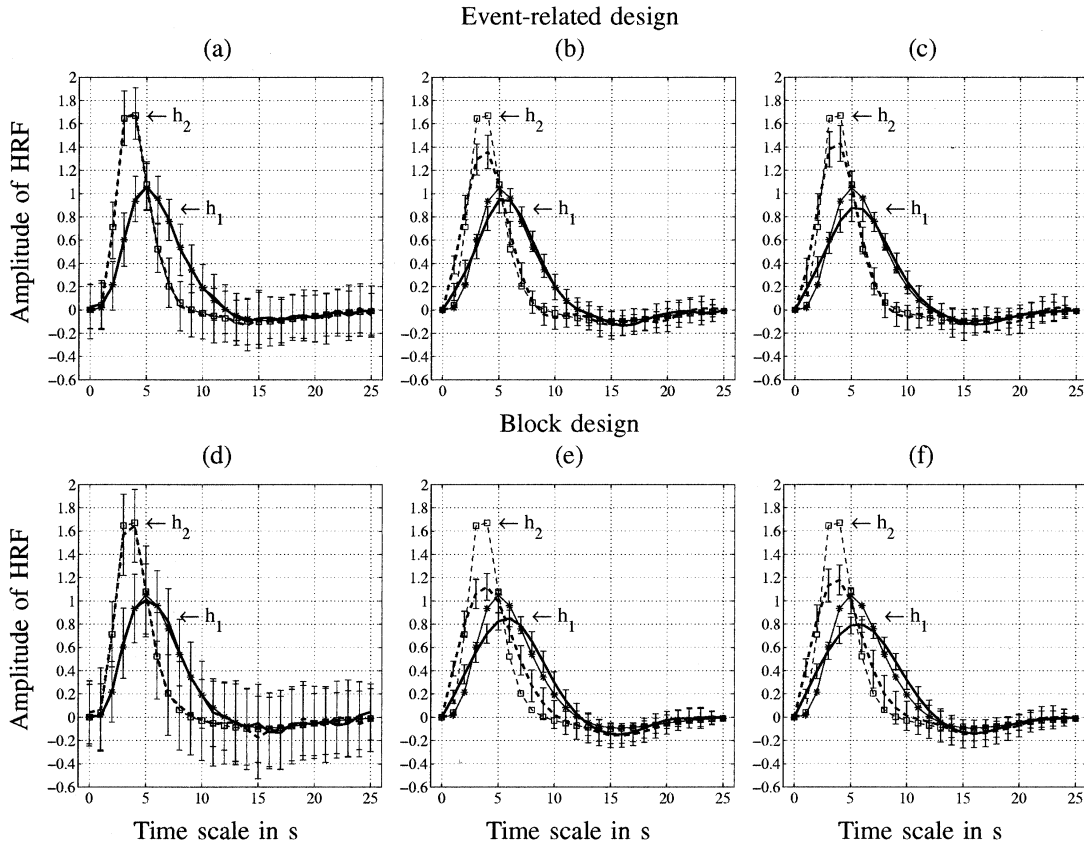|		$gMSE(\cdot\,;\boldsymbol{h}_m)\times 100$			
		ER design		Block design	
		$\boldsymbol{h}_1$	$\boldsymbol{h}_2$	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$
CNR=0.3	ML estimates	5.45	5.37	8.83	11.26
	MAP estimates ($\tau_1=\tau_2$)	1.47	3.39	1.58	5.31
	MAP estimates ($\tau_1\neq\tau_2$)	1.46	2.97	1.6	4.78
CNR=1.53	ML estimates	0.21	0.22		
	MAP estimates ($\tau_1=\tau_2$)	0.15	0.19		
	MAP estimates ($\tau_1\neq\tau_2$)	0.12	0.18		
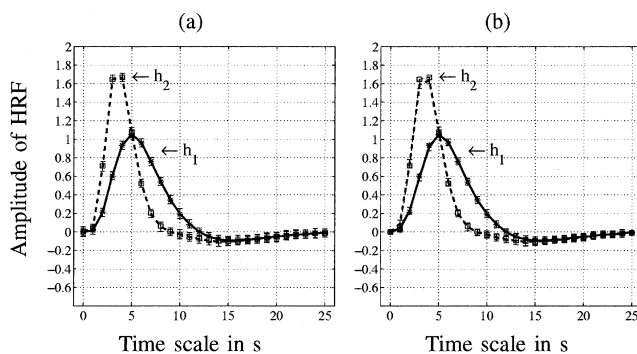
Fig. 3. Average HRF estimates over $J = 100$ drawings of the noise distribution in the case of synchronous paradigm ($\Delta t = \mathrm{TR} = 1$ s) and for a high CNR case. (a): ML HRF estimates. (b): MAP HRF estimates computed with a constant prior model. Solid and dashed lines represent $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$, respectively. Fine and thick lines code for true and HRF estimates, respectively but cannot be distinguished at this level of CNR. The time samples of $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ are marked with $\star$ and $\square$, respectively.

constant but the bias significantly increases compared with the event-related situation. To summarize, working with a poor design for HRF estimation amounts to decreasing the CNR both for ML and Bayesian procedures.

Fig. 2(a)-(b) provides the corresponding MSE over all time samples for h_1 and h_2, respectively. Fig. 2(a) shows that the MSE is always lower for the MAP solutions when dealing with h_1. The MAP estimate computed with an adaptative prior model has a larger bias around the peak but a lower gMSE compared with ML as reported in Table IV. By contrast, Fig. 2(b) shows that both MAP solutions have a larger MSE around the peak of h_2 even if the adaptative prior model allows to significantly reduce this error. As shown in Table IV, the MAP estimates still remain more attractive than the ML solution since their gMSE is always lower. It is also important to compare both MAP estimates and to note that the adaptative prior model provides better solutions in terms of gMSE at the expense of the computational cost (twice more expensive).

The MSE has also been computed for the block design experiment. As illustrated in Fig. 2(c)-(d) as well as in Table IV, these results show the significant increase of the mean square error both for the ML and MAP estimates, but more importantly for the ML solution. To conclude, the designer of the paradigm should carefully select the objective function to be optimized according to the underlying question (detection HRF estimation) [32].

2) High CNR or "Asymptotic" Data Case: Fig. 3(a) shows the average ML estimates in the nearly noise-free case. As illustrated, the ML estimate remains unbiased and is *asymptotically*

consistent since its estimation variance tends to zero. The average MAP estimates computed with a constant prior model[10] are plotted in Fig. 3(b) for comparison. As shown on this figure, there is a strong evidence that the MAP converges to the true HRF too, so that our solution is *asymptotically consistent*. It should be stressed that such a result is a direct consequence of the automatic tuning of the hyperparameters. Indeed, if $\lambda = r_b/\tau$ was kept constant (as in supervised estimation) the MAP estimate would be biased when the noise variance r_b decreases. Since $\lambda \to 0$ when $r_b \to 0$, the weight of the prior model tends to vanish, so that asymptotically the MAP estimate identifies with the ML solution. As reported in Table IV, the gMSE index provides similar results for both estimates.

C. How to Choose the HRF Sampling Rate?

The present simulation addresses tradeoff between the bias and computational burden depending on the chosen sampling period Δt given δt and TR. Here, we have considered a sufficiently high CNR to be able to assess the statistical properties from one single realization of the noise process.

Since the onsets of the trials occur with a temporal resolution of δt seconds, they are not synchronised with the data. Our aim

---

[10]The MAP solution computed with an adaptative prior model is not reported here since it provides the same average time courses.
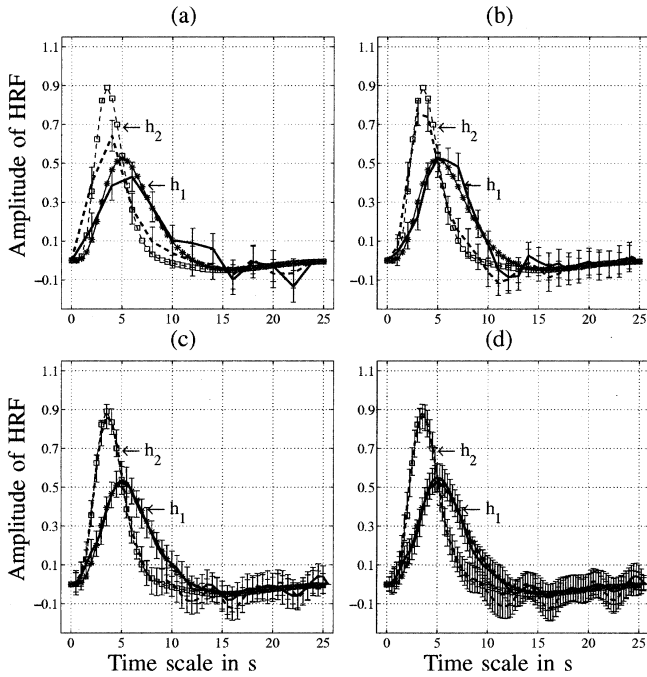
Fig. 4. HRF estimates for synchronous asynchronous estimation technique. Top row: HRFs computed using $\Delta t = \mathrm{TR}$ (a) and $\Delta t = \mathrm{TR}/2$ (b). Bottom row: HRFs estimated using $\Delta t = \mathrm{TR}/4$ (c) and $\Delta t = \mathrm{TR}/8$ (d). Solid and dashed lines represent h_1 and h_2, respectively. Fine and thick lines code for true and HRF estimates, respectively. The time samples of h_1 and h_2 are marked with $\star$ and $\square$, respectively.

TABLE V
QUADRATIC ERROR $\mathcal{E}$ AND MEAN STANDARD DEVIATION $\mathcal{S}$ OF THE MAP ESTIMATES COMPUTED FOR DIFFERENT VALUES OF Δt

	$\mathcal{E}(\cdot\,;\boldsymbol{h}_m)$		$\mathcal{S}(\widehat{\boldsymbol{h}}_m)$	
	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$
$\Delta t = \mathrm{TR}$	0.02	0.021	0.078	0.077
$\Delta t = \mathrm{TR}/2$	$9.3\ 10^{-3}$	$9.2\ 10^{-3}$	0.07	0.067
$\Delta t = \mathrm{TR}/4$	$3.5\ 10^{-3}$	$4.2\ 10^{-3}$	0.069	0.065
$\Delta t = \mathrm{TR}/8$	$3.1\ 10^{-3}$	$5.3\ 10^{-3}$	0.068	0.065

is high have been obtained in 0.03, 0.08, 0.33, and 3.1 s, respectively, on a Pentium IV 1-GHz. Decreasing the CNR induces a slower convergence, but for $\Delta t = \delta t$, the HRF estimate is usuallly computed in about one second for a dataset of 200 samples. The computational burden is, therefore, low enough to make this approach feasible for several hundred time series (voxels or region of interest).

Moreover, choosing $\Delta t = \widehat{\Delta t}$ compared with larger values leads to a better estimate of the noise variance r_b. This is likely to be due to the fact that some data variance may be introduced by the undersampling step.

Finally, we note that the hyperparameters ($\lambda = r_b/\tau$, if $\tau_m = \tau$, $\forall m \in \mathbb{N}_M^*$, or $\lambda_m = r_b/\tau_m$ otherwise) automatically adapt to the level of discretization. In the simulation performed for different sampling periods, we have checked that $\lambda_K \approx 8\lambda_{2K}$ (or $\lambda_{\Delta t} \approx 8\lambda_{\Delta t/2}$) is approximately satisfied.

D. Improvements Brought by Successive Models

The goal of this section is to classify the performances of the HRF estimates for the models that have been successively designed in Section II, in terms of quadratic error and variance reduction.

1) the influence of modeling drift terms in (2);
2) the *session-dependent* model (2), where each HRF is estimated from each session (before a possible averaging step over the sessions), compared with the *multisession* model ($\mathcal{M}$), where each HRF admits a single estimate derived from the whole set of sessions.

a) Comparison of rows 1–3 and 2–4 of Table VI proves that the quadratic error of the HRF estimate diminishes when modeling a drift component (more significantly for $\boldsymbol{h}_2$): the criterion $\mathcal{E}$ is smaller when the nuisance variables are jointly estimated. The variance reduction or the decrease of $\mathcal{S}$ is rather marginal, probably because the number of unknown parameters is larger when a trend is modeled with a constant number of data. These improvements are emphasized when dealing with multisession estimation. Fig. 5(b) illustrates the gain in robustness brought by model ($\mathcal{M}$), which takes a session-dependent drift into account, compared with the results when the drift has not been modeled in Fig. 5(c). Quantitative results reinforce these improvements, as shown in rows 7–8 of Table VI.

b) Rows 5–6 of Table VI as well as the comparison of Fig. 5(a)–(c) shows that multisession modeling has a greater influence on the accuracy of the HRF estimates than the consideration of nuisance variables in the model. Assuming stability

is to determine the optimal oversampling period $\widehat{\Delta t}$. This optimal value should be able to control the instant-matching error in the estimate, and help us to best set the tradeoff between bias reduction and potential variance increase.

Here, we have tested the single session asynchronous model (2), in which the drift component has been discarded. In addition, we have imposed identical prior variances ($\tau_1 = \tau_2$).

Fig. 4 shows a qualitative comparison between $\Delta t = \mathrm{TR}$ and $\Delta t < \mathrm{TR}$ models when dealing with asynchronous paradigms. The influence of the oversampling period appears clearly: the closer Δt is to δt, the less biased the estimate is.

We have quantitatively checked this result measuring the *quadratic error* $\mathcal{E}$ and the *mean standard deviation* $\mathcal{S}$ of the HRF estimates with

$$\mathcal{E}(\widehat{\boldsymbol{h}}_m;\boldsymbol{h}_m) = \frac{\|\widehat{\boldsymbol{h}}_m - \boldsymbol{h}_m\|^2}{K-1}, \quad m = 1,2$$

$$\mathcal{S}(\widehat{\boldsymbol{h}}_m) = \frac{1}{K-1}\sum_{p=1}^{K-1}\sigma_{m,p}. \qquad (17)$$

For $\boldsymbol{h}_1$ that fits well with the prior model, both criteria $\mathcal{E}$ and $\mathcal{S}$ decrease with Δt, even when the sampling period goes below $\widehat{\Delta t}$ (see Table V). By contrast, for $\boldsymbol{h}_2$, $\mathcal{E}$ decreases until $\Delta t = \delta t = \mathrm{TR}/4$ and increases for lower values, as reported in Table V. Therefore, selecting a sampling period under $\widehat{\Delta t} = \mathrm{TR}/4$ does not really improve the accuracy and robustness of the global estimate $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$. These results are also obtained for lower CNR values.

Choosing $\Delta t < \widehat{\Delta t}$ leads to an increase of the computational cost. More precisely, the results of Fig. 4(a)–(d) where the CNR

TABLE VI
QUANTITATIVE ASSESSMENT OF THE HRFs ESTIMATED BY THE SUCCESSIVE MODELS OF SECTION II

		Quadratic error $\mathcal{E}(\cdot\ ;\ \boldsymbol{h}_m)$		Dispersion $\mathcal{S}(\widehat{\boldsymbol{h}}_m)$	
		$\boldsymbol{h}_1$	$\boldsymbol{h}_2$	$\boldsymbol{h}_1$	$\boldsymbol{h}_2$
	Single Session & No drift & $\tau_1 = \tau_2$	0.017	0.018	0.13	0.13
	Single Session & No drift & $\tau_1 \neq \tau_2$	0.016	0.017	0.12	0.13
	Single Session & drift & $\tau_1 = \tau_2$	0.015	0.015	0.11	0.1
Fig.5(a)	Single Session & drift & $\tau_1 \neq \tau_2$	0.015	0.014	0.08	0.11
	multisession & No drift & $\tau_1 = \tau_2$	$5\ 10^{-3}$	0.01	0.1	0.1
Fig.5(c)	multisession & No drift & $\tau_1 \neq \tau_2$	$6\ 10^{-3}$	$8\ 10^{-3}$	0.08	0.1
	multisession & drift & $\tau_1 = \tau_2$	$5\ 10^{-3}$	$7\ 10^{-3}$	0.06	0.06
Fig.5(b)	multisession & drift & $\tau_1 \neq \tau_2$	$4\ 10^{-3}$	$6,5\ 10^{-3}$	0.04	0.05
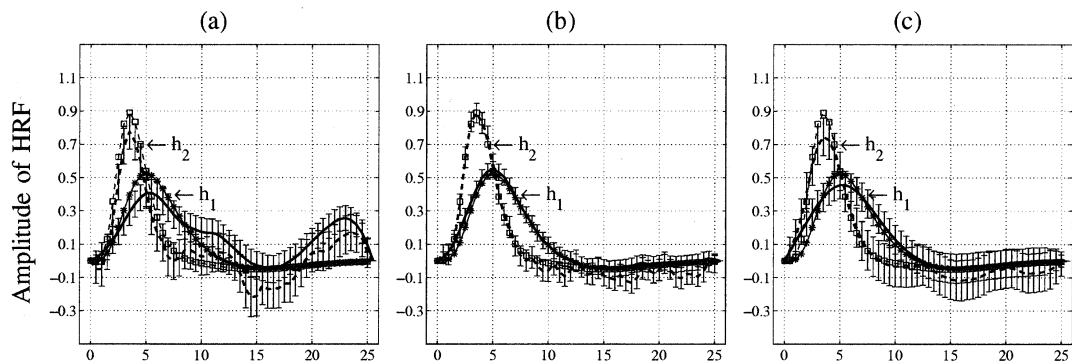
Fig. 5. (a),(b): comparison of the HRF estimates computed from *one* versus *four* sessions, respectively when the low frequency drift included in the data is modeled with $Q_i = 4$ nuisance variables for each session. (b),(c): comparison of drift modeled (b) and drift not modeled (c) for HRF estimates computed from four sessions. In all cases, we have considered the adaptative prior model ($\tau_1 \neq \tau_2$).

of the HRFs across sessions in a given voxel actually allows to bring more information and, therefore, leads to a significant bias reduction and a slight variance decrease. The variance reduction is even greater on data without drift component (results not shown). Nonetheless, the drift embedded in the data being random and session-varying, the $\mathcal{S}$ values reported in Table VI do not go down when a session-dependent drift has not been modeled. Comparison of Fig. 5(a)-(b) demonstrates the relevance of the HRF estimates computed from all available datasets ($I = 4$).

E. Validity of the Error Bars

A concern is to know whether the approximation leading to (8) has a good chance to be satisfied. This approximation is at the origin of the variance underestimation of $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$. Our argument is that with a small number of drift parameters compared with the number of data (when Q/N is small enough) and a standard CNR our approach provides an accurate estimate of the nuisance variables such that the extra source of error related to these parameters should be negligeable. Such cases are realistic and occur when $Q/N \leqslant 0.02$, for instance when the lowest frequency $f_{\min}$ is 1/120, $N = 180$ and $\mathrm{TR} = 2$ s we can choose $Q_i = 4$. Consequently, we are interested in quantifying the bias on the variance estimate provided by our ap-

proach. To achieve this goal, we performed several Monte Carlo studies[11] in which we analyzed the influence of the number of nuisance variables ($Q_i = 4$ and $Q_i = 18$) as well as the the noise level (CNR $= 0.73$ as in Section V-D and CNR $= 0.29$) on the estimation variance of $\underline{l}$. The drift terms were held to the same value for a complete run. First of all, we checked that our estimate $\widehat{\underline{l}}^{\mathrm{ML}}$ is unbiased. For a small number of drift terms ($Q_i = 4$, $Q_i/N_i = 0.02$) and a standard CNR (CNR $= 0.73$), the standard deviation of our nuisance variable estimate $\widehat{\underline{l}}^{\mathrm{ML}}$ was no larger than 3% of the maximal amplitude of the drift. This analysis only gives the behavior of our estimate $\widehat{\underline{l}}^{\mathrm{ML}}$, but even indirectly, the dispersion of $\widehat{\underline{l}}^{\mathrm{ML}}$ is related to the computed error bars on $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$. In such cases, our approach essentially provides an accurate approximation of the dispersion of $\widehat{\boldsymbol{h}}^{\mathrm{MMAP}}$ at low cost.

When increasing the noise level (CNR $= 0.29$) or multiplying the number of nuisance variables ($Q_i = 18$, $Q_i/N_i = 0.1$), we observed that the standard deviation of $\widehat{\underline{l}}^{\mathrm{ML}}$ was at least multiplied by three. In these more difficult cases, the solutions $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$ and $\widehat{\boldsymbol{h}}^{\mathrm{MMAP}}$ are close to each other but the main difference is that the error bars derived on $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$ will be underestimated since they will not be able to capture the dispersion at-

[11]$J = 10^3$ realizations of the noise have been drawn to accurately estimate the dispersion of the nuisance variable solution $\widehat{\underline{l}}^{\mathrm{ML}}$.
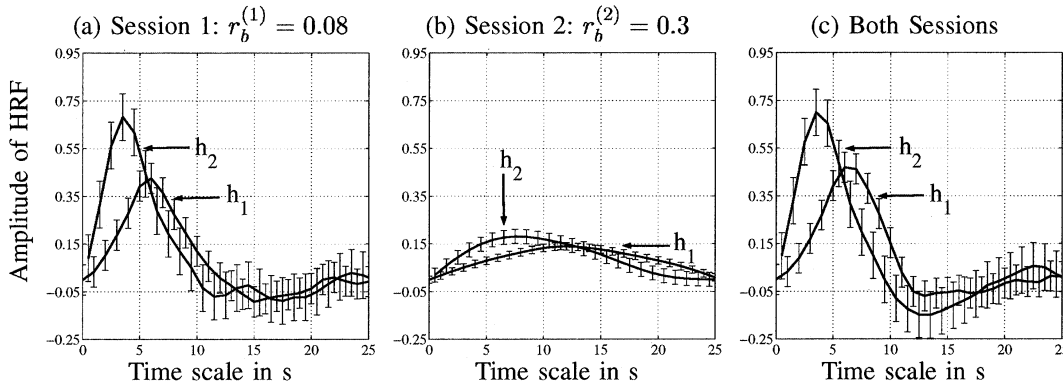
Fig. 6.   Testing for the robustness to the equal noise variance hypothesis. (a)-(b) HRFs estimated from a single session where the Gaussian noise has for variance $r_b^{(1)} = 0.08$ and $r_b^{(2)} = 0.3$, respectively. In (c), the HRFs have been estimated from both sessions with model ($\mathcal{M}$), that is assuming $r_b^{(1)} = r_b^{(2)}$.

tributable to the drift terms. In other words, the estimation variance of $\widehat{\underline{l}}^{\mathrm{ML}}$ is not taken into account in $\widehat{\boldsymbol{h}}^{\mathrm{MAP}}$. Hence, we should compute $\widehat{\boldsymbol{h}}^{\mathrm{MMAP}}$ and the marginal posterior covariance matrix $\Sigma_{\boldsymbol{H}}$[12] to remove the existing bias on our error bars. These developments are beyond the scope of this paper (see Section VII). We, therefore, limit the field of applicability of our technique to small number of nuisance variables ($Q_i \leqslant 5$) and reasonable CNR (roughly $\mathrm{CNR} \geqslant 0.5$).

F. Robustness to the Equal Noise Variance Hypothesis

We are also interested in testing the robustness of the HRF estimate to a departure from the hypothesis that the noise is constant across sessions. For doing so, we have considered two sessions with respective noise variances $r_b^{(1)} = 0.08$ and $r_b^{(2)} = 0.3$. We stress here that $r_b^{(2)} = 0.3$ is an extremely high noise level, not found in actual datasets but used to emphasize the robustness of our approach in a multisession framework.

As shown in Fig. 6(a) and as expected, the HRF estimates are closer to the original ones when the noise level is low. Fig. 6(b) demonstrates that increasing the noise variance provides oversmoothed results and stresses the limits of the prior model for such low CNR values. Moreover, small error bars that are visible in Fig. 6(b) result from overconfidence in the prior. In this extremely noisy situation, the problem discussed in Section V-E is amplified. Here, this variance underestimation concerns the hyperparameters and is due to the large uncertainty on the hyperparameter estimates. With a more important computational effort, this uncertainty could be computed using the methodology derived in [33]. The opposite result would be observed in the standard ML framework, where the higher the noise level, the larger the error bars. Note also that the noise variance is better estimated when the CNR is not too low. Finally, the HRFs plotted in Fig. 6(c) have been computed from both available datasets, assuming they have been generated with the same noise variances, since this assumption holds in model ($\mathcal{M}$). Clearly, Fig. 6(c) shows that the proposed technique remains robust with regard to some departures from this hypothesis. In other words, the multisession approach provides better results than the session-dependent technique, even with session-varying noise levels.

---

[12]Given by the left superior block of the joint covariance matrix $\Sigma_{\boldsymbol{H},\boldsymbol{L}}$ of size $M(K-1)Q \times M(K-1)Q$.

G. Overparametrization of the Number of HRFs

Here, we have used two trial types: one generating the peaked HRF $\boldsymbol{h}_2$, depicted in Fig. 5 for instance, and one yielding no response (zero function). Two different HRFs were modeled at the estimation stage. The question addressed here is whether modeling the nonexisting function as a HRF influences bias and variance of the estimate of $\boldsymbol{h}_2$. Indeed, it is well known in the ML framework that, the larger the number of unknowns, the higher the variance of estimation. Simulations (not reported here) show that inclusion in the model of a *uncorrelated* and *nonexistent* HRF does not have any influence on the estimation of the existing HRF, whatever the noise level.

Note that it is also possible to design a statistical test either for assessing whether an estimated HRF is zero or not ($\boldsymbol{h}_1 = 0$?), or for comparing both estimates ($\boldsymbol{h}_1 = \boldsymbol{h}_2$?), since the sum of squares of the difference between $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ follows a χ^2 distribution (see [34] for details).

VI. Experiment

A. MRI Parameters

The experiment was performed on a 3-T whole-body system (Bruker, Germany) equipped with a quadrature birdcage radio frequency (RF) coil and a head-gradient coil insert designed for echoplanar imaging. Functional images were obtained with a T2*-weighted gradient echo, echo planar imaging sequence (TR $= 3.3$ s, TE $= 30$ ms, FOV $= 240 \times 240$ mm^2, matrix $= 64 \times 64$). Each image, acquired in 1.3 s, comprised 22 4-mm-thick axial slices covering most of the brain. A high-resolution ($1 \times 1 \times 1.2$ mm) anatomical image using a 3-D gradient-echo inversion-recovery sequence, was also acquired for each participant.

B. Description of the Paradigm

The method was assessed on real data acquired in a speech discrimination experiment. The experiment consisted of six sessions ($I = 6$) comprising 100 trials ($N_i = 100$) lasting 3.3 seconds each. In each trial, the participant heard two pseudo-words over headphones. His task was to indicate whether he had perceived or not a difference between the two stimuli. There were three types of trials ($M = 3$): "Phonological," "Acoustic," and "Control." In trials belonging to the "Control" condition, the
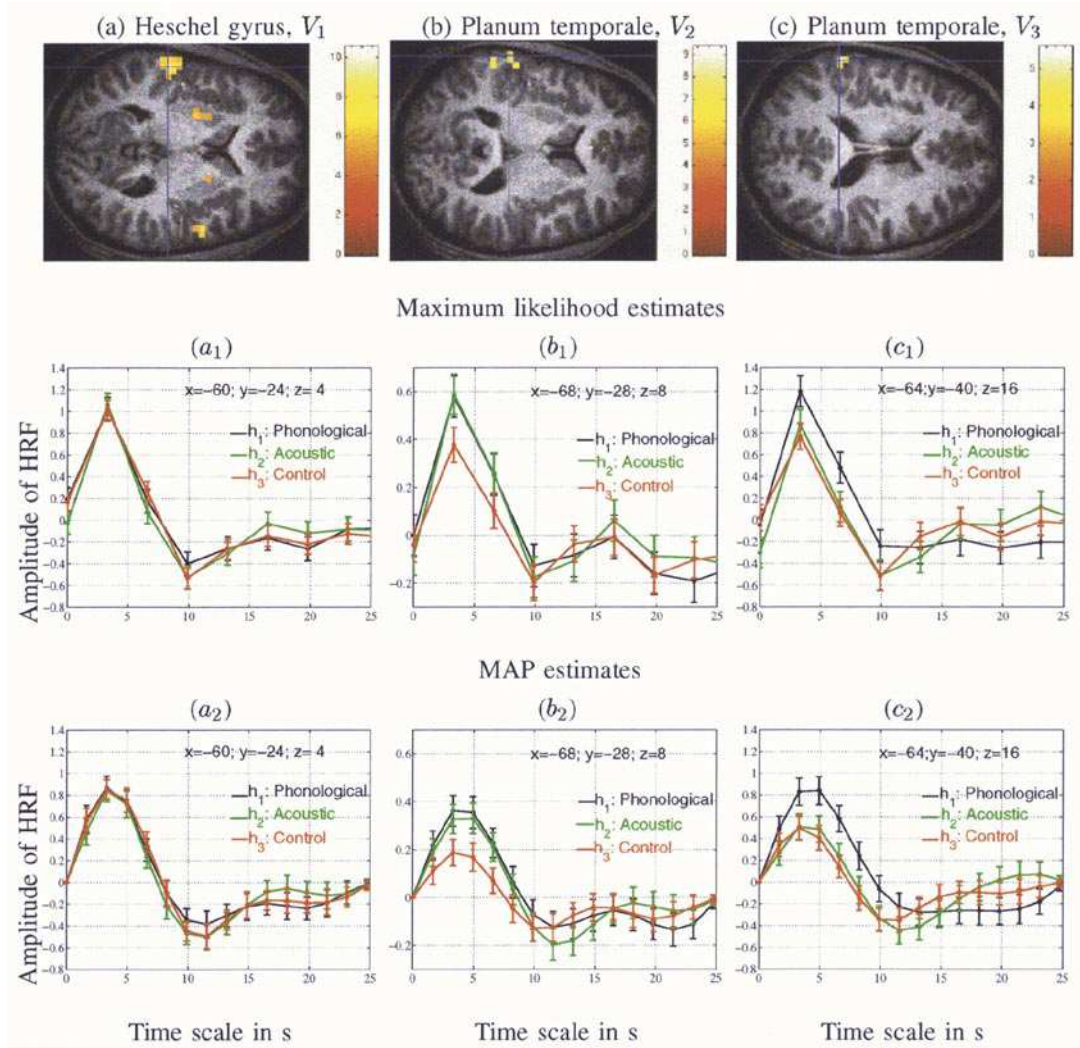
Fig. 7. Real data originating from a speech perception experiment. Top row: statistical t maps yielded by SPM 99 (thresholded at $P = 0.001$), superimposed on axial slices of averaged T1-weighted images. Middle row (a1)–(c1): Maximum likelihood HRF estimates computed from six sessions in voxels marked by blue crosses ($V_1 - V_3$). Bottom row (a2)–(c2): MAP HRF estimates computed from six sessions in the same voxels. Voxel coordinates are indicated near the time courses. "Phonological," "Acoustic," and "Contro" conditions are coupled with h_1, h_2, h_3, respectively.

two auditory stimuli in the pair were exactly the same. In the "Phonological" condition, the stimuli differed along a contrast used to distinguish words in the language of the participant (it was *linguistically relevant*: path versus bath in English). In the "Acoustic" condition, the stimuli also differed but the contrast between the stimuli was not relevant in the language of the participant (e.g., beat versus beet in English).

The stimuli pairs were presented during the silent gaps lasting 2 seconds between two succesive acquisitions (the TR was 3.3 s and the time of acquisition of one volume was 1.3 s). The onsets of events were aligned with the start of the second stimulus in a pair (i.e. at 1.65 s), which felt in between two successive acquisitions. $\Delta t = \mathrm{TR}/2$ is, therefore, an appropriate choice for the sampling period of the HRF.

C. Results

The top of Fig. 7 shows t maps (thresholded at $P = 0.001$ corrected for multiple comparisons) superimposed on axial slices of averaged T1-weighted images, computed with SPM99.

To obtain such results, one first has to specify an fMRI model that uses a canonical HRF $\boldsymbol{h}_c$. Second, least squares estimation and inference on relevant contrasts of the parameter estimates is performed. The first half of the regressors of the fMRI model is defined by $f_m(t_n) \triangleq (\boldsymbol{x}^{(m)} \star \boldsymbol{h}_c)(t_n)$ for $n = 1, \ldots, N = \Sigma_i N_i = 600$ and $m = 1, \ldots, M = 3$. The second half derives from the convolution of the first-order derivative of $\boldsymbol{h}_c$ with the onsets. The use of derivatives allows to gain robustness against variations of the maximum amplitude or the delay of this maximum.

The middle row of Fig. 7 shows the ML HRF estimates in three different voxels from the left superior temporal gyrus of one participant: (Talairach coordinates in millimeters: $V_1(X = -60,\ Y = -24,\ Z = 4)$; $V_2(X = -68,\ Y = -28,\ Z = 8)$; $V_3(X = -64,\ Y = -40,\ Z = 16)$. These time courses have been computed without oversampling, i.e., just by estimating the amplitude of each HRF at every TR. The bottom row of Fig. 7 shows the MAP HRF estimates in the same voxels. These results have been obtained from model ($\mathcal{M}$), that is discarding

the influence of the trend and the baseline and taking the six datasets into account. We have also considered the same prior variance for the three modeled hemodynamic responses.

Not surprisingly, Fig. 7(a1)-(a2) proves that the stimuli elicited very similar responses in Heschel gyrus (primary auditory cortex, V_1). The two other voxels were located in the planum temporale. Fig. 7(b1)-(b2) shows for V_2 that there is differential treatment when the stimuli differed, regardless of the type of difference (phonological or acoustic). By contrast, Fig. 7(c1)-(c2) shows a specific increment for phonological contrasts (speech processing).

The main error that appears on the ML solutions in Fig. 7(a1)–(c1) concerns the *hemodynamic* delay, i.e., the time-to-peak since the instant-matching error is about $\mathrm{TR}/2$. To circumvent this problem, the HRFs should also oversampled but this leads to an increased number of parameters and, thus, to unstable ML solutions. This situation, therefore, requires regularization.

VII. Discussion

In this paper, we have described and tested a general method for estimating the hemodynamic response function in fMRI data. The method is general enough to deal with all specific features of fMRI data, including the ability to work on several sessions and several experimental conditions in the context of an asynchronous sampling in event-related paradigm. In addition, our optimization scheme is sufficiently efficient to allow large fMRI time series to be processed. Finally, physiological artifacts can be correctly taken into account, provided that they are modeled with a small number of parameters (typically, four per session). This work has been implemented in a MATLAB toolbox[13] and interfaced with the SPM99 software.

To our knowledge, this work presents the only comprehensive robust nonparametric estimation of the fMRI brain response to a task or a stimulus. Applications of the technique are manifold. This approach should improve the observation of significant differences between the HRFs estimated for different stimuli within the same brain region. Using the HRF estimate to specify a *subject-dependent* general linear model, we could address the problem of the validity of the choice of the regressors[14] of such a model and may improve the localization of the signal (see [35] for a multivariate approach of this problem). In the domain of fusion of information (in particular with EEG signals) it is also important to have such a robust estimate of the HRF, to investigate the relationship between metabolic and electrical measurements.

Physiologically, the HRF should be zero-valued at $k = 0$. However, in practice, one sometimes detects voxels where this assertion is not true. The most common cause of this effect seems to be uncorrected stimulus-correlated motion. Since motion correction in fMRI data is hardly perfect, it is often important to remove the zero constraint at $k = 0$. In the proposed software, this constraint can be either changed to a less stringent prior (for instance, the first-order derivative at the extremal

---

[13]The HRF toolbox can be downloaded at the following URL site: http://www.madic.org/download/HRFTBx.

[14]These regressors has been selected empirically so far.

time samples of the HRF) or shifted in time such that the start of the HRF estimate can be chosen arbitrarily before the actual occurrence of the stimuli (for instance, a few seconds before). This allows the user to check the influence of this constraint on the HRF estimation.

The method can be extended in several ways:

First, to reduce the bias on the error bars we could resort to a second order Taylor expansion on the drift estimate as proposed for hyperparameters in [33]. This scheme needs an important computational effort. Alternatively, the usual Bayesian methodology leads to integrate the joint posterior distribution $p(\boldsymbol{h}, \underline{l} | y, \mathcal{M}; \boldsymbol{\theta})$ over $\underline{l}$. Further developments should be done to develop the MMAP estimate and check its ability to provide error bars that allow for the dispersion of the nuisance variables, in particular if a large number of nuisance variables is necessary to model drift terms as well as other additive effects of no interest. In this framework, a particular attention should be paid to hyperparameter estimation. If an EM algorithm was considered to perform this step, the *complete* data should be $(y, \boldsymbol{h}, \underline{l})$ and function $\mathcal{Q}$ should be defined by

$$
\begin{aligned}
\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^k; \underline{y}, \mathcal{M}) &= \cdots \\
&= \int \log p(\underline{y}, \boldsymbol{h}, \underline{l} | \mathcal{M}; \boldsymbol{\theta}) \\
&\quad \times p(\boldsymbol{h}, \underline{l} | \underline{y}, \mathcal{M}; \boldsymbol{\theta}^k) d\boldsymbol{h}\, d\underline{l} \\
&= \mathrm{E}_{\boldsymbol{H}, \underline{L}} \left[\log p(\underline{y}, \boldsymbol{h}, \underline{l} | \mathcal{M}; \boldsymbol{\theta}) | \underline{y}, \mathcal{M}; \boldsymbol{\theta}^k \right].
\end{aligned}
$$

Maximizing $\mathcal{Q}$ would require the computation of terms similar to (24), which could depend on the joint covariance matrix $\Sigma_{\boldsymbol{H}, \underline{L}}$ of size $M(K-1)Q \times M(K-1)Q$. This matrix is obtained after a block matrix inversion that will be more expensive than the computation of the covariance matrix $\boldsymbol{\Sigma}$ [see (5)] since the number of nuisance variables $Q = \Sigma_i Q_i$ would be larger than the number of HRF coefficients $M(K-1)$. On the other hand, since vector $\boldsymbol{\theta}$ would have a smaller size, the EM algorithm should converge in fewer iterations with a higher cost per iteration.

Second, another interesting extension would consist in considering a spatial model. Since the BOLD signal is known to have some spatial structure [36], estimation of the HRF over a region of interest should also provide a more robust estimation, as demonstrated by [37] who use a general prior on the spatial extension of the signal.

Third, the model presented here assumes that the response is constant in time. While this assumption is reasonable in a first instance (as long as the ISIs do not decrease below about two seconds), it is likely that there exists some variations in time due to physiological or neural adaptation to the stimulus or task. This is the subject of ongoing research.

Fourth, it is not yet clear if the responses coming from different subjects can easily be averaged in a single response. It may be that several subjects have too different brain responses such that the averaging of those signals would be difficult to interpret. The extension of the method to deal with multisubject data should be developed, although data can be analyzed at the same time using a different HRF per subject, with a generalization of model ($\mathcal{M}$) that takes different noise variances across

subjects into account. Such a generalization requires to overcome the actual difficulties encountered in group analysis: the most salient ones consist in removing subject-dependent movement artifacts and normalizing all subjects in the same space reference.

Lastly, we hope that this method can be the basis for some work that would take advantage of the recent advance in joint recording of electrical activity at the surface of the scalp in the MR scanner. In particular, using such (joint) recordings, it may be possible to partially retrieve the local field potential information from the BOLD signal, which would provide a better understanding of the neural computation well resolved in time and space.

APPENDIX I
DECOUPLING THE COMPLETE LIKELIHOOLD OPTIMIZATION

Vector $\widetilde{\boldsymbol{\theta}}$ can be partitioned into two subvectors $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$ and $\boldsymbol{\theta}_{\boldsymbol{H}}$ which respectively control the conditional pdfs $p(\underline{y}|\boldsymbol{h}, \mathcal{M}; \boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}})$ and $p(\boldsymbol{h}; \boldsymbol{R}, \boldsymbol{\theta}_{\boldsymbol{H}})$. The M-step of the EM algorithm can be divided into two simpler independent maximization problems. The complete likelihood which enters in the definition of $\mathcal{Q}$ in (12) can be expressed as

$$p(\underline{y}, \boldsymbol{h}|\mathcal{M}; \widetilde{\boldsymbol{\theta}}) = p(\underline{y}|\boldsymbol{h}, \mathcal{M}; \boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}) p(\boldsymbol{h}; \boldsymbol{R}, \boldsymbol{\theta}_{\boldsymbol{H}}). \quad (18)$$

For any set value of parameters vector $\widetilde{\boldsymbol{\theta}}^0$, define functions $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ and $\mathcal{Q}_{\boldsymbol{H}}$ as

$$\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}(\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}^0; \underline{y}) = \mathrm{E}\left[\log p(\underline{y}|\boldsymbol{h}, \mathcal{M}; \boldsymbol{\theta}_{Y|\boldsymbol{H}})|\underline{y}; \widetilde{\boldsymbol{\theta}}^0\right] \quad (19)$$

$$\mathcal{Q}_{\boldsymbol{H}}(\boldsymbol{\theta}_{\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}^0; \underline{y}) = \mathrm{E}\left[\log p(\boldsymbol{h}; \boldsymbol{R}, \boldsymbol{\theta}_{\boldsymbol{H}})|\underline{y}; \widetilde{\boldsymbol{\theta}}^0\right]. \quad (20)$$

It can be immediately deduced from (12) and (18) that function $\mathcal{Q}$ can be expressed as

$$\mathcal{Q}(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\theta}}^0; \underline{y}) = \mathcal{Q}_{\underline{Y}|\boldsymbol{H}}\left(\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}^0; \underline{y}\right) + \mathcal{Q}_{\boldsymbol{H}}\left(\boldsymbol{\theta}_{\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}^0; \underline{y}\right) \quad (21)$$

which shows that the M step of the EM algorithm can be decoupled into two operations: maximization of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$ and maximization of $\mathcal{Q}_{\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}_{\boldsymbol{H}}$.

APPENDIX II
UPDATING $\boldsymbol{\theta}_{\underline{Y}|\boldsymbol{H}}$

In the present case, $\underline{y}$ is made up of independent realizations $(\boldsymbol{y}_i)_{1 \leqslant i \leqslant I}$. As a consequence, the expression of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ can be rewritten as

$$\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}\left(\boldsymbol{\theta}'_{\underline{Y}|\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}; \underline{y}\right) = \sum_{i=1}^{I} \mathcal{Q}_{Y_i|\boldsymbol{H}}\left(\boldsymbol{\theta}'_{Y_i|\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}; \underline{y}\right). \quad (22)$$

Handling successively the likelihood (3), the Gaussianity of $\boldsymbol{H}$, definition (19) and (22), we are able to derive a closed-form expression for $\mathcal{Q}_{Y_i|\boldsymbol{H}}$

$$\mathcal{Q}_{Y_i|\boldsymbol{H}}\left(\boldsymbol{\theta}'_{Y_i|\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}; \underline{y}\right) = -\frac{1}{2}\left[N_i \log r'_b + \frac{1}{r'_b}\mathcal{E}_i\left(\boldsymbol{l}'_i\right)\right] \quad (23)$$

where function $\mathcal{E}_i(\cdot)$ is defined by the following expectation:

$$\mathcal{E}_i\left(\boldsymbol{l}'_i\right) = \mathrm{E}_{\boldsymbol{H}}\left[\left\|\boldsymbol{y}_i - \boldsymbol{P}_i\boldsymbol{l}'_i - \boldsymbol{X}_i\boldsymbol{h}\right\|^2 |\underline{y}, \widetilde{\boldsymbol{\theta}}\right]. \quad (24)$$

Since the posterior pdf $p(\boldsymbol{h}|\underline{y}, \mathcal{M}; \widetilde{\boldsymbol{\theta}})$ is $\mathcal{N}(\widehat{\boldsymbol{h}}^{\mathrm{MAP}}, \boldsymbol{\Sigma})$-distributed, analytic calculations give for $\mathcal{E}_i$

$$\mathcal{E}_i(\boldsymbol{l}'_i) = \left\|\boldsymbol{y}_i - \boldsymbol{P}_i\boldsymbol{l}'_i - \boldsymbol{X}_i\widehat{\boldsymbol{h}}^{\mathrm{MAP}}\right\|^2 + \mathrm{tr}\left(\boldsymbol{X}_i\boldsymbol{\Sigma}\boldsymbol{X}_i^t\right).$$

Here, we would like to stress the separability property, expressed through (22), that allows to think about a parallel update of vectors $\boldsymbol{\theta}_{Y_i|\boldsymbol{H}}$, for $i \in \mathbb{N}_I^*$, except for r_b. Such a property results from model $(\mathcal{M})$ and implies that the maximization of $\mathcal{Q}_{\underline{Y}|\boldsymbol{H}}$ is block-wise decoupled for vectors $\boldsymbol{l}_i$. Finally, replacing $\mathcal{E}_i$ by its value in (23), taking the first derivative of $\mathcal{Q}_{Y_i|\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}'_{Y_i|\boldsymbol{H}}$ and equating to zero yields (14) and (15) for $\boldsymbol{l}_i$ and r_b, respectively.

APPENDIX III
UPDATING $\boldsymbol{\theta}_{\boldsymbol{H}}$

From the prior pdf (4) and the definition (20), function $\mathcal{Q}_{\boldsymbol{H}}$ is defined by

$$\mathcal{Q}_{\boldsymbol{H}}\left(\boldsymbol{\theta}'_{\boldsymbol{H}}, \widetilde{\boldsymbol{\theta}}; \underline{y}\right) = -\frac{(K-1)}{2}\sum_{m=1}^{M}\log\tau'_m - \frac{\mathcal{E}\left(\boldsymbol{\theta}'_{\boldsymbol{H}}\right)}{2}$$
$$-\frac{M}{2}\log\det(\boldsymbol{R}) \quad (25)$$

with

$$\mathcal{E}\left(\boldsymbol{\theta}'_{\boldsymbol{H}}\right) = \mathrm{E}_{\boldsymbol{H}}\left[\boldsymbol{h}^{\mathrm{t}}\boldsymbol{R}'_{\boldsymbol{H}}\boldsymbol{h}|\underline{y}, \mathcal{M}, \widetilde{\boldsymbol{\theta}}\right]$$
$$= (\widehat{\boldsymbol{h}}^{\mathrm{MAP}})^{\mathrm{t}}\boldsymbol{R}'_{\boldsymbol{H}}\widehat{\boldsymbol{h}}^{\mathrm{MAP}} + \mathrm{tr}\left(\boldsymbol{R}_{\boldsymbol{H}|\underline{Y}}\boldsymbol{R}'_{\boldsymbol{H}}\right) \quad (26)$$

and $\boldsymbol{R}'_{\boldsymbol{H}} = [\tau'_1\boldsymbol{R}, \tau'_2\boldsymbol{R}, \ldots, \tau'_M\boldsymbol{R}]$. Here again, replacing $\mathcal{E}$ by (26) in (25) and taking the first derivative of $\mathcal{Q}_{\boldsymbol{H}}$ with respect to $\boldsymbol{\theta}'_{\boldsymbol{H}}$ allow us to find the updating (16).

ACKNOWLEDGMENT

REFERENCES

[1] B. R. Rosen, R. L. Buckner, and A. M. Dale, "Event-related functional MRI: past, present and future," *Proc. Nat. Acad. Sci. USA*, vol. 95, pp. 773–780, 1998.

[2] S. Ogawa, T. Lee, A. Kay, and D. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proc. Nat. Acad. Sci. USA*, vol. 87, no. 24, pp. 9868–9872, 1990.

[3] P. A. Bandettini, A. Jesmanowicz, E. C. Wong, and J. S. Hyde, "Processing strategies for time-course data sets in functional mri of the human brain," *Magn. Reson. Med.*, vol. 30, pp. 161–173, 1993.

[4] R.L. Buckner, J. Goodman, M. Burock, M. Rotte, W. Koutstaal, D. L. Schachter, B. R. Rosen, and A. M. Dale, "Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI," *Neuron*, vol. 20, pp. 285–296, Feb. 1998.

[5] R. Buxton and L. Frank, "A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation," *J. Cereb. Blood Flow Metab.*, vol. 17, no. 1, pp. 64–72, 1997.

[6] R. B. Buxton, E. C. Wong, and F. L. R, "Dynamics of blood flow and oxygenation changes during brain activation: the balloon model," *Magn. Reson. Med.*, vol. 39, pp. 855–864, June 1998.

[7] T.-Q. Li, T. N. Haefelin, B. Chan, A. Kastrup, T. Jonsson, G. Glover, and M. E. Moseley, "Assessment of hemodynamic response during focal neural activity in human using bolus tracking, arterial spin labeling, and BOLD techniques," *NeuroImage*, vol. 12, no. 4, pp. 442–451, Oct. 2000.

[8] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann, "Neurophysiological investigation of the basis of the fMRI signal," *Nature*, vol. 412, no. 6843, pp. 150–157, July 12, 2001.

[9] K. J. Friston, "Statistical parametric mapping," in *Functional Neuroimaging: Technical Foundations*, R. Thatcher, M. Hallet, T. Zeffiro, E. John, and M. Huerta, Eds. Orlando, FL: Academic, 1994, pp. 79–93.

[10] G. M. Boynton, S. A. Engel, G. H. Glover, and D. J. Heeger, "Linear systems analysis of functional magnetic resonance imaging in human V1," *J. Neurosci.*, vol. 16, pp. 4207–4221, 1996.

[11] N. Lange, "Empirical and substantive models, the Bayesian paradigm, and meta-analysis in functional brain imaging," *Hum. Brain Mapp.*, vol. 5, pp. 259–263, 1997.

[12] M. S. Cohen, "Parametric analysis of MRI data using linear systems methods," *NeuroImage*, vol. 6, pp. 93–103, 1997.

[13] J. C. Rajapakse, F. Kruggel, J. M. Maisog, and D. Von Cramon, "Modeling hemodynamic response for analysis of functional MRI time-series," *Hum. Brain Mapp.*, vol. 6, pp. 283–300, 1998.

[14] G. H. Glover, "Deconvolution of impulse response in event-related BOLD fMRI," *NeuroImage*, vol. 9, pp. 416–429, 1999.

[15] C. Goutte, F. A. Nielsen, and L. K. Hansen, "Modeling the haemodynamic response in fMRI using smooth filters," *IEEE Trans. Med. Imag.*, vol. 19, pp. 1188–1201, Dec. 2000.

[16] C. Gössl, L. Fahrmeir, and D. P. Auer, "Bayesian modeling of the hemodynamic response function in BOLD fMRI," *NeuroImage*, vol. 14, pp. 140–148, 2001.

[17] F. A. Nielsen, L. K. Hansen, P. Toft, C. Goutte, N. Lange, S. C. Stroher, N. Morch, C. Svarer, R. Savoy, B. Rosen, E. Rostrup, and P. Born, "Comparison of two convolution models for fMRI time series," *NeuroImage*, vol. 5, p. S473, 1997.

[18] G. Marrelec, H. Benali, P. Ciuciu, and J.-B. Poline, "Bayesian estimation of the hemodynamic response function in functional MRI," in *Bayesian Inference and Maximum Entropy Methods*, R. Fry, Ed. Baltimore, MD: MaxEnt Workshops, 2001.

[19] E. T. Bullmore, M. Brammer, S. C. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," *Magn. Reson. Med.*, vol. 35, pp. 261–277, 1996.

[20] R. Birn, R. Cox, and P. A. Bandettini, "Detection versus estimation in event-related fMRI: choosing the optimal stimulus timing," *NeuroImage*, vol. 15, no. 1, pp. 252–264, Jan. 2002.

[21] K. Worsley, C. Liao, J. Aston, V. Petre, G. Duncan, F. Morales, and A. Evans, "A general statistical analysis for fMRI data," *NeuroImage*, vol. 15, no. 1, pp. 1–15, Jan. 2002.

[22] E. Zarahn, G. K. Aguirre, and M. D'Esposito, "Empirical analysis of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions," *NeuroImage*, vol. 5, pp. 179–197, 1997.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[24] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 2988–2998, Dec. 1996.

[25] S. S. Saquib, C. A. Bouman, and K. D. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," *IEEE Trans. Image Processing*, vol. 7, pp. 1029–1044, July 1998.

[26] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. B*, vol. 39, pp. 1–38, 1977.

[27] J. Idier, Y. Goussard, and A. Ridolfi, "Unsupervised image segmentation using a telegraph parameterization of Pickard random fields," in *Spatial Statistics. Methodological Aspects and Some Applications, ser. Lecture Notes in Statistics*. New York: Springer-Verlag, 2001, vol. 159, pp. 115–140.

[28] X.-L. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: a general framework," *Biometrika*, vol. 80, pp. 267–278, 1993.

[29] C. F. J. Wu, "On the convergence of the EM algorithm," *Ann. Statist.*, vol. 11, no. 1, pp. 95–103, 1983.

[30] T. Liu, L. Frank, E. C. Wong, and R. B. Buxton, "Detection power, estimation efficiency, and predictability in event-related fMRI," *NeuroImage*, vol. 13, no. 4, pp. 759–773, Apr. 2001.

[31] A. M. Dale, "Optimal experimental design for event-related fMRI," *Hum. Brain Mapp.*, vol. 8, pp. 109–114, 1999.

[32] T. D. Wager and T. E. Nichols, "Optimization of experimental design in fMRI: a general framework using a genetic algorithm," *NeuroImage*, vol. 18, pp. 293–309, 2003.

[33] R. E. Kass and D. Steffey, "Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models)," *J. Amer. Statist. Soc.*, vol. 84, no. 407, pp. 717–726, Sept. 1989.

[34] G. Marrelec, H. Benali, P. Ciuciu, M. Plgrini-Issac, and J.-B. Poline, "Robust bayesian estimation of the hemodynamic response function in event-related BOLD MRI using basic physiological information," *Hum. Brain Mapp.*, vol. 19, no. 1, pp. 1–17, May 2003.

[35] F. Kherif, J.-B. Poline, G. Flandin, H. Benali, O. Simon, S. Dehaene, and K. Worsley, "Multivariate model specification for fMRI data," *NeuroImage*, vol. 16, no. 4, pp. 1068–1083, Aug. 2002.

[36] X. Descombes, F. Kruggel, and D. Y. von Cramon, "Spatio-temporal fMRI analysis using Markov random fields," *IEEE Trans. Med. Imag.*, vol. 17, pp. 1028–1039, Dec. 1998.

[37] C. Gössl, D. P. Auer, and L. Fahrmeir, "Bayesian spatio-temporal inference in functional magnetic resonance imaging," *Biometrics*, vol. 57, 2001.

VII.3   Vision sous l'angle des réseaux bayésiens

[a.7] G. Marrelec, P. Ciuciu, M. Pélégrini-Issac et H. Benali, « Estimation of the Hemodynamic Response in Event-Related Functional MRI : Bayesian Networks as a Framework for Efficient Bayesian Modeling and Inference », *IEEE Trans. Med. Imag.*, vol. 23, 8, pp. 959–967, août 2004.

Estimation of the Hemodynamic Response in Event-Related Functional MRI: Bayesian Networks as a Framework for Efficient Bayesian Modeling and Inference

Guillaume Marrelec*, Philippe Ciuciu, *Member, IEEE*, Mélanie Pélégrini-Issac, and Habib Benali, *Senior Member, IEEE*

Abstract—A convenient way to analyze blood-oxygen-level-dependent functional magnetic resonance imaging data consists of modeling the whole brain as a stationary, linear system characterized by its transfer function: the hemodynamic response function (HRF). HRF estimation, though of the greatest interest, is still under investigation, for the problem is ill-conditioned. In this paper, we recall the most general Bayesian model for HRF estimation and show how it can beneficially be translated in terms of Bayesian graphical models, leading to 1) a clear and efficient representation of all structural and functional relationships entailed by the model, and 2) a straightforward numerical scheme to approximate the joint posterior distribution, allowing for estimation of the HRF, as well as all other model parameters. We finally apply this novel technique on both simulations and real data.

Index Terms—Bayesian inference, Bayesian networks, functional MRI, hemodynamic response function.

I. INTRODUCTION

FUNCTIONAL magnetic resonance imaging (fMRI) is a noninvasive technique allowing for the evolution of brain processes to be dynamically followed in various cognitive and behavioral tasks [1]. In the most common fMRI technique, based on the so-called blood-oxygen-level-dependent (BOLD) contrast, the measure is only indirectly related to neuronal activity through a process that is still under investigation [2]–[4]. For this reason, a convenient way to analyze BOLD fMRI data consists of modeling the whole brain as a stationary, linear "black box" system characterized by its transfer response function, also called hemodynamic response function (HRF)

[5]. This model, called general linear model (GLM), fairly well accounts for the properties of the real system as long as the inter-stimulus interval does not decrease beyond about two seconds [6], [7]. When this constraint is not respected, other models have to be developed [8], [9].

Estimation of the HRF is of the greatest interest when analyzing fMRI data, since it can give a deep insight into the underlying dynamics of brain activation and the relationships between activated areas. HRFs are increasingly suspected to vary from region to region, from task to task, and from subject to subject [10]–[12]. Age and disease are also more and more believed to have a significant influence on the BOLD response [13], [14]. Nevertheless, accurate estimation of the response function still belongs to ongoing research, since the problem is badly conditioned. Various nonparametric methods have been developed so far in an attempt to infer the HRF at each time sample, such as selective averaging [6], averaging over regions [15], introduction of nondiagonal models for the temporal covariance of the noise [16], or temporal regularization [17].

In [18] and [19], we proposed a Bayesian nonparametric estimation of the HRF for event-related designs. Basic yet relevant physiological information was introduced to temporally constrain the problem and calculate robust estimators of the parameters of interest. In [20]–[22], the model was extended to account for asynchronous event-related designs, different trial types, and several fMRI sessions, further improving the estimation. For calculation reasons, all variants proposed so far have, however, the drawback of not integrating the hyperparameter uncertainty. Furthermore, probabilistic treatment of the drift parameters in the extended model was possible [23], [24], but at a significantly higher computational cost.

In this paper, we propose to cast a new light on the GLM. We still place ourselves in a Bayesian framework, permitting integration of information originating from various sources and efficient inference on the parameters of interest. A general model is set to account for most event-related fMRI data. In a conventional Bayesian approach, we would then calculate the joint posterior distribution of all parameters, which would be the pivotal quantity for all further inference. Since direct sampling from this probability density function (pdf) would prove impossible, Monte Carlo Markov chain (MCMC) sampling would be required, such as Gibbs sampling [25], [26]. In this case, posterior conditional pdfs should be derived. In this perspective,

we advocate that calculation of the posterior pdf is unncessary. Instead, we resort to a novel approach that focuses on graphical modeling and that, once the model has been properly set, makes it possible to directly lead probabilistic inference about all parameters. More precisely, we utilize graph theory [27] to conveniently deal with the model. Indeed, graphs give a very simple and efficient representation of the model, however complex it may be. In this framework, we translate the model into a Bayesian network. Using Markov properties of such networks, drawing inference becomes straightforward, and Gibbs sampling finally provides us with a numerical approximation of the joint posterior pdf.

In the next part of this paper, we develop the general Bayesian framework for HRF estimation, presenting an extended version of the GLM. In Section III, the GLM is translated in terms of graphical model, and it is shown how inference can readily be performed from there. We briefly present simulations and finally apply our model to real data.

II. HRF ESTIMATION IN fMRI DATA ANALYSIS

A. Notations

In the following, x denotes a real number, $\boldsymbol{x}$ a vector, and $\boldsymbol{X}$ a matrix. For the sake of simplicity, (x_i)—between parentheses—is a shortcut for $(x_i)_{1 \leq i \leq I}$. "t" is the regular matrix transposition. $\boldsymbol{I}_N$ stands for the $N \times N$ identity matrix. "$\propto$" relates two expressions that are proportional. For two variables x and y, "$x \mid y$" stands for "x given y," and $\mathrm{p}(x)$ for the probability of x. $\mathcal{N}(\boldsymbol{m}, \boldsymbol{V}; \boldsymbol{x})$ is the Gaussian density function with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{V}$ calculated at sample $\boldsymbol{x}$. Inv-$\chi^2(d, r^2; u)$ is the scaled inverse-chi-square density function[1] with d degrees of freedom and scale parameter r^2 evaluated at sample u.

B. General Linear Model

Let an fMRI experiment be composed of S sessions, each session involving I different stimulus types. Define $\boldsymbol{y}_s = (y_{s,t_{s,n}})_{1 \leq n \leq N_s}$ as the BOLD fMRI time course of a voxel (i.e., volume element) at (not-necessarily uniformly sampled) times $(t_{s,n})$ for session s, and $\boldsymbol{x}_{s,i} = (x_{s,i,t})_{t_{s,0} \leq t \leq t_{s,N_s}}$ the corresponding binary time series, composed of the ith stimulus onsets. The following discrete linear convolution model (H) is assumed to hold between the stimuli and the data

$$y_{s,t_{s,n}} = \sum_{i=1}^{I} \sum_{k=0}^{K_i} x_{s,i,t_{s,n}-k\Delta t} h_{i,k\Delta t} + \sum_{m=1}^{M_s} d_{m,t_{s,n}} \lambda_{s,m} + e_{t_n}$$

$n = n_s + 1, \ldots, N_s$, where n_s is the largest integer so, that $t_{s,n} - K_i \Delta t < t_{s,1}$ for all i. The $(K_i + 1)$-dimensional vector $\boldsymbol{h}_i = (h_{i,k\Delta t})^t$ represents the ith unknown HRF to be estimated, sampled every Δt. All HRFs are assumed to be constant across sessions. $L_s = N_s - n_s$ is the actual amount of data used in the calculation for each session. $\boldsymbol{X}_{s,i} = (x_{s,i,t_{s,n}-k\Delta t})$ is the regular L_s-by-$(K_i + 1)$ design matrix, consisting of the lagged stimulus covariates. In the L_s-by-M_s matrix $\boldsymbol{D}_s = (d_{m,t_{s,n}})$

are the values at times $(t_{s,n})$ of a basis of M_s functions that takes a potential drift and any other nuisance effect into account, and $\boldsymbol{\lambda}_s = (\lambda_{s,m})^t$ contains the corresponding coefficients. For the sake of simplicity, the basis is assumed to be orthonormal, so that $(1/L_s)\boldsymbol{D}_s \boldsymbol{D}_s^t = \boldsymbol{I}_{L_s}$. Vector $\boldsymbol{e}_s = (e_{s,t_{s,n}})^t$ accounts for noise and is supposed to consist of independent and identically distributed Gaussian variables of unknown variance σ_s^2, assumed to be independent from the HRFs. In matrix form, (H) boils down to

$$\boldsymbol{y}_s = \sum_{i=1}^{I} \boldsymbol{X}_{s,i} \boldsymbol{h}_i + \boldsymbol{D}_s \boldsymbol{\lambda}_s + \boldsymbol{e}_s, \quad s = 1, \ldots, S$$

also called general linear model (GLM). In this model, the likelihood of the data yields

$$\mathrm{p}\left((\boldsymbol{y}_s) \mid H, (\boldsymbol{h}_i), (\sigma_s^2), (\boldsymbol{\lambda}_s)\right)$$
$$= \prod_{s=1}^{S} \mathrm{p}\left(\boldsymbol{y}_s \mid H, (\boldsymbol{h}_i), \sigma_s^2, \boldsymbol{\lambda}_s\right) \quad (1)$$

with each term in the product reading

$$\mathrm{p}\left(\boldsymbol{y}_s \mid H, (\boldsymbol{h}_i), \sigma_s^2, \boldsymbol{\lambda}_s\right)$$
$$= \mathcal{N}\left(\sum_{i=1}^{I} \boldsymbol{X}_{s,i} \boldsymbol{h}_i + \boldsymbol{D}_s \boldsymbol{\lambda}_s, \sigma_s^2 \boldsymbol{I}_{L_s}; \boldsymbol{y}_s\right). \quad (2)$$

C. HRFs and Hyperparameters

In its general form, the GLM is usually ill-conditioned, for there are too many parameters to estimate compared to the information brought by the data. Prior information must, hence, be incorporated in order to constrain the problem. Since the underlying physiological process of BOLD fMRI is as of yet only partially understood, we set the following soft constaints [18], [21].

P1)   The HRFs start and end at 0. This amounts to setting the first and last samples of each HRF to 0, so that only $K_i - 1$ parameters (instead of $K_i + 1$) are now unknown.

P2)   The HRFs are smooth. Quantification is achieved by setting Gaussian priors for the norm of the second derivative of the HRFs, whose variances are adjusted by hyperparameters ϵ_i's. More precisely, we assume that

$$\mathrm{p}\left(\boldsymbol{h}_i \mid H, \epsilon_i^2\right) \propto \exp\left[\frac{1}{2\epsilon_i^2} \|\partial^2 \boldsymbol{h}_i\|^2\right]$$
$$i = 1, \ldots, I.$$

Following usual practice, $\partial^2 \boldsymbol{h}_i$ can be discretized as

$$(\partial^2 \boldsymbol{h})_{i,k\Delta t} \approx \frac{h_{i,(k+1)\Delta t} - 2h_{i,k\Delta t} + h_{i,(k-1\Delta t)}}{(\Delta t)^2}$$

for $k = 1, \ldots, K_i - 1$. Taking into account that $h_{i,0} = h_{i,K} = 0$, we obtain in matrix form

$$\partial^2 \boldsymbol{h}_i = \boldsymbol{D}_{2,i} \boldsymbol{h}_i$$

where $\boldsymbol{D}_{2,i}$ is the following $(K_i - 1) \times (K_i - 1)$ matrix:

$$\boldsymbol{D}_{2,i} = \frac{1}{(\Delta t)^2} \begin{pmatrix} -2 & 1 & 0 & & & & & \\ 1 & -2 & 1 & 0 & & & & \\ 0 & 1 & -2 & 1 & 0 & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & 0 & 1 & -2 & 1 & 0 \\ & & & & 0 & 1 & -2 & 1 \\ & & & & & 0 & 1 & -2 \end{pmatrix}.$$

---

[1]If u is chi-square distributed with d degrees of freedom, then dr^2/u is scaled inverse-chi-square distributed with d degrees of freedom and scale parameter r^2. Equivalently, a scaled inverse-chi-square distribution Inv-$\chi^2(d, r^2)$ is a special case of the inverse Gamma distribution Inv-$\Gamma(\alpha, \beta)$, with $\alpha = d/2$ and $\beta = dr^2/2$. See the Appendix for the exact expression of the corresponding pdf.

Hence

$$\|\partial^2 \boldsymbol{h}_i\|^2 = (\boldsymbol{D}_{2,i}\boldsymbol{h}_i)^t(\boldsymbol{D}_{2,i}\boldsymbol{h}_i)$$
$$= \boldsymbol{h}_i^t \left(\boldsymbol{D}_{2,i}^t\boldsymbol{D}_{2,i}\right)\boldsymbol{h}_i.$$

Finally, we obtain for $\boldsymbol{h}_i$:

$$\mathrm{p}\left(\boldsymbol{h}_i \mid H, \epsilon_i^2\right) = \mathcal{N}\left(\boldsymbol{0}, \epsilon_i^2\boldsymbol{R}_i^{-1}; \boldsymbol{h}_i\right) \quad i = 1, \ldots, I \quad (3)$$

where $\boldsymbol{R}_i = \boldsymbol{D}_{2,i}^t\boldsymbol{D}_{2,i}$ is the following $(K_i - 1)$-by-$(K_i - 1)$ symmetrical positive definite matrix

$$\boldsymbol{R}_i = \frac{1}{(\Delta t)^4}$$
$$\times \begin{pmatrix} 5 & -4 & 1 & 0 & & \cdots & & & 0 \\ -4 & 6 & -4 & 1 & 0 & & & & \\ 1 & -4 & 6 & -4 & 1 & 0 & & & \\ 0 & 1 & -4 & 6 & -4 & 1 & 0 & & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ \vdots & & 0 & 1 & -4 & 6 & -4 & 1 & 0 \\ & & & 0 & 1 & -4 & 6 & -4 & 1 \\ & & & & 0 & 1 & -4 & 6 & -4 \\ 0 & & \cdots & & & 0 & 1 & -4 & 5 \end{pmatrix}.$$

P3) No prior dependence is assumed between HRFs, so that

$$\mathrm{p}\left((\boldsymbol{h}_i), \left(\epsilon_i^2\right) \mid H\right) = \prod_{i=1}^{I} \mathrm{p}\left(\boldsymbol{h}_i \mid H, \epsilon_i^2\right) \cdot \mathrm{p}\left(\epsilon_i^2 \mid H\right).$$

For convenience reasons, the priors for ϵ_i^2's are set as conjugate priors, i.e., these parameters are assumed to be *a priori* independent identically distributed with common pdf a scaled inverse-χ^2 with n_ϵ degrees of freedom and scale parameter r_ϵ^2 given in the model (n_ϵ set to a small value to obtain a "hardly" informative prior). This setting is further precised later and also analyzed in the discussion.

D. Drifts and Noise Variances

Unlike the HRFs, noise variances and drift parameters may vary across sessions. Each σ_s^2 is assumed to follow a scaled inverse-χ^2 distribution with $n_{\sigma,s}$ degrees of freedom and scale parameter $r_{\sigma,s}^2$. Each $\boldsymbol{\lambda}_s$ is assumed to be Gaussian distributed with mean $\boldsymbol{m}_{\lambda,s}$ and covariance matrix $\boldsymbol{V}_{\lambda,s}$.

E. Joint Posterior Distribution

Considering the model so constructed and assuming no further prior dependence between parameters, formal application of the chain rule yields

$$\mathrm{p}\left((\boldsymbol{y}_s), \left(\epsilon_i^2\right), (\boldsymbol{h}_i), \left(\sigma_s^2\right), (\boldsymbol{\lambda}_s) \mid H\right)$$
$$= \prod_{s=1}^{S} \mathrm{p}\left(\boldsymbol{y}_s \mid H, (\boldsymbol{h}_i), \sigma_s^2, \boldsymbol{\lambda}_s\right)$$
$$\times \mathrm{p}(\boldsymbol{\lambda}_s \mid H) \cdot \mathrm{p}\left(\sigma_s^2 \mid H\right)$$
$$\cdot \prod_{i=1}^{I} \mathrm{p}\left(\boldsymbol{h}_i \mid H, \epsilon_i^2\right) \cdot \mathrm{p}\left(\epsilon_i^2 \mid H\right). \quad (4)$$

Given data $(\boldsymbol{y}_s)$, our knowledge relative to the model parameters can easily be updated using the conditioning formula

$$\mathrm{p}\left(\left(\epsilon_i^2\right), (\boldsymbol{h}_i), \left(\sigma_s^2\right), (\boldsymbol{\lambda}_s) \mid H, (\boldsymbol{y}_s)\right) =$$
$$\frac{\mathrm{p}\left((\boldsymbol{y}_s), \left(\epsilon_i^2\right), (\boldsymbol{h}_i), \left(\sigma_s^2\right), (\boldsymbol{\lambda}_s) \mid H\right)}{\mathrm{p}((\boldsymbol{y}_s) \mid H)}.$$

In words, the joint posterior probability distribution is proportional to the joint probability of (4). Replacing all distributions by their functional forms, this joint posterior pdf could be calculated in closed form, as is indeed done in most works applying Bayesian analysis. Since direct sampling from the joint posterior pdf is impossible, we must resort to MCMC, e.g., Gibbs sampling where the conditional pdfs should be derived. In this perspective, we propose to avoid calculating the joint posterior pdf to directly proceed to inference. In order to do so, we beforehand embed our model in a framework that allows for convenient representation, handling, and numerical inference: Bayesian graphical models.

III. Graphical Modeling

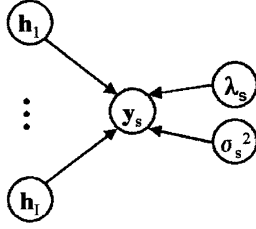#### A. Directed Acyclic Graphs and Bayesian Networks

A graph G is a mathematical object that relates a set of vertices, or nodes, V, to a set of edges, E, consisting of pairs of elements taken from V. There is a directed edge or arrow between vertices $\boldsymbol{z}_n$ and $\boldsymbol{z}_m$ in V if the set E contains the ordered pair $(\boldsymbol{z}_n, \boldsymbol{z}_m)$; vertex $\boldsymbol{z}_n$ is a parent of vertex $\boldsymbol{z}_m$, and vertex $\boldsymbol{z}_m$ is a child of vertex $\boldsymbol{z}_n$. A directed graph is a graph whose edges are all directed. A path is a sequence of distinct vertices $\boldsymbol{z}_{n_1}, \ldots, \boldsymbol{z}_{n_m}$ for which $(\boldsymbol{z}_{n_l}, \boldsymbol{z}_{n_{l+1}})$ is in E for each $l = 1, \ldots, m - 1$. The path is a cycle if the end points are allowed to be the same, $\boldsymbol{z}_{n_1} = \boldsymbol{z}_{n_m}$. An oriented graph with no cycle is called a directed acyclic graph (DAG).

A distribution p over $\boldsymbol{z}$ is compatible with a DAG G if it satisfies all independence relationships entailed by G. (G, p) is then called a Bayesian network. For more details, the reader is referred to [27]. The major feature of Bayesian networks is that $\mathrm{p}(\boldsymbol{z})$ must factorize according to the so-called factorization property

$$\mathrm{p}(\boldsymbol{z}) = \prod_{n=1}^{N} \mathrm{p}(\boldsymbol{z}_n \mid \mathrm{pa}(\boldsymbol{z}_n)) \quad (5)$$

where $\mathrm{pa}(n)$ is the set of parents of vertex $\boldsymbol{z}_n$. This is nothing but a multidimensional generalization of the Markov chain rule. Defining a Bayesian network, hence, amounts to 1) defining relevant variables (i.e., nodes) $\boldsymbol{z}_n$, 2) defining structural relationships (i.e., edges) $\boldsymbol{z}_n \to \boldsymbol{z}_m$, and 3) defining functional relationships $\mathrm{p}(\boldsymbol{z}_n \mid \mathrm{pa}(n))$. Pearl [28] showed a property that proves to be very efficient for numerical sampling, namely that nothing more is required to calculate the conditional probability of any node: the probability distribution of any variable $\boldsymbol{z}_n$ in the network, conditioned on the state of all other variables, is given by the product

$$\mathrm{p}(\boldsymbol{z}_n \mid \mathrm{r.v.}) \propto \mathrm{p}(\boldsymbol{z}_n \mid \mathrm{pa}(\boldsymbol{z}_n)) \cdot \prod_{\boldsymbol{z}_m \in \mathrm{ch}(\boldsymbol{z}_n)} \mathrm{p}(\boldsymbol{z}_m \mid \mathrm{pa}(\boldsymbol{z}_m))$$

$$(6)$$

Fig. 1.   Structural dependence of $\boldsymbol{y}_s$.

Fig. 2.   Structural dependence of $\boldsymbol{h}_i$.

where r.v. stands for "remaining variables" and $\mathrm{ch}(\boldsymbol{z}_n)$ for the children nodes of $\boldsymbol{z}_n$. Note that $\boldsymbol{z}_m$ can have several parents and, hence, $\mathrm{pa}(\boldsymbol{z}_m)$ may not be restricted to $\boldsymbol{z}_n$. This formula states that the conditional probabilities can be derived from local quantities that are part of the model specification.

B. Constructing the GLM Graphical Model

The GLM can easily be expressed in terms of a Bayesian network. This translation requires two major steps: representing the structural relationships, and then the functional relationships.

Equation (1) states that each $\boldsymbol{y}_s$ depends on the corresponding λ_s and σ_s^2, as well as all $\boldsymbol{h}_i$'s. This equation can, hence, be represented by the graph depicted in Fig. 1.

As to the model relative to the prior information on the HRFs, it expresses that each $\boldsymbol{h}_i$ depends exclusively on ϵ_i^2. Fig. 2 is, hence, a good model for it.

Gathering all parts leads to the graph proposed in Fig. 3. Irrespective of the functional relationships between nodes, application of (5) to the graphical model states that the joint pdf for all variables decomposes as

$$
\begin{aligned}
&\mathrm{p}\left((\boldsymbol{y}_s),\left(\epsilon_i^2\right),(\boldsymbol{h}_i),\left(\sigma_s^2\right),(\lambda_s)\mid H\right)\\
&=\prod_{s=1}^{S}\mathrm{p}(\boldsymbol{y}_s\mid H,\mathrm{pa}(\boldsymbol{y}_s))\\
&\quad\times\mathrm{p}(\lambda_s\mid H,\mathrm{pa}(\lambda_s))\cdot\mathrm{p}\left(\sigma_s^2\mid H,\mathrm{pa}\left(\sigma_s^2\right)\right)\\
&\quad\times\prod_{i=1}^{I}\mathrm{p}\left(\boldsymbol{h}_i\mid H,\mathrm{pa}(\boldsymbol{h}_i)\right)\cdot\mathrm{p}\left(\epsilon_i^2\mid H,\mathrm{pa}\left(\epsilon_i^2\right)\right)
\end{aligned}
$$

which, once developed, is exactly (4). Our graphical model, thus, unambiguously embeds all structural relationships of the GLM. However, complicated (H) may be, it is still much simpler to conceptualize it in graph form than as it was presented before. Whereas determination of structural relationships between two given variables in model (H) remains a tough problem to tackle, the corresponding DAG clearly and unambiguously represents all possible independence relationships, that can be read off the graph using its Markov properties.

Among others, a direct consequence of this modeling is that it is now possible to apply (6) to the GLM graph model, expressing the conditional pdfs of all vari-
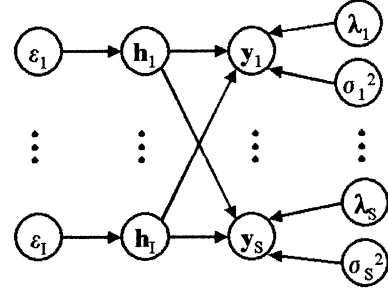

Fig. 3.   DAG corresponding to the GLM.

ables, $\mathrm{p}(\epsilon_i^2\mid H,\mathrm{r.v.}),\mathrm{p}(\boldsymbol{h}_i\mid H,\mathrm{r.v.}),\mathrm{p}(\sigma_s^2\mid H,\mathrm{r.v.})$,   and $\mathrm{p}(\lambda_s\mid H,\mathrm{r.v.})$ as a function of the distributions already defined

$$
\mathrm{p}\left(\epsilon_i^2\mid H,\mathrm{r.v.}\right)\propto\mathrm{p}\left(\epsilon_i^2\mid H\right)\cdot\mathrm{p}\left(\boldsymbol{h}_i\mid H,\epsilon_i^2\right) \tag{7a}
$$

$$
\begin{aligned}
\mathrm{p}(\boldsymbol{h}_i\mid H,\mathrm{r.v.})&\propto\mathrm{p}\left(\boldsymbol{h}_i\mid H,\epsilon_i^2\right)\\
&\times\prod_{s=1}^{S}\mathrm{p}\left(\boldsymbol{y}_s\mid H,(\boldsymbol{h}_i),\sigma_s^2,\lambda_s\right)
\end{aligned} \tag{7b}
$$

$$
\mathrm{p}\left(\sigma_s^2\mid H,\mathrm{r.v.}\right)\propto\mathrm{p}\left(\sigma_s^2\mid H\right)\cdot\mathrm{p}\left(\boldsymbol{y}_s\mid H,(\boldsymbol{h}_i),\sigma_s^2,\lambda_s\right) \tag{7c}
$$

$$
\mathrm{p}(\lambda_s\mid H,\mathrm{r.v.})\propto\mathrm{p}(\lambda_s\mid H)\cdot\mathrm{p}\left(\boldsymbol{y}_s\mid H,(\boldsymbol{h}_i),\sigma_s^2,\lambda_s\right). \tag{7d}
$$

Note that the functional relationships have not been defined yet—all the properties abovementioned are entailed by the sole structural relationships. As a matter of fact, the graphical representation is much more general than the GLM. The only constraints set by the graph is that the functional relationships be of the form $\mathrm{p}(\epsilon_i^2),\mathrm{p}(\boldsymbol{h}_i\mid\epsilon_i^2),\mathrm{p}(\sigma_s^2),\mathrm{p}(\lambda_s)$, and $\mathrm{p}(\boldsymbol{y}_s\mid H,(\boldsymbol{h}_i),\sigma_s^2,\lambda_s)$. But these conditional distributions can, in turn, be chosen at will. On the other hand, the graphical model can be made more specific for our purpose, so that it exactly fits the GLM. Identifying all functional relationships of the network to their counterparts for model (H) then makes the DAG a perfect representation of the GLM. $\mathrm{p}(\epsilon_i^2)$ and $\mathrm{p}(\boldsymbol{h}_i\mid\epsilon_i^2)$ can be chosen to a scaled inverse-χ^2 and a Gaussian distribution, respectively, as detailed in Section II-C. $\mathrm{p}(\sigma_s^2)$ and $\mathrm{p}(\lambda_s)$ can be set as in Section II-D. Finally, $\mathrm{p}(\boldsymbol{y}_s\mid H,(\boldsymbol{h}_i),\sigma_s^2,\lambda_s)$ can be set as in (2).

C. Numerical Inference

To obtain a numerical approximation of the joint posterior pdf, we apply Gibbs sampling. This consists of starting with a seed vector and sequentially modifying one vector component at a time by sampling according to the conditional pdf of that component given the remaining variables. Samples are composed of the set of all vectors whose components have been updated an equal amount of times.

A key issue with Gibbs sampling is to partition the vector of all parameters into components whose conditional sampling can easily be performed. Another one is derivation of the conditional pdfs corresponding to the chosen clustering. In our case, both questions are answered at once, thanks to the previous step of graph modeling. As a matter of fact, it first allows us to decompose the parameter vector onto its $2I+2S$ canonical components: I ϵ_i^2's and $\boldsymbol{h}_i$'s, S σ_s^2's, and λ_s's. All $\boldsymbol{y}_s$'s being given, no sampling needs to be done on these variables. The updating steps are performed on these

variables; we, therefore, need access to the following conditional pdfs: $\mathrm{p}(\epsilon_i^2 \,|\, H, \mathrm{r.v.})$, $\mathrm{p}(\boldsymbol{h}_i \,|\, H, \mathrm{r.v.})$, $\mathrm{p}(\sigma_s^2 \,|\, H, \mathrm{r.v.})$, and $\mathrm{p}(\boldsymbol{\lambda}_s \,|\, H, \mathrm{r.v.})$. But these are just the conditional distributions given by Pearl's theorem and structurally developed in (7). Integration of the exact functional relationships into (7) leads to (see the Appendix for a summary of the properties used)

- According to (7a), $(\epsilon_i^2 \,|\, H, \mathrm{r.v.})$ is proportional to the product of two inverse-chi-square distributions in ϵ_i^2; it is, hence, also inverse-chi-square distributed

$$\mathrm{p}\left(\epsilon_i^2 \,\big|\, H, \mathrm{r.v.}\right) = \mathrm{Inv}\text{-}\chi^2\left(\mu_i, \tau_i^2; \epsilon_i\right)$$

  with

$$\mu_i = n_\epsilon + (K_i - 1)$$
$$\tau_i^2 = \frac{n_\epsilon r_\epsilon^2 + \boldsymbol{h}_i^t R \boldsymbol{h}_i}{n_\epsilon + (K_i - 1)}.$$

- According to (7b), $(\boldsymbol{h}_i \,|\, H, \mathrm{r.v.})$ is proportional to the product of two multivariate Normal distributions in $\boldsymbol{h}_i$; it is, hence, also multivariate Normal distributed

$$\mathrm{p}(\boldsymbol{h}_i \,|\, H, \mathrm{r.v.}) = \mathcal{N}(\boldsymbol{\delta}_i, \boldsymbol{\Delta}_i; \boldsymbol{h}_i)$$

  with

$$\boldsymbol{\Delta}_i = \left(\frac{1}{\epsilon_i^2} R + \sum_s \frac{1}{\sigma_s^2} \boldsymbol{X}_{s,i}^t \boldsymbol{X}_{s,i}\right)^{-1}$$

$$\boldsymbol{\delta}_i = \boldsymbol{\Delta}_i \left(\sum_s \frac{1}{\sigma_s^2} \boldsymbol{X}_{s,i}^t \right.$$
$$\left. \cdot \left(\boldsymbol{y}_s - \sum_{j \neq i} \boldsymbol{X}_{s,j} \boldsymbol{h}_j - D_s \boldsymbol{\lambda}_s\right)\right).$$

- According to (7c), $(\sigma_s^2 \,|\, H, \mathrm{r.v.})$ is proportional to the product of two inverse-chi-square distributions in σ_s^2; it is, hence, also inverse-chi-square distributed

$$\mathrm{p}\left(\sigma_s^2 \,\big|\, H, \mathrm{r.v.}\right) = \mathrm{Inv}\text{-}\chi^2\left(\nu_s, \omega_s^2; \sigma_s^2\right)$$

  with

$$\nu_s = n_{\sigma,i} + L_s$$
$$\omega_s^2 = \frac{n_{\sigma,s} r_{\sigma,i}^2 + \|\boldsymbol{y}_s - \sum_i \boldsymbol{X}_{s,i} \boldsymbol{h}_i - D_s \boldsymbol{\lambda}_s\|^2}{n_{\sigma,i} + L_s}.$$

- According to (7d), $(\boldsymbol{\lambda}_s \,|\, H, \mathrm{r.v.})$ is proportional to the product of two multivariate Normal distributions in $\boldsymbol{\lambda}_s$; it is, hence, also multivariate Normal distributed

$$\mathrm{p}(\boldsymbol{\lambda}_s \,|\, H, \mathrm{r.v.}) = \mathcal{N}(\boldsymbol{\gamma}_s, \boldsymbol{\Gamma}_s; \boldsymbol{\lambda}_s)$$

  with

$$\boldsymbol{\Gamma}_s = \left(V_{\lambda,s}^{-1} + \frac{1}{\sigma_s^2} D_s^t D_s\right)^{-1}$$

$$\boldsymbol{\gamma}_s = \boldsymbol{\Gamma}_s \left(V_{\lambda,s}^{-1} \boldsymbol{m}_{\lambda,s} + \frac{1}{\sigma_s^2} D_s^t \left(\boldsymbol{y}_s - \sum_i \boldsymbol{X}_{s,i} \boldsymbol{h}_i\right)\right).$$

The sampling can then be performed by sequentially updating the ϵ_i^2's, the $\boldsymbol{h}_i$'s, then the σ_s^2's, and finally the $\boldsymbol{\lambda}_s$'s. Convergence monitoring is performed component-wise using parallel sampling as detailed in [29]. More precisely, we first take the logarithm of ϵ_i^2 and σ_s^2, so that all variables are spanned from $-\infty$ to $+\infty$. For each estimand ϕ, $\phi = \log \epsilon_i^2, \log \sigma_s^2, h_{i,k}$, and $\lambda_{s,m}$, we draw B parallel sequences of length C (we typically

took $B = 10$ and $C = 50$), each sample being denoted $\phi^{[bc]}$, with $b = 1, \ldots, B$ and $c = 1, \ldots, C$. We then compute the between-sequence variance BV, and the within-variance sequence WV as follows:

$$\mathrm{BV} = \frac{C}{B-1} \sum_{b=1}^{B} (\bar{\phi}^{[b\cdot]} - \bar{\phi}^{[\cdot\cdot]})^2$$

with

$$\bar{\phi}^{[b\cdot]} = \frac{1}{C} \sum_{c=1}^{C} \phi^{[bc]} \quad \text{and} \quad \bar{\phi}^{[\cdot\cdot]} = \frac{1}{B} \sum_{b=1}^{B} \bar{\phi}^{[b\cdot]}$$

and

$$WV = \frac{1}{B} \sum_{b=1}^{B} (s^2)^{[b]}$$

where

$$(s^2)^{[b]} = \frac{1}{C-1} \sum_{c=1}^{C} \left(\phi^{[bc]} - \bar{\phi}^{[b\cdot]}\right)^2.$$

We then calculate

$$\sqrt{\hat{R}} = \sqrt{1 + \frac{1}{C}\left(\frac{\mathrm{BV}}{WV} - 1\right)}$$

for each scalar estimand. These quantities are supposed to decline to 1 as the sampling converges. We stop the algorithm when all $\sqrt{\hat{R}}$ are close enough to 1, e.g., smaller than 1.1, and remove α percent of each chain to account for a burn-in period.

We are admittedly mostly interested in the HRFs, but knowledge of the values taken by the other parameters are relevant as well for our analysis and a better understanding of brain processing. Gibbs sampling gives us access to estimates for all parameters or any quantity of interest related to them. For instance, in this paper, parameter estimators are given as

$$\text{posterior mean} \pm \text{posterior standard deviation}.$$

Once Gibbs sampling has converged, these quantities are approximated by their sample counterparts.

IV. SIMULATIONS

We simulated data with two HRFs $(I = 2)$, as depicted in Fig. 4. To obtain the ϵ_i's corresponding to these HRFs, we calculated them as follows. If we knew that $\boldsymbol{h}_i = \boldsymbol{h}_i^0$, then we could infer ϵ_i^2 using Bayes' theorem

$$\mathrm{p}\left(\epsilon_i^2 \,|\, H', \boldsymbol{h}_i = \boldsymbol{h}_i^0\right) \propto \mathrm{p}\left(\epsilon_i^2 \,|\, H'\right) \cdot \mathrm{p}\left(\boldsymbol{h}_i = \boldsymbol{h}_i^0 \,|\, H', \epsilon_i^2\right)$$

where $\mathrm{p}(\boldsymbol{h}_i = \boldsymbol{h}_i^0 \,|\, H', \epsilon_i^2)$ would be given by (3), and $\mathrm{p}(\epsilon_i^2 \,|\, H')$ could be taken as a uniform prior, assuming no particular prior knowledge. $(\epsilon_i^2 \,|\, H', \boldsymbol{h}_i = \boldsymbol{h}_i^0)$ would then be Inv-χ^2 distributed and

$$\mathrm{E}\left[\epsilon_i^2\right] = \frac{1}{K-1} \cdot \boldsymbol{h}_i^{0t} R \boldsymbol{h}_i^0.$$

This last relation can, in turn, be taken as estimate for ϵ_i^2, leading to

$$\epsilon_1^2 = 1.21$$
$$\epsilon_2^2 = 0.30.$$

For the simulation, we also took two sessions of $N = 100$ time samples. Δt and the sampling interval were both set to
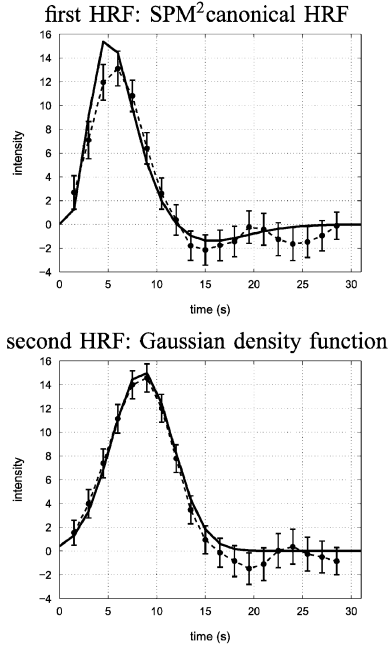
first HRF: SPM^{2}canonical HRF

second HRF: Gaussian density function

Fig. 4. Simulations. Estimated (*dashed line*) and simulated (*solid line*) HRF (Available: www.fil.ion.ucl.ac.uk/spm/spm99.html).
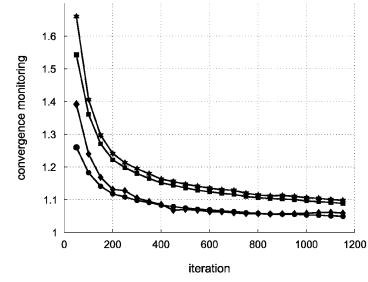

Fig. 5. Convergence monitoring. Every 50 steps is represented $\max[\sqrt{\hat{R}}]$ taken among all ϵ_i^2 (diamonds), σ_s^2 (circles), $h_{i,k}$ (squares), and $\lambda_{s,m}$ (stars).

1.5 s. Quadratic drifts ($p_1(t) = 846 + 0.2 \cdot t + 0.001 \cdot t^2$ and $p_2(t) = 950 + 0.15 \cdot t + 0.0011 \cdot t^2$) and Gaussian white noises ($\sigma_1^2 = 50, \sigma_2^2 = 100$) were also added. Note that the noise standard deviations are about of the same amplitude as the HRF. For the analysis, we set both orders to $K = 20$, and took a quadratic drift in consideration ($M = 3$) with each $\boldsymbol{m}_{\lambda,s}$ set to $(\mathrm{E}[\boldsymbol{y}_s], 0, 0)^t$, and $\boldsymbol{V}_{\lambda,s}$ to a diagonal matrix such, that $\mathrm{diag}(\boldsymbol{V}_{\lambda,s}) = (10\,000^2 \ 1\,000^2 \ 1\,000^2)$. All $n_{\sigma,s}$ and n_ϵ were set to 1 to implement vague priors for σ_s^2's and ϵ_i^2's. As to the remaining hyperparameters, the goal was to set them, so that they provide a magnitude order for the corresponding variables as follows:

- each $r_{\sigma,s}^2$ was set to $\mathrm{Var}[\boldsymbol{y}_s]$;
- r_ϵ^2 was set to

$$\frac{\mathrm{E}_s[\mathrm{Var}[\boldsymbol{y}_s]]}{\max[\boldsymbol{h}_{\mathrm{SPM}}]} \cdot \left[\frac{1}{K-1} \cdot \boldsymbol{h}_{\mathrm{SPM}}^t \boldsymbol{R} \boldsymbol{h}_{\mathrm{SPM}} \right]$$

where $\boldsymbol{h}_{\mathrm{SPM}}$ is SPM canonical HRF and $\mathrm{E}_s[\mathrm{Var}[\boldsymbol{y}_s]]$ is a scaling factor.

Gibbs sampling took about 2 250 updates for each of the 10 parallel chains to converge, and we kept the last 10% of each chain. We obtained the following estimates:

$$\epsilon_1^2 \approx 1.29 \pm 0.80$$
$$\epsilon_2^2 \approx 0.936 \pm 0.524$$
$$\sigma_1^2 \approx 38.6 \pm 7.8$$
$$\sigma_2^2 \approx 109 \pm 20$$
$$\lambda_{1,0} \approx 858 \pm 4$$
$$\lambda_{2,0} \approx 956 \pm 4.$$

Estimates for ϵ_i^2, σ_s^2, and $\boldsymbol{\lambda}_s$ were accurate. As shown in Fig. 4, HRF estimates were also very accurate for the noise level considered.

V. REAL DATA

Eleven healthy subjects (age 18–40) were scanned while performing a motor sequence learning task. Using a joystick, they were asked to reach a target projected on a screen for 3 s, following an elliptic curve as precisely and rapidly as possible. They had to complete 64 trials of sequence (SEQ) mode (the targets appeared in a predefined order, unknown to the subject, to form a 8-item-long sequence) and 16 trials of random (RAN) mode (the targets appeared pseudorandomly). The time interval beween two consecutive trials, or inter-stimulus interval, was randomly selected to uniformly lie between 3 and 4 s. Functional T_2^*-weighted acquisitions were performed on a 3 T Bruker MEDSPEC 30/80 MR system (TR: 3 486 ms, TE: 35 ms, flip angle: 90°, matrix $64 \times 64 \times 42$, voxel size $3 \times 3 \times 3$ mm).

The data imply to work with $I = 2$ HRFs ($i = 1$ and $i = 2$ corresponding to SEQ and RAN, respectively) and $S = 5$ sessions. For the analysis, we first adjusted the stimulus on a grid of interval $\Delta t = \mathrm{TR}/5$. Both HRFs were assumed to have a common order $K = 5 \times 5$, for a total duration of 5 TRs. The prior hyperparameters were set as in Section IV. To illustrate the method, we selected two voxels, v_1 and v_2. v_1 was located in the right cerebellum and v_2 in the right inferotemporal lobe. Our goal was to estimate both HRFs corresponding to conditions SEQ and RAN, respectively.

Based on our monitoring system, convergence took around 1 300 iterations to occur. Fig. 5 shows a typical convergence curve. As showed in Fig. 6, the method was able to extract different HRF behaviors for different conditions, despite a very low signal-to-noise ratio. The high noise level was reflected in the large estimate error bars but did not prevent discrimination between conditions. The estimated values for the other parameters are given in Table I.

VI. DISCUSSION

Our approach made it possible to associate the well-known GLM for HRF estimation in fMRI data analysis to a directed acyclic graph. This had the first advantage of making clear all modeling hypotheses. Moreover, in a Bayesian framework, the complex, yet central, step of calculating the joint posterior pdf was avoided. Instead, the graph provided us with a very convenient tool to first break down the set of all variables into coherent subsets, namely its nodes. Using the Markov properties, it was direct to derive all conditional pdfs that were required for Gibbs sampling as products of conditional pdfs that have been
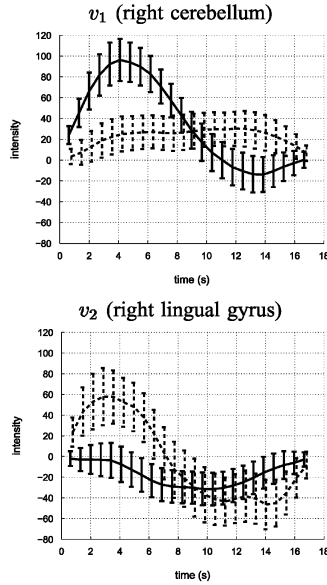
Fig. 6. Real data. HRFs corresponding to the SEQ (*solid line*) and RAN (*dashed line*) stimuli. Each sample has been slightly shifted to the left (SEQ) or to the right (RAN) for a better graphical rendering.

TABLE I
Estimates of the Smoothing Parameters ϵ_i, the Noise Variances σ_s^2, and the Baselines $\lambda_{s,1}$

		ϵ_i
v_1	$i = 1$	$3.83e4 \pm 2.90e4$
	$i = 2$	$2.76e4 \pm 5.30e4$
v_2	$i = 1$	$2.43e4 \pm 1.90e4$
	$i = 2$	$1.11e5 \pm 1.33e5$

		σ_s^2
v_1	$s = 1$	$1.75e4 \pm 2.67e3$
	$s = 2$	$5.02e3 \pm 8.24e2$
	$s = 3$	$6.82e3 \pm 1.10e3$
	$s = 4$	$5.38e3 \pm 8.55e2$
	$s = 5$	$5.37e3 \pm 8.36e2$
v_2	$s = 1$	$2.53e4 \pm 3.91e3$
	$s = 2$	$4.53e3 \pm 7.08e2$
	$s = 3$	$1.05e4 \pm 1.73e3$
	$s = 4$	$9.11e3 \pm 1.46e3$
	$s = 5$	$5.08e3 \pm 8.26e2$

		$\lambda_{s,1}$
v_1	$s = 1$	$3.79e3 \pm 21.3$
	$s = 2$	$3.86e3 \pm 18.3$
	$s = 3$	$3.50e3 \pm 18.4$
	$s = 4$	$3.65e3 \pm 18.1$
	$s = 5$	$3.37e3 \pm 17.8$
v_2	$s = 1$	$3.94e4 \pm 25.1$
	$s = 2$	$3.48e3 \pm 20.2$
	$s = 3$	$3.74e3 \pm 22.0$
	$s = 4$	$3.52e3 \pm 21.4$
	$s = 5$	$3.26e3 \pm 20.4$

specified with the modeling. Fully probabilistic numerical inference was then straightforward at a reasonable time cost. Furthermore, solving the same problem with several variables set to certain values was an easy matter, since all that must be done is removing these variables from the sampling scheme.

Because variable partitioning is implied by the graph structure, our application of Gibbs sampling differs from classical applications. For instance, we sequentially sampled each HRF, resp. drift vector, resp. noise variance, whereas a conventional procedure would simultaneously sample all HRFs, then all drift parameters, and so on. What influence this difference makes on the convergence speed of the Markov chain is a matter that needs

to be further investigated. Another point would be to compare our method to a more conventional procedure where part, or all, of the nuisance parameters (e.g., drift parameters) are integrated out of the joint posterior pdf, and inference is done on the posterior marginal pdf of the HRFs. In the "marginal" scheme, each sample is performed on a lower dimension, and the Markov Chain is also of lower dimension. For these reasons, it is expected that convergence will be faster. This scheme is, however, very sensitive to model changes and makes inference on other parameters tedious.

Influence of the prior parameters is of importance to check sensitivity of the estimates. $\boldsymbol{m}_\lambda$ and $\boldsymbol{V}_\lambda$ were found to have very little weight on the inference, n_σ and r_σ some more. Selection of n_ϵ and r_ϵ did not much matter for drift and noise parameters, but had a dramatic influence on HRF estimation. The ad hockery proposed to manually set these hyperparameters seems to be useful, but its influence on the sampling scheme is still unclear. Indeed, except for this, we did not even have a prior idea of the magnitude order of ϵ_i^2, whereas scaled inverse-chi-square pdfs are relatively localized around their mode. To remedy this flaw, we suggest that setting priors for $\log(\epsilon_i^2)$ would prove more adapted to the state of ignorance that we are in relative to these parameters than conjugate priors. A general procedure is proposed in [30] to sample from pdfs that have the structure implied by (6) using rejection sampling. Another improvement would be to modify the model, so that ϵ_i becomes independent of the HRF intensity. This could be achieved by consideration of a normalized HRF and a scaling factor. Choice of noninformative, i.e., improper,[2] priors might also help to circumvent that problem. The reason why we did not consider this option here is that it is safe to use improper priors as long as the posterior pdf can be proved to be proper. Since one advantage of our method is to avoid calculation of the posterior pdf, this step cannot be performed in our setting without making it lose some of its appeal. Nonetheless, we are optimistic and believe that further investigation could justify use of improper priors in a restricted fashion.

In the framework of DAG modeling, the local properties of relationships renders the model very simple to structurally or functionally modify at a local level, either because it does not correctly explain the phenomenon under interest, or because a more complex model is sought. As a matter of fact, the proposed model for HRF estimation can already be seen as an improvement of the graphical model associated to the basic one-HRF, one-session linear model. Fig. 7 illustrates how this model can successively be expanded by introduction of smoothing priors, several sessions, and several stimulus types. These models can be dealt with as efficiently as the one detailed in this paper. Their joint posterior distribution would be given by application of (5), whereas all conditional distributions required by Gibbs sampling could easily be obtained through (6).

Besides, consideration of local spatial information, as in [23], [24], [31] could be achieved by gathering all voxel graphical models that were here assumed to be independent from each other and adding relationships between neighboring $\boldsymbol{h}_{v,i}$'s. Another point would be to relax the assumption that all HRFs are

[2]A prior pdf p is improper if its integral is not finite, i.e.,
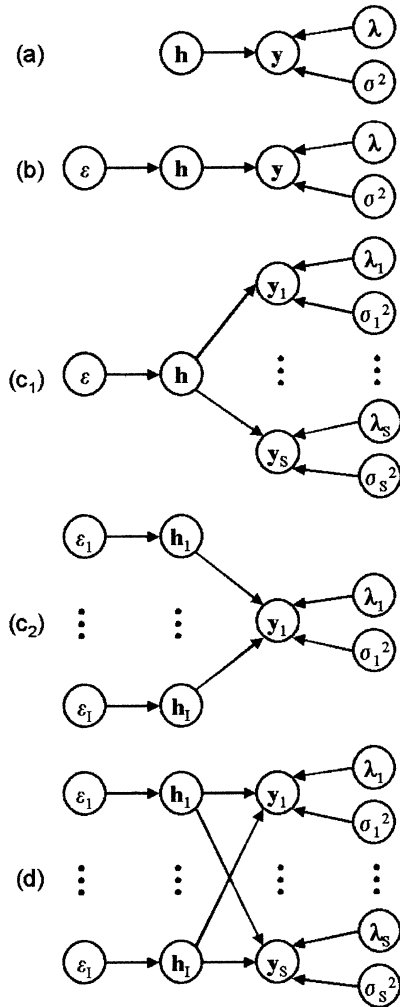
$$\int p(x)\, dx = \infty.$$

Fig. 7. Successives complexification of the HRF model. (a) basic model; (b) model with hyperparameter for smoothness regularization; (c_1) model accounting for one stimulus type and several sessions; (c_2) model accounting for several stimulus types and one sessions; (d) GLM considered in this paper.

constant across sessions and rather assume that they share the same shape accross sessions, with an amplitude that can vary, as discussed in [22]. As more and more information is incorporated into the model, the corresponding graph will become more and more complex. However, tools have been developed to deal with such graphs. Parallel processing of Gibbs sampling can be implemented. To avoid the problem of simultaneous updating of neighboring variables, one has to apply the so-called "edge reversal" control policy, as detailed in [28]. For huge graphs, [32] proposed an efficient variant of Gibbs sampling.

We finally believe that this novel approach has a much broader application range than just fMRI data analysis. Indeed, we are confident in the fact that any Bayesian model can be embedded in a graphical framework. This would allow to concentrate on the modeling, since efficient and automated inference would directly derive from the model.

VII. Conclusion

In this paper, we proposed a novel Bayesian inference framework for HRF estimation in fMRI data analysis, based on trans-

lating the existing Bayesian model into a Bayesian network to combine the features of graphical modeling and Bayesian analysis. This approach makes extensive use of Bayesian networks to 1) represent the model in a compact, yet efficient way, and 2) lead probabilistic inference through Gibbs sampling. This technique takes advantage of Markov properties of DAGs. Models can easily be designed, and both structural (i.e., of independence) and functional relationships are clearly presented. Moreover, using Gibbs sampling on the DAG, fully probabilistic numerical inference is straightforward. Ongoing research includes integration of more diffuse prior pdfs when necessary, as well as spatial constraints for the HRFs.

Appendix

The results detailed here originate from [29].

A. Multivariate Normal Distribution

If x is a d-dimensional multivariate Normal distributed variable with mean m and covariance matrix V, then

$$p(x) = (2\pi)^{d/2}|V|^{-1/2}\exp\left[-\frac{1}{2}(x-m)^t V^{-1}(x-m)\right].$$

If x has a probability distribution defined by

$$p(x) \propto f_1(x) \cdot f_2(x)$$

where f_i is multivariate Gaussian with mean m_i and covariance matrix V_i, then x is multivarite Gaussian with mean

$$m = \left(V_1^{-1} + V_2^{-1}\right)^{-1}\left(V_1^{-1}m_1 + V_2^{-1}m_2\right)$$

and covariance matrix

$$V = \left(V_1^{-1} + V_2^{-1}\right)^{-1}.$$

B. Inverse-Chi-Square Distribution

If y is inverse-chi-square distributed with n degrees of freedom and scale s, then

$$p(y) = \frac{(n/2)^{n/2}}{\Gamma(n/2)} s^n y^{-(n/2+1)} e^{-ns^2/(2y)}.$$

If y has a probability distribution defined by

$$p(y) \propto f_1(y) \cdot f_2(y)$$

where f_i is inverse-chi-square distributed with n_i degrees of freedom scale s_i, then y is inverse-chi-square with

$$n = n_1 + n_2$$

degrees of freedom and scale s so, that

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}.$$

References

[1] W. Chen and S. Ogawa, "Principles of BOLD functional MRI," in *Functional MRI*, C. Moonen and P. Bandettini, Eds, Berlin: Springer, 1999, pp. 103–113.

[2] T.-Q. Li, T. N. Haefelin, B. Chan, A. Kastrup, T. Jonsson, G. H. Glover, and M. E. Moseley, "Assessment of hemodynamic response during focal neural activity in human using bolus tracking, arterial spin labeling and BOLD techniques," *Neuroimage*, vol. 12, pp. 442–451, 2000.

[3] A. Aubert and R. Costalat, "A model of the coupling between brain electrical activity, metabolism, and hemodynamics: Application to the interpretation of functional neuroimaging," *Neuroimage*, vol. 17, pp. 1162–1181, 2002.

[4] K. Uğurbil, L. Toth, and D.-S. Kim, "How accurate is magnetic resonance imaging of brain function?," *Trends Neurosci.*, vol. 26, pp. 108–114, 2003.

[5] K. J. Friston, P. Jezzard, and R. Turner, "Analysis of functional MRI time-series," *Hum. Brain Mapp.*, vol. 1, pp. 153–171, 1994.

[6] A. M. Dale and R. L. Buckner, "Selective averaging of rapidly presented individual trials using fMRI," *Hum. Brain Mapp.*, vol. 5, pp. 329–340, 1997.

[7] R. L. Buckner, "Event-related fMRI and the hemodynamic response," *Hum. Brain Mapp.*, vol. 6, pp. 373–377, 1998.

[8] C. R. Genovese, "A Bayesian time-course model for functional magnetic resonance imaging data," *J. Amer. Statist. Assoc.*, vol. 95, pp. 691–719, 2000.

[9] C. Gössl, L. Fahrmeir, and D. P. Auer, "Bayesian modeling of the hemodynamic response function in BOLD fMRI," *Neuroimage*, vol. 14, pp. 140–148, 2001.

[10] R. L. Buckner, W. Koutstaal, D. L. Schacter, A. D. Wagner, and B. R. Rosen, "Functional-anatomic study of episodic retrieval using fMRI (I)," *Neuroimage*, vol. 7, pp. 151–162, 1998.

[11] R. L. Buckner, W. Koutstaal, D. L. Schacter, A. M. Dale, M. Rotte, and B. R. Rosen, "Functional-anatomic study of episodic retrieval using fMRI (II)," *Neuroimage*, vol. 7, pp. 163–175, 1998.

[12] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of human, BOLD hemodynamic responses," *Neuroimage*, vol. 8, pp. 360–369, 1998.

[13] M. D'Esposito, E. Zarahn, G. K. Aguirre, and B. Rypma, "The effect of normal aging on the coupling of neural activity to the BOLD hemodynamic response," *Neuroimage*, vol. 10, pp. 6–14, 1999.

[14] M. D'Esposito, L. Y. Deouel, and A. Gazzaley, "Alterations in the BOLD fMRI signal with ageing and disease: A challenge for neuroimaging," *Nat. Neurosci.*, vol. 4, pp. 863–872, 2003.

[15] J. Kershaw, S. Abe, K. Kashikura, X. Zhang, and I. Kanno, "A Bayesian approach to estimating the haemodynamic response function in event-related fMRI," *Neuroimage*, vol. 11, p. S474, 2000.

[16] M. A. Burock and A. M. Dale, "Estimation and detection of event-related fMRI signals with temporally correlated noise: A statistically efficient and unbiased approach," *Hum. Brain Mapp.*, vol. 11, pp. 249–260, 2000.

[17] C. Goutte, F. Å. Nielsen, and L. K. Hansen, "Modeling the haemodynamic response in fMRI using smooth FIR filters," *IEEE Trans. Med. Imag.*, vol. 19, pp. 1188–1201, Dec. 2000.

[18] G. Marrelec, H. Benali, P. Ciuciu, and J.-B. Poline, "Bayesian estimation of the hemodynamic response function in functional MRI," in *Proc. MaxEnt Workshop*, R. Fry, Ed., 2001, pp. 229–247.

[19] G. Marrelec, H. Benali, P. Ciuciu, M. Pélégrini-Issac, and J.-B. Poline, "Robust Bayesian estimation of the hemodynamic response function in event-related BOLD fMRI using basic physiological information," *Hum. Brain Mapp.*, vol. 19, pp. 1–17, 2003.

[20] P. Ciuciu, G. Marrelec, J. Idier, H. Benali, and J.-B. Poline. A general tool to estimate the hemodynamic response function in fMRI data. presented at OHBM Proc. [CD-ROM].

[21] P. Ciuciu, G. Marrelec, J.-B. Poline, J. Idier, and H. Benali, "Robust estimation of the hemodynamic response function in asynchronous multi-tasks multisessions event-related fMRI paradigms," in *Proc. IEEE ISBI*, 2002, pp. 847–850.

[22] P. Ciuciu, J.-B. Poline, G. Marrelec, J. Idier, C. Pallier, and H. Benali, "Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment," *IEEE Trans. Med. Imag.*, vol. 22, pp. 1235–1251, Oct. 2003.

[23] P. Ciuciu, J. Idier, G. Marrelec, and J.-B. Poline, "Estimation régularisée et non supervisée de la fonction de réponse hémodynamique en imagerie cérébrale fonctionnelle (IRMf)," in *Proc. GRETSI*, 2003, pp. 312–315.

[24] P. Ciuciu, J. Idier, and C. Pallier. (2003) Outlier detection for robust region-based estimation of the hemodynamic response function in event-related fMRI. [Online] SHFJ/CEA Tech. Rep.. Available: http://www.madic.org/biblio/en/Author/CIUCIU-P.php

[25] C. Robert, *Simulations par la méthode MCMC*, Paris: Economica, 1997.

[26] M.-H. Chen, Q.-M. Shao, and J. G. Ibrahim, *Monte Carlo Methods in Bayesian Computation*. New York: Springer, 2000.

[27] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge, U.K.: Cambridge Univ. Press, 2001.

[28] ——, "Evidential reasoning using stochastic simulation of causal models," *Artif. Intell.*, vol. 32, pp. 245–257, 1987.

[29] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, London, U.K.: Chapman & Hall, 1998.

[30] G. Marrelec and H. Benali, "Automated rejection sampling from product of distributions," *Comput. Statist.*, 2004, to be published.

[31] C. Gössl, D. Auer, and L. Fahrmeir, "Bayesian spatiotemporal inference in functional magnetic resonance imaging," *Biometrics*, vol. 57, pp. 554–562, 2001.

[32] C. S. Jensen, A. Kong, and U. Kjærulff, "Blocking Gibbs sampling in very large probabilistic expert systems," in *Int. J. Hum. Comput. Studies (Special Issue Real-World Applications of Uncertain Reasoning)*, vol. 42, 1993, Available: http://citeseer.nj.nec.com/jensen93blocking.html, pp. 647–666.

[33] G. Marrelec, P. Ciuciu, M. Pélégrini-Issac, and H. Benali, "Estimation of the hemodynamic response function in event-related functional MRI: Directed acyclic graphs for a general Bayesian framework," in *Lecture Notes in Computer Science*. ser. , C. Taylor and A. Noble, Eds, Berlin, Germany: Springer-Verlag, 2003, vol. 2732, Information Processing in Medical Imaging, pp. 635–646.

VII.4 Application en sciences cognitives : compréhension du langage

[a.4] Dehaene-Lambertz, Ghislaine and Dehaene, Stanislas and Anton, Jean-Luc and Campagne, Aurelie and Ciuciu, Philippe and Dehaene, Guillaume P. and Denghien, Isabelle and Jobert, Antoinette and Le Bihan, Denis and Sigman, Mariano and Pallier, Christophe and Poline, Jean-Baptiste, « Functional Segregation of Cortical Language Areas by Sentence Repetition », *Human Brain Mapping*, vol. 27,pp. 360–371, 2006.

Functional Segregation of Cortical Language Areas by Sentence Repetition

Ghislaine Dehaene-Lambertz,[1–5]* **Stanislas Dehaene,**[1,2,5] **Jean-Luc Anton,**[3]
Aurelie Campagne,[3] **Philippe Ciuciu,**[4,5] **Guillaume P. Dehaene,**[1]
Isabelle Denghien,[5] **Antoinette Jobert,**[1–5] **Denis LeBihan,**[4,5]
Mariano Sigman,[1–5] **Christophe Pallier,**[1–5] **and Jean-Baptiste Poline**[4,5]

[1]*Unité INSERM 562, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV, Orsay, France*
[2]*Collège de France, Paris, France*
[3]*Centre IRMf, IFR 131, CHU Timone, Marseille, France*
[4]*UNAF, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV, Orsay, France*
[5]*IFR 49, Institut d'imagerie neurofonctionnelle, France*

◆ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ◆

Abstract: The functional organization of the perisylvian language network was examined using a functional MRI (fMRI) adaptation paradigm with spoken sentences. In Experiment 1, a given sentence was presented every 14.4 s and repeated two, three, or four times in a row. The study of the temporal properties of the BOLD response revealed a temporal gradient along the dorsal–ventral and rostral–caudal directions: From Heschl's gyrus, where the fastest responses were recorded, responses became increasingly slower toward the posterior part of the superior temporal gyrus and toward the temporal poles and the left inferior frontal gyrus, where the slowest responses were observed. Repetition induced a decrease in amplitude and a speeding up of the BOLD response in the superior temporal sulcus (STS), while the most superior temporal regions were not affected. In Experiment 2, small blocks of six sentences were presented in which either the speaker voice or the linguistic content of the sentence, or both, were repeated. Data analyses revealed a clear asymmetry: While two clusters in the left superior temporal sulcus showed identical repetition suppression whether the sentences were produced by the same speaker or different speakers, the homologous right regions were sensitive to sentence repetition only when the speaker voice remained constant. Thus, hemispheric left regions encode linguistic content while homologous right regions encode more details about extralinguistic features like speaker voice. The results demonstrate the feasibility of using sentence-level adaptation to probe the functional organization of cortical language areas. *Hum Brain Mapp 27:360–371, 2006.* © 2006 Wiley-Liss, Inc.

Key words: MRI; brain; voice; speech; habituation; asymmetry; auditory cortex

◆ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ◆

INTRODUCTION

Linguistic tasks, from syllable discrimination to story listening, activate a large network of perisylvian areas. Contrary to neuropsychological studies, in which specific deficits, such as articulatory deficits [Dronkers, 1996], phonological deficits [Caplan et al., 1995; Caramazza et al., 2000], and semantic deficits [Hart et al., 1985], have begun to dissect the brain circuitry involved in language, brain imaging studies have been relatively less successful in the segregation of functionally distinct regions within this perisylvian network. For example, activations are detected in the left inferior frontal gyrus and insula in syllable discrimination tasks [Dehaene-Lambertz et al., 2005; Zatorre et al., 1992], metaphonological tasks [Burton et al., 2000], syntactical processing [Hashimoto and Sakai, 2002; Kaan and Swaab, 2002],

and sentence listening [Dehaene et al., 1997; Pallier et al., 2003]. Although careful subtractions between adequate conditions have identified clusters within this network, their precise involvement in the studied task is still under question after more than a decade of brain imaging research. In this article we tackle this question from another angle, combining timing information to identify clusters with different dynamic and functional properties within the cortical language network and the priming method.

The adaptation or priming method is based on the observation that stimulus repetition induces a decrease in brain activity (repetition suppression), which can be measured with event-related potentials (ERPs) [Dehaene-Lambertz and Dehaene, 1994] and functional MRI (fMRI) [Grill-Spector et al., 1998; Naccache and Dehaene, 2001]. More importantly, the same decrease in activity can be observed when the repeated stimuli are not exactly similar but only share a common property that is extracted at the probed cortical location. This method is thus sensitive to the code used in a particular brain region. By varying the property that is repeated, it becomes possible to target a specific representation and to uncover its cerebral bases. Priming designs have been successfully used to characterize the processing steps in visual object perception [Grill-Spector and Malach, 2001], reading [Dehaene et al., 2004], or number representation [Naccache and Dehaene, 2001; Piazza et al., 2004]. In this approach, the first step consists in isolating the network involved in the object representation by studying which brain regions display repetition suppression when the same object is repeated. Then in successive experiments shared properties between the repeated stimuli are varied in order to identify which coding variations are relevant and which are not relevant to obtain a repetition suppression effect in a given brain region.

Here we applied this reasoning to speech perception. In the first experiment, we aimed to isolate which brain regions are sensitive to the mere effect of repetition of the same sentence. In the second experiment, analyzed in the Functional Imaging Analysis Contest (FIAC), we moved forward to characterize the functional properties of these perisylvian regions by repeating only one property of the sentences, either the sentence content or the carrier of this content: the voice of the speaker.

EXPERIMENT 1

In this experiment, sentences were repeated two to four times in a row in a slow-event design. Our goal was to examine whether the linguistic network can be decomposed into spatially organized functional circuits based on the timing of their BOLD response and on the dynamics of their adaptation. Previous priming experiments used brief presentations in the visual domain (faces, objects, words, and numbers), but repetition suppression has been described in auditory neurons [Ulanovsky et al., 2003] as well as in visual neurons [Miller et al., 1991]. Furthermore, repetition of syllables and pseudowords induces a decrease in the amplitude of evoked responses or MRI activations [Belin and Zatorre,

2003; Cohen et al., 2004; Dehaene-Lambertz and Gliga, 2004], suggesting that repetition suppression may be a general property of the brain, not limited to visual perception. Nevertheless, sentences typically last a few seconds and multiple processes at different integration levels (acoustic, phonetic, lexical, syntactic, local semantic, contextual semantic) are required to perceive their similarity. The observation of repetition suppression effects at a long lag (14.4 s) would suggest that the priming method can also target such long time range representations.

Methods

Participants

Ten right-handed young French adults (5 women and 5 men, ages 21–35 years), with no history of oral or written language impairment, neurological or psychiatric disease, nor hearing deficits were tested. All participants gave their written informed consent and the study was approved by the local ethics committee.

Stimuli

A female French speaker was recorded reading a well-known children's story, "The Three Little Pigs." People know the characters and the general scenario of this story. However, sentences were rewritten and thus were original. The advantage of a children's story is to elicit attention from the listener because of frequent changes of register (dialogue, exclamation, description, etc.). The speaker used a theatrical intonation to reinforce this aspect. Sixty sentences with a mean duration of 2200 ms (1580–2863 ms) were extracted from the original story. The overlearned and expected sentences of the story were not presented. More details on the stimuli can be found in Dehaene-Lambertz and Houston [1998].

Experimental procedure

A randomly chosen sentence was presented every 14.4 seconds in a slow-event-related design. Each sentence was then repeated two, three, or four times in a row still with a 14.4-second intersentence interval. Each participant listened to a total of 180 sentences, 20 different sentences in each of the three repetition conditions. These repetition conditions were randomly ordered. The experiment was divided in five runs, each comprising 36 sentences and lasting 8 minutes 30 seconds. The participants passively listened to the sentences. To keep them minimally attentive during the experiment, they were instructed that they would have to recognize the story from which the sentences were extracted and to answer questions about the sentences at the end of the experiment. All participants recognized the story.

Image acquisition and analysis

The experiment was performed on a 3T whole body system (Bruker, Germany) equipped with a quadrature birdcage radio frequency (RF) coil and a head only gradient coil

insert designed for echoplanar imaging (26 contiguous axial slices, thickness 4.5 mm, TR = 2.4 s, TE = 40 ms, flip angle = 90°, field-of-view = 192 × 256 mm, 64 × 64 pixels). A high-resolution (1 × 1 × 1.2 mm^3), T1-weighted anatomical image using a 3-D gradient-echo inversion-recovery sequence was also acquired for each participant.

fMRI data analysis was performed using Statistical Parametric Mapping (SPM99, http://www.fil.ion.ucl.ac.uk/spm/). For each participant, after image reconstruction of the functional runs the first four volumes of each run were discarded to eliminate nonequilibrium effects of magnetization. The remaining volumes were corrected for the temporal delays between slices and realigned to the first volume. The mean realigned image was used to check the correct alignment of the functional images with the structural image. The T1-weighted structural MRI scan was spatially normalized to the Montreal Neurological Institute (MNI) template and the normalization parameters were applied to the functional images. These images were then resampled every 4 mm using bilinear interpolation and spatially smoothed with an isotropic Gaussian filter (kernel = 5 mm). The time series in each voxel was highpass-filtered (cutoff = 60 s), smoothed with a 4-s Gaussian kernel, and globally normalized with proportional scaling.

A linear model was generated by entering, for each run, four distinct variables corresponding to the four repetition positions of a sentence: first, second, third, and fourth presentations. Due to the paradigm construction, there were 60 events for the first and second presentation, 40 for the third presentation, and only 20 for the fourth presentation. The variables convolved by the standard SPM hemodynamic response function (HRF) and their temporal derivatives were included in the model. For random effect group analyses, the contrast images from individual participant's analysis were smoothed with an 8-mm Gaussian kernel and submitted to one-sample t-tests. We first examined the activations separately for the first to fourth presentations of each sentence. Second, we tested whether activations for a particular sentence repetition significantly decreased relative to the preceding sentence presentation. Finally, we tested for a significant linear decrease across repetition from the first presentation of the sentence to the fourth.

In order to better estimate the periodicity and phase of the event-related BOLD response, a distinct model was also designed by convolving the same four variables with a single cycle of a sine and a cosine waveform at a frequency of 14.4 s. The ratio of the regression weights of the sines and cosines was then transformed with the inverse tangent function to yield a BOLD response phase between 0 and 2π. This phase was multiplied by the stimulation period (14.4 s) and divided by 2π to yield a phase lag expressed in seconds. BOLD response amplitude was computed as the square root of the sum of the squares of the sine and cosine coefficients. To estimate the presence of a significant response at the stimulation period, an analysis across sessions was performed using the Rayleigh circular statistic, which contrasts the observed distribution of phases to the null hypothesis of a uniform distribution across the phase circle. To assess the presence of repetition effects, the images of amplitude and phase computed for each participant and for each repetition position were entered in separate random-effect group analyses and submitted to one-sample t tests.

If not reported otherwise, all the reported effects passed a voxelwise threshold of $P < 0.001$ uncorrected for multiple comparisons and a $P < 0.05$ threshold on the extent of clusters.

Results

Activations to the first presentation of a sentence involved both superior temporal gyri and sulci, the left inferior frontal region and insula, both thalami, and the right caudate nucleus, whereas for the fourth presentation, activations were limited to the middle part of the superior temporal gyrus, anteriorly and posteriorly to Heschl's gyri. As illustrated in Figure 1, a decrease in amplitude was already observed for the second presentation of the sentence and was significant in both temporal regions and in the left inferior frontal gyrus and insula (Table I). Although additional decreases with successive repetition were small, a significant linear effect was present mainly over the left superior temporal region (252 voxels, Fig. 1), while on the right side the linear effect of repetition was limited to 50 voxels (Table I). Some regions, such as Heschl's gyrus and its vicinity, presented no effect of repetition (Fig. 1).

We analyzed also the phase and periodicity of the BOLD response: Two effects are apparent in the hemodynamic response curves shown in Figure 1. First, the latency of the BOLD response seems to vary across regions, and second, it seems to decrease with sentence repetition. To rigorously quantify both effects, we fitted the whole-brain data with a set of sine and cosine functions in order to extract the phase and amplitude of the BOLD response (see Methods, above). We then submitted both parameters to a between-participants random effect analysis. The amplitude data essentially replicated the above SPM analysis, showing both a strong activation to the first sentence, in bilateral temporal and left inferior frontal cortex, as well as an adaptation effect particularly evident from the first to the second repetition. The phase analysis, however, yielded novel evidence in favor of a hierarchical temporal organization of these areas. In the image of the phase of the response to the first sentence, we observed, in all participants, a temporal progression with the earliest responses observed in Heschl's gyrus, followed by successively slower responses in posterior and middle STS, temporal pole, and inferior frontal gyrus. The phase averaged over subjects to the first and to the second sentence (converted in seconds) is shown in Figure 2. As illustrated in this figure, in several regions the phase showed a slight acceleration with sentence repetition. A random effect analysis was used to compare the phases to the first sentence and to the mean of the phases of repetitions 2–4. Search was confined to the network activated by the presentation of the first sentence (Fig. 1), and to increase sensitivity, the voxel threshold was decreased to $P < 0.01$ (still with corrected

Figure 1.
Repetition suppression related to sentence repetition. **A:** In Experiment 1: sagittal slice at $x = -49$ mm (standard Talairach coordinates) displaying the regions in which repetition induces a linear decrease in amplitude (blue-yellow scale), superimposed on the activations to the first sentence (yellow-red scale). Graphs show the adaptation of the mean BOLD response with sentence repetition at four locations that illustrate the different patterns of responses to sentence repetition. Note also the different shapes of the BOLD responses, with, for example, an earlier peak in the middle STS than in the posterior STS. Coordinates are given in standard Talairach coordinates. **B:** In Experiment 2: sagittal slice at $x = -55$ mm (Talairach coordinates) displaying the regions significantly more activated when both parameters, sentences and speakers, varied than when both were repeated (blue-yellow scale) superimposed on the activations to the first sentence (yellow-red scale).
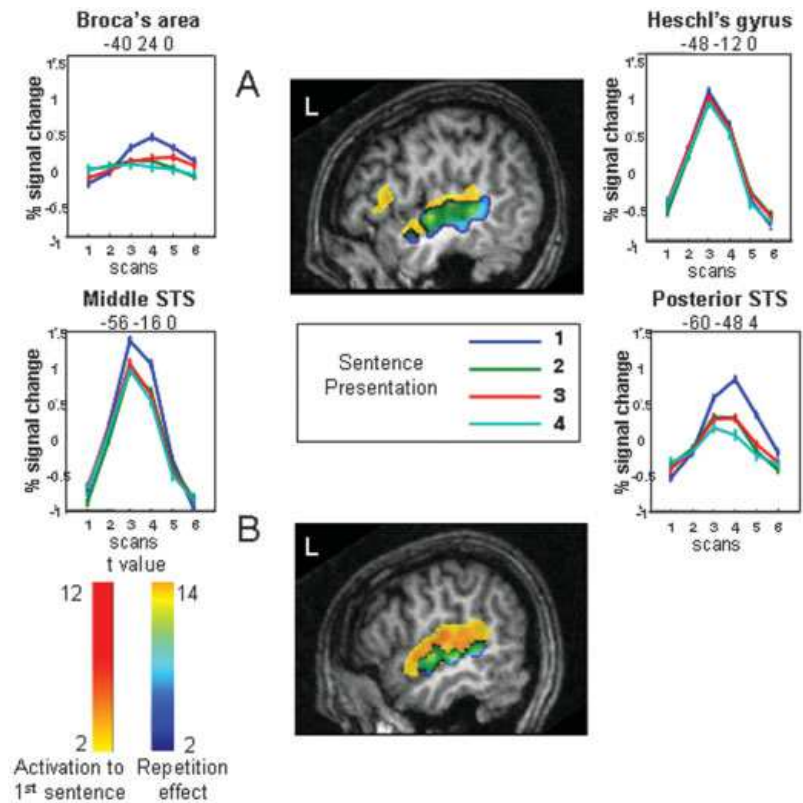
cluster-level $P < 0.05$). This analysis revealed a significant acceleration with repetition only in the right STS (Talairach coordinates: 60, -12, 0; Z = 3.75 and 56, -24, 0; Z = 2.96), and a marginally significant effect in the left STS (-56, -8, 0; Z = 3.43 and -56, -20, 0; Z = 3.24).

Discussion

Using a slow event-related paradigm, we observed a hierarchical temporal organization along the superior temporal regions with the earliest response in the primary auditory cortices and the slowest in associative regions with a caudal–rostral and a dorsal–ventral gradient along the superior temporal sulcus (Figs. 1, 2). Repetition of sentences affected both amplitude and phase of the HRF response. This adaptation effect was particularly evident between the first and second sentence, although it continues with the following repetitions. The pattern of adaptation was different across regions (Fig. 1) with a set of regions demonstrating the same response each time a sentence was presented (e.g., Heschl's gyrus), regions showing a strong decrease between the first and second presentation and more subtle decrease with the following repetitions (e.g., Broca's area) and regions that show a close to linear adaptation with repetition (e.g., superior temporal sulci).

The phase delay between the fastest region (Heschl's gyrus) and the slowest (temporal poles and inferior frontal regions) varies from 2 to 8 s. This delay cannot be related to

an acquisition delay between the different slices. First, we applied a temporal correction for slice acquisition times to the original data. Second, acquisition was from bottom to top, whereas the slowest phases are observed in the more basal slices. Third, the observed delays are much larger that the slice acquisition lags and, in several cases, multiple delays are seen in the same slice.

Several studies have already reported delays in the time course of activations in frontal areas relative to posterior areas [Schacter et al., 1997; Thierry et al., 1999]. Note that here we are measuring phase, a measure sensitive not only to the rise of the hemodynamic response, and thus to the peak latency as measured in previous experiments [Henson et al., 2002; Schacter et al., 1997; Thierry et al., 1999, 2003], but also to its plateau (possibly including neural responses posterior to sentence offset). This methodological difference may explain the large interval of phase lags observed (2–8 s). Although hemodynamic differences between cortical areas may contribute to the observed differences, the observed delays seem too large to be due solely to factors other than differences in cognitive processing. Indeed, several previous experiments also showed that BOLD response delays could be affected by cognitive factors. By comparing brain activity evoked by decision tasks bearing on the first or second presentation of a pair of nouns, Thierry et al. [2003] demonstrated that they could further delay the peak of the BOLD response in frontal areas. Conversely, Henson et al. [2002]

◆ Dehaene-Lambertz et al. ◆

TABLE I. fMRI activations in Experiment I

Area	No. voxels in cluster	Cluster-level P value (corrected)	Z value at local maximum	Talairach coordinates: x, y, z
1. First sentence				
Right STG and STS	332	<0.001	5.51	44, −28, 0
			4.57	60, −24, 4
			4.04	40, −48, 16
Left STG and STS	253	<0.001	5.15	−52, −32, 4
			3.88	−48, 0, −12
			3.50	−60, −56, 12
Left inferior frontal and insula	61	<0.001	4.02	−48, 16, 8
Thalami	140	<0.001	3.93	−16, −20, 16
			3.87	8, −12, 12
Right caudate nucleus	37	0.002	3.78	24, 24, 4
2. Fourth sentence				
Right STG	147	<0.001	4.23	48, −36, 0
			3.54	36, −20, −8
Left STG	78	<0.001	4.38	−60, −32, 4
			4.28	−52, −16, 0
3. First > second presentation				
Right STG and STS	252	<0.001	5.35	60, −36, 4
			5.12	48, −48, 8
			4.11	52, −12, −4
Left STG and MTG	147	<0.001	4.91	−52, −20, 0
			4.65	−60, −44, 4
Left insula	32	.02	3.67	−32, 16, 8
4. Regions displaying linear habituation				
Right STG and STS	50	0.005	3.92	44, −32, 4
			5.12	48, −16, 0
Left STG and STS	223	<0.001	5.27	−56, −16, 0
			4.79	−60, −48, 8

STG: superior temporal gyrus; STS: superior temporal sulcus; MTG: middle temporal gyrus.

studied repetition effects in face processing using a design comparable to the present one, and observed that the second presentation of a face elicited an earlier response in the right fusiform area. Similarly, in the present experiment sentence repetition accelerated the phase of the BOLD response, which would not be expected if the delays were due solely to hemodynamics.

While previous studies observed solely a difference in BOLD response delay between frontal regions and temporal regions [Schacter et al., 1997; Thierry et al., 2003], we demonstrate here a gradual organization of phases along the superior temporal regions with a slowing down extending in both anterior and posterior directions from Heschl's gyrus. Along with this anterior–posterior gradient, there is also a dorsal–ventral gradient with fastest phase along the sylvian fissure than along the superior temporal sulcus. These gradients are compatible with the known organization of auditory connections leading from the superior temporal region towards the temporal pole and the ventrolateral prefrontal cortex [Petrides and Pandya, 2002], through distinct ventral and dorsal pathways [Romanski et al., 1999]. The observed fMRI delays in excess of 1 or 2 s are unlikely to reflect the transmission of neural information from one region to the next, which occurs on the scale of a few tens of milliseconds. This gradient might rather be the result of different cognitive operations that integrate larger and pos-

sibly more abstract speech units that may require longer processing time, and/or with a more sustained activity.

This hypothesis of hierarchical integration of larger units along the superior temporal regions is confirmed by the repetition suppression effect. Regions coding segmental properties (acoustical and phonological) were expected to show no effect of repetition contrary to those processing supra-segmental properties (prosodic, semantic, and syntactic properties). Indeed, a similar anterior–posterior, dorsal–ventral gradient was observed for repetition effects, with both superior temporal sulci strongly affected by repetition. Roughly three types of adaptation patterns can be described: The fastest regions, Heschl's gyrus and its vicinity in the superior temporal gyrus, are not affected by repetition, either because they are unable to code an entire sentence, or to maintain this code during the time lag of 14.4 s separating two repetitions. The fast response of this region, its insensitivity to sentence repetition, confirms the role of segmental coding proposed for this auditory region with a time unit of a few tens or hundred milliseconds [Boemio et al., 2005]. At the opposite, Broca's area, one of the areas showing the slowest phase, is significantly activated only for the first presentation, displaying a sharp adaptation effect. Although Broca's region has been classically associated with syntactic processing, its role in syntax is now reassessed and discussed in a more general framework of working memory
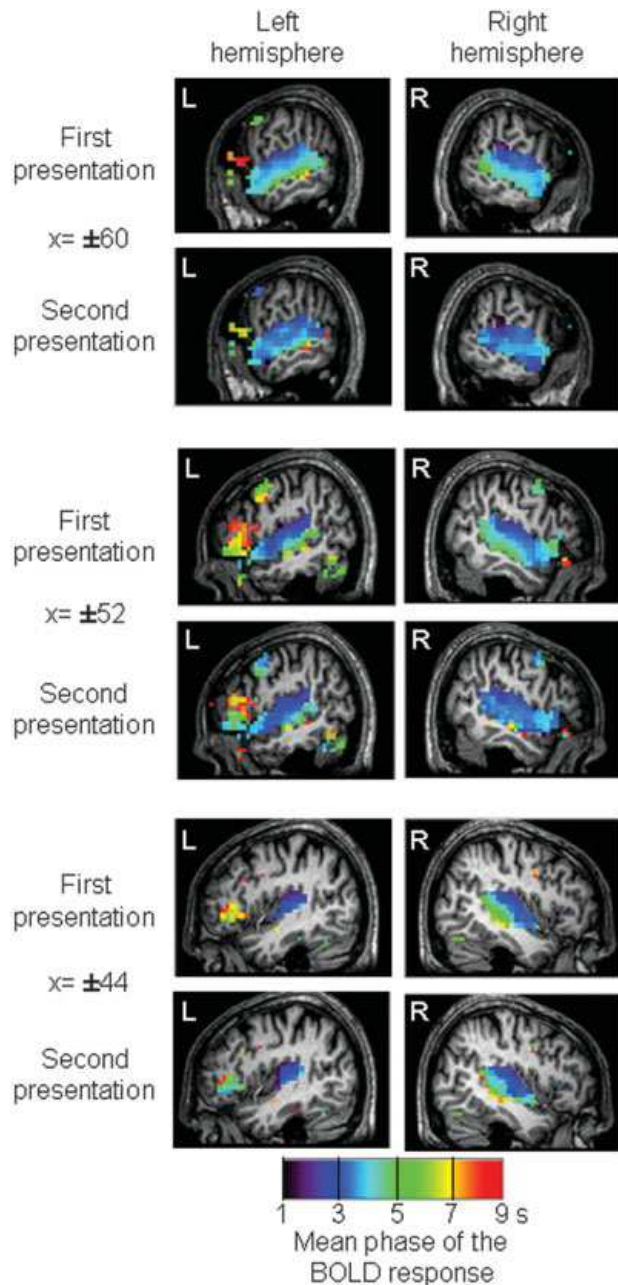
Figure 2.

Temporal organization of cortical responses to sentences for the first and second presentation of the same sentence. Colors encode the circular mean of the phase of the BOLD response, expressed in seconds relative to sentence onset. Fastest responses, in purple, are visible in Heschl's gyrus, while the slowest responses, in yellow and red, are encountered in Broca's area. A delay in phase is visible along the superior temporal region with a dorsal–ventral and anterior–posterior gradient. The second presentation of a sentence speeds up the phase in all these regions.

load [Kaan and Swaab, 2002]. In our experiment, the syntax was highly varied across sentences but the sentences were short, with no particular syntactical ambiguities or difficulties. A significant activation present in this region only for the first presentation suggests two different hypotheses. First, once the syntactic tree is computed for the first presentation, it might be easily reapplied to successive presentations with no renewed computations. Second, this region might be incidental in sentence perception and might be recruited only when participants have time to rehearse the sentence in order to integrate it in the global context of the story. This single experiment cannot separate these hypotheses.

Finally, temporal regions with a median phase response showed a clear effect of repetition suppression but remained significantly activated even when the sentence was presented for the fourth time. The decrease of the amplitude of the BOLD response and the speeding-up of the phase (Figs. 1, 2) along the superior temporal region might reflect either a general decrease of attention and/or an improved contextual integration of the sentence elements in the local context of the sentence and in the global context of the story.

To summarize, repetition induced a decrease in amplitude and a speeding up of the phase of the BOLD response with a gradient roughly similar to what is obtained by measuring the phase of the BOLD response. In order to deepen our understanding of the role of these regions, our second step involved manipulating which parameter was repeated. Among the regions that show repetition suppression in Experiment 1, can we separate those sensitive to the linguistic content and those sensitive to the carrier of this content, i.e., the speaker's voice?

EXPERIMENT 2

A spoken utterance conveys not only linguistic information but also information about the speaker identity and their emotional state. Each individual possesses voice characteristics due to the configuration of his/her vocal tract, his/her pronunciation, his/her dialectal accent making him/her recognizable even when the listener does not previously know the speaker. The capacity to identify speaker voice is present very early on. Infants [Mehler et al., 1978], even fetuses [Kisilevsky et al., 2003], are able to recognize their mother and to discriminate stranger's voices [Dehaene-Lambertz, 2000].

Classical theories of speech perception and word recognition postulate the existence of an early processing stage of "speaker normalization" that deletes voice-specific features from the acoustic signal. Yet some researchers believe that such a processing stage is not necessary and have proposed that word recognition proceeds by comparing the acoustic input with multiple exemplars (or templates) stored in lexical memory [Goldinger, 1998]. Features of the speaker voice are assumed to remain encoded in those exemplars. Evidence in favor of this claim comes from word identification experiments using repetition priming that showed that a given word is better identified when it is repeated with

similar acoustic features (e.g., when spoken by the same speaker) than when it is repeated with different acoustic details [Church and Schacter, 1994; Goldinger, 1996].

Brain imaging provides the opportunity to disentangle brain regions that are sensitive to the repetition of abstract linguistic content, and regions that are sensitive to the repetition of more detailed acoustic features. The goal of our second experiment was thus to separate surface and abstract representations of sentences by using a repetition paradigm. We manipulated independently the repetition of two parameters: speaker and sentence content. Either the same sentence was repeated but said by different speakers (SSt-DSp) or the speaker was constant but produced different sentences (DSt-SSp). These conditions were compared with two other conditions in which the same sentence produced by the same speaker was repeated several times (SSt-SSp), and different sentences were produced by different speakers (DSt-DSp). We expected a decrease of activity in regions that code for linguistic information when the same sentence was repeated. If there is normalization, the same decrease should be present when the speaker is held constant or when it is varied. Similarly, regions that code for speaker identity should show a decrease when the speaker is held constant, even when the linguistic content is varied. In order to avoid an exceedingly long experimental duration, we used short blocks of six sentences (block-design) that were either all similar or all different for the parameter of interest.

Methods

Participants

Sixteen right-handed young French adults with no history of oral or written impairment, neurological or psychiatric disease, or hearing deficits were participants. They gave their written informed consent, and the study was approved by ethics committee. Acquisition was stopped in Participant 5, leaving 15 participants complete the data set (8 women and 7 men; mean age, 27 years; range, 21–35).

Stimuli

A female French speaker was recorded reading a well-known children's story, "The Three Little Pigs," with a theatrical intonation. Sentences were rewritten and thus were original. We extracted 64 sentences from the same story used in Experiment 1. Each sentence was secondarily recorded by 9 other French speakers with the instruction to not only repeat the sentence but to follow as much as possible the intonation and rhythm of the first speaker. We obtained 640 sentences; i.e., 64 sentences produced by 5 men and 5 women with a mean duration of 2277 ms (1515–2793 ms).

Experimental procedure

The experiment was divided into four runs, two runs with a block design and two with an event-related design. The order of the four runs was counterbalanced across participants. In both designs, the instructions given to the participants were to

lie still in the scanner with eyes closed and to attentively listen to the sentences, because they would be asked whether sentences presented after scanning were presented during the experiment. Only the block design is analyzed below, whereas the event-related design is analyzed in other papers in this special issue (see Poline et al., Table I).

Block design. The sentences were presented in small blocks of 6 sentences, one every 3333 ms (thus the silence between sentences was variable because of the variation in sentence length). There were four types of blocks: Same Sentence-Same Speaker (SSt-SSp), the same sentence said by the same speaker was repeated 6 times; Same Sentence-Different Speakers (SSt-DSp), the same sentence was repeated by 6 different speakers (3 men and 3 women); Different Sentences-Same Speaker (DSt-SSp), the same speaker produced 6 different sentences; Different Sentences-Different Speakers (DSt-DSp), 6 different speakers (3 men and 3 women) produced 6 different sentences. For each block, the speakers were randomly selected among the 10 original speakers with the constraint that male and female voices were equally distributed within each block and across blocks. For all conditions, sentences were randomly selected on line with the constraint that a sentence could not be repeated in the first 8 blocks. The selection of the sentences for the condition (SSt-DSp) was limited to a subset of 96 sentences (16 sentences × 6 speakers) in which prosody was the most homogeneous among speakers. The order of speakers was randomly selected for each block in which different speakers were present. Each block lasted 20 seconds and was followed by a silence of 9 seconds. The order of the conditions was randomly selected for each participant with the constraint that no more than two blocks of the same condition were presented in a row. Across the two runs used in the block design, a total of 8 blocks were obtained in each condition presented (4 blocks per run and per condition).

Event-related design. One sentence was presented every 3333 ms. The same conditions as above were presented but were defined in this design by the transition between two sentences. Four types of transitions were thus possible: The sentence was similar to the previous one (SSt-SSp), or there was a change of speaker (SSt-DSp), or a change of sentence (DSt-SSp), or a change of sentence and speaker (DSt-DSp). Note that the change of speaker could be either within the same sex, or across sexes. Participants performed two runs of 141 sentences each giving 72 transitions in each condition.

Image acquisition

The experiment was performed on a 3-T whole body system (Bruker, Germany), equipped with a quadrature birdcage radio frequency (RF) coil. Functional images comprising 30 axial slices (thickness = 4 mm) covering most of the brain were obtained with a T2-weighted gradient echo, EPI sequence (interleaved acquisition with the following parameters: TR, 2.5 s; TE, 35 ms; flip angle, 80°; field-of-view, 192 × 192 mm; 64 × 64 pixels). A high-resolution (1 × 0.9 × 1.4 mm), T1-weighted, anatomical image using a 3-D gradi-

◆ **A**daptation to **S**entence **R**epetition ◆

TABLE II. fMRI activations in Experiment 2

Area	No. voxels in cluster	Cluster-level P value (corrected)	Z value at local maximum	Talairach coordinates: x, y, z
1a. Main effect of sentence repetition				
Left middle STS	61	0.016	4.21	$-63, -15, 0$
Left posterior MTG	25	0.246	3.46	$-57, -42, 0$
1b. Asymmetries in the main effect of sentence repetition				
Left posterior MTG	89	<0.001	3.96	$-54, -42, 0$
Left middle STS	15	0.342	3.57	$-60, 15, 6$
2. Main effect of speaker repetition				
No significant regions				
3. Interactions speaker × sentence repetitions				
No significant regions				
4a. Effect of sentence repetition restricted to same speaker condition				
Left middle STS	49	0.055	4.54	$-63, -12, -3$
Left posterior MTG	62	0.024	4.07	$-60, -42, 0$
4b. Asymmetries in the effect of sentence repetition restricted to same speaker condition				
Left posterior MTG				
and STS	70	0.001	3.99	$-57, -42, 0$
5a. Effect of sentence repetition restricted to different speaker condition				
Left middle STS	52	0.023	3.84	$-63, -15, -3$
5b. Asymmetries in the effect of sentence repetition restricted to different speaker condition				
Left middle STS	41	0.001	4.62	$-57, -12, 6$
Left posterior MTG	46	0.006	3.90	$-60, -39, 0$

STG: superior temporal gyrus; STS: superior temporal sulcus, MTG: middle temporal gyrus.

ent-echo inversion-recovery sequence was also acquired for each participant.

Image analysis

fMRI data analysis was performed using Statistical Parametric Mapping (SPM2, http://www.fil.ion.ucl.ac.uk/spm/). Preprocessing was the same as Experiment 1 except that the images were resampled every 3 mm and that the time series in each voxel was highpass-filtered with a cutoff of 128 s.

Block design analysis. For each participant a linear model was generated by entering, for each run, five distinct variables corresponding to the first sentence pooled across all conditions and to the second to sixth sentences, separately for the four conditions. The logic was that the conditions differed only after the second sentence and that the response to the first sentence might differ because it follows a long period of silence. The variables were convolved by the standard SPM hemodynamic response function (HRF). For random effect group analyses, the individual contrast images from individual participants were smoothed with an 8-mm Gaussian kernel and submitted to one-sample t-tests. The design being a 2×2 factorial design, we computed the main effects of sentence and of speaker and their interactions. From a cognitive point of view, we were interested in the brain regions that were sensitive to sentence repetition even when acoustic variability due to a change of speaker was present, and to the brain regions sensitive to speaker repetition when sentences were constant or varied. Thus, we also computed the corresponding four contrasts.

We tested also for significant left–right asymmetries in these analyses: For each participant the transformation matrix of the anatomical image toward its flipped image by a nonlinear normalization was computed and then applied to the smoothed contrast images of the previous analyses to obtain their flipped images. Individual asymmetry images were obtained by subtracting the original contrast image from its flipped version. Those images were then entered into a random-effect analysis, which tested whether the amount of activation for the contrast of interest was significantly larger in one hemisphere relative to the other.

If not reported otherwise, all the reported effects passed a voxelwise threshold of $P < 0.001$ uncorrected for multiple comparisons and a $P < 0.05$ threshold on cluster extent.

Results

The first sentence of all blocks activated bilateral superior temporal regions and the left posterior part of the brainstem (colliculi). A significant asymmetry favoring the left side was observed in the posterior part of the superior temporal sulcus extending dorsally and medially over the planum temporale (272 voxels, Z = 4.75 at $x = -39$, $y = -36$, $z = 12$, and Z = 4.7 at $x = -57$, $y = -39$, $z = 6$).

A main effect of sentence repetition was observed in the middle part of the left temporal sulcus. This region became habituated even when speakers varied (DStDSp-SStDSp), demonstrating a normalization property. A significant asymmetry favoring the left side was present for the latter comparison in this cluster and in a second more posterior cluster located in the middle temporal gyrus (Table II).
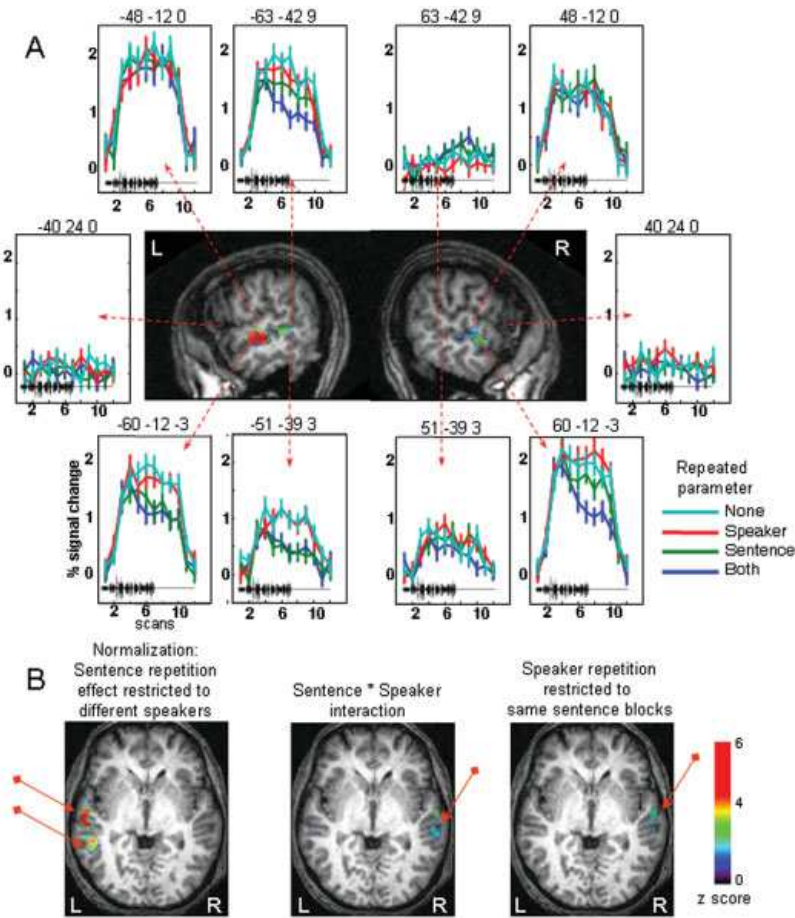
Figure 3.

Repetition suppression related to speaker and sentence repetition in Experiment 2. **A:** Graphs, surrounding sagittal slices at $x = \pm 60$ mm (standard Talairach coordinates), show the adaptation of the mean BOLD response in the different conditions. The black tracing on each plot represents the duration of the speech stimuli (~ 19 s). Several patterns are seen. The superior temporal regions (e.g., -48 mm, -12 mm, 0 mm) are activated and do not adapt, while inferior frontal regions (-40 mm, 24 mm, 0 mm) are inactive. The left STS show adaptation to sentence repetition, even when speakers varied (e.g., -60 mm, -12 mm, -3 mm). Right homologous regions either did not adapt (e.g., 51 mm, -39 mm, 3 mm) or adapted only when the speaker was maintained constant (60 mm, -12 mm, -3 mm). **B:** Axial slices at z = 0 mm (Talairach coordinates), displaying the clusters activated in the different comparisons.

This classical SPM analysis did not reveal any effect of speaker repetition nor significant interaction between speaker and sentence repetition factors. There were also no significant asymmetries for these contrasts. However, more sensitive analyses based on permutations and mixed-effect model, as presented by Meriaux et al. [2006], found a significant main effect of speaker repetition in the left superior temporal sulcus ($x = -63, y = -42, z = -9$) and a significant interaction between Sentence and Speaker repetition in the right middle temporal gyrus ($x = 60, y = -12, z = -3$) (Fig. 3; see tables in Meriaux et al. [2006]). The latter region significantly habituated when the speaker was kept constant relative to blocks of varied speakers; however, it did so only when the sentence was maintained constant (SStDSp-SStSSp). In the latter comparison, a second cluster was also significant in the left superior temporal gyrus ($x = -60, y = -30, z = 6$).

Discussion

In this second experiment where a change of sentence was contrasted with a change of speaker, we observed a reliable effect of adaptation to the linguistic content in the left superior temporal sulcus that was independent of the speaker, while the converse adaptation to speaker identity was not observed. Adaptation to sentences was significantly asymmetric, present only in the middle and posterior left temporal region, while the contralateral right regions were unable to normalize across speakers. The right temporal region habituated to sentence repetition only when the speaker was maintained constant.

Speech normalization

Speech is produced by different vocal tracts inducing variations in the acoustical cues that support linguistic content. Normalization of speech perception across many different speakers has long been recognized as one of the crucial difficulties that the brain has to resolve. Liberman [1985] proposed that normalization across speakers is realized through the recognition of the motor pattern or gesture that underlies the overt phoneme realization. The recent discovery of mirror neurons in a possible equivalent of Broca's area in the macaque has given biological support to the notion of a common representation between auditory or visual percepts and motor patterns [Kohler et al., 2002]. However, we found no evidence of Broca's involvement in the normalization process engaged here (Fig. 3). Because the

same sentences were used in Experiments 1 and 2, the activation that we found in Broca's area in Experiment 1 might be related to the participants' rehearsal of the sentence, which is possible during the pauses of the slow-event paradigm. The incidental involvement of the left inferior frontal region in sentence comprehension is further underscored by a transcranial magnetic stimulation (TMS) study showing that magnetic stimulation over that region had no effect on a participant's sentence perception, contrary to the same stimulation over the posterior temporal area that facilitates native language processing [Andoh et al., 2006]. These results are in contradiction to those reported by Sakai et al. [2002] and Hashimoto and Sakai [2002] underscoring the involvement of the left inferior frontal gyrus in sentence comprehension. However, as the authors pointed out, the involvement of this region in speech comprehension was demonstrated only when the task explicitly requires the use of syntactic rules or when participants have to recover words in degraded speech [Davis and Johnsrude, 2003]. Broca's area might thus be involved in speech listening when there is a conscious effort in sentence processing to integrate the different elements of the sentence.

The only regions displaying a normalization property were in the superior temporal sulcus, a unimodal auditory region [Poremba et al., 2003] that reacts more to speech than to other auditory stimuli (see Binder et al. [2000] for a meta-analysis of speech vs. nonspeech studies). In particular, using sinewave speech we observed that an area encompassing the two maxima observed here was more activated when stimuli were perceived as speech than when the same stimuli were perceived as whistles [Dehaene-Lambertz et al., 2005]. The same clusters are also activated by different types of degraded speech but only when they are intelligible [Davis and Johnsrude, 2003; Narain et al., 2003]. These results underscore that this part of the STS is not sensitive to the surface form of the auditory stimuli, but rather to the linguistic representations elicited by them. In an experiment of word repetition contrasting oral and visual modalities, Cohen et al. [2004] proposed that an equivalent of the visual word form area might exist in the superior temporal sulcus. This auditory word form area (AWFA), which showed repetition suppression within the auditory modality but not across different modalities, would be tuned to recognize words in the auditory environment irrespective of irrelevant features, such as speaker identity, pitch, speech rate, etc. It is noteworthy that the tentative coordinates that they propose for the AWFA ($-60, -8, -4$) fall close to the main peak of speaker independent adaptation observed in the present study ($-63, -15, -3$). We also observed a second spot presenting these characteristics, posterior to the auditory cortex. Based on this experiment we cannot conclude whether the code unit in these regions is limited to words or whether it can integrate an entire sentence. However, one might speculate, following the classical distinction between dorsal and ventral pathways [Hickok and Poeppel, 2000; Scott and Johnsrude, 2003], that these two clusters might be associated with different coding schemes: the anterior region would

map the word form onto lexical representations, while the posterior region, through the dorsal pathway, would maintain the phonological word form in order to interface with the working memory and the motor systems. In a similar priming study, but using written sentences, Noppeney and Price [2004] reported syntactic adaptation in the left temporal pole (Talairach coordinates: $-42, 3, -27$). We could have missed this region because of magnetic susceptibility artifacts in the temporal pole. However, the major difference between both studies, aside from the visual vs. auditory modality, is that the depth of syntactic computations might be different in both tasks. Similarity between sentences could be detected at several levels here. Our sentences were also simple to understand and never comprised ambiguous syntactic structures that might have pushed the participants in Noppeney and Price's study to rely more on syntactic processing in order to understand the sentences even in nonambiguous cases (see Hahne and Friederici [1999] for the effect of the proportion of syntactically ambiguous sentences).

Hemispheric asymmetry

The functional asymmetry between the two superior temporal sulci is striking. Repetition suppression is observed in left and right homologous clusters in the middle superior temporal sulcus. However, on the right side this effect is limited to the condition where both the sentence and the speaker remain constant. Contrary to the left cluster, the right cluster does not possess normalization capacities. These complementary codes, one abstract on the left side, the other exemplar-dependent on the right side, might explain performances in behavioral priming experiments. Lexical decision, which is primed by previous presentations of the word, even produced by different voices [Luce and Lyons, 1998], would be based on the left abstract code while explicit word recognition, which is affected by voice differences [Church and Schacter, 1994], might be informed by both regions. This left–right difference is reminiscent of another similar asymmetry described in the fusiform area for visual stimuli. Whereas the left fusiform area displays repetition suppression when a different view of the object is repeated, the homologous right region shows repetition suppression only when the same view is repeated [Vuilleumier et al., 2002]. The authors argued that this difference was independent of lexico-semantic factors but related to view-invariant properties of the visual code in the left hemisphere. In both visual and auditory perception the left hemisphere appears to compute an abstract categorical representation, stripped of sensory details, while the right homologous region maintains more surface details.

Adaptation to the speaker

While we observed regions that habituated to sentence, we did not find regions that habituated to speaker repetition. In our experiment, although the participants did not know the speakers, we could have expected a difference because blocks with different speakers comprise male and

female voices, contrary to blocks with the same speaker. The right anterior STS has been identified as an important region involved in identifying speakers [Belin et al., 2004; von Kriegstein and Giraud, 2004]. In an adaptation paradigm, Belin and Zatorre [2003] compared blocks of the same repeated syllable produced by different speakers with blocks of different syllables produced by the same voice. The right anterior STS was the only region that differed between these conditions. Here, in a similar passive task but with complex linguistic stimuli, we were not able to observe a similar adaptation effect. However, we might also have missed this region because of signal loss due to magnetic susceptibility artifacts (the peak maximum Talairach coordinates: 58, 2, −8, reported by Belin and Zatorre [2003] is outside the intersection of all the participants' masks). Von Kriegstein and Giraud [2004] also noticed that the right posterior STS was activated when participants had to recognize a precise voice, particularly if the voice was previously unknown. This finding is compatible with the exemplar-based representation of sentences that we observed in that region. In order to recognize a particular voice, the participants may emphasize the surface details encoded in this region.

CONCLUSION

Our results demonstrate the feasibility of using sentence-level adaptation to probe the functional organization of cortical language areas. Experiment 2 was a first step to segregate the perisylvian network by contrasting voice and linguistic content. Hierarchical comparisons between the different linguistic parameters should normally follow. A recent experiment [Kouider and Dupoux, 2005] demonstrating that it is possible to obtain subliminal priming with auditory stimuli opens the possibility of more refined tools to study speech comprehension, as was done for visual perception.

REFERENCES

Andoh J, Artiges E, Pallier C, Riviere D, Mangin JF, Cachia A, Plaze M, Paillere-Martinot ML, Martinot JL (2006): Modulation of language areas with functional MR image-guided magnetic stimulation. Neuroimage 29:619–627.

Belin P, Zatorre RJ (2003): Adaptation to speaker's voice in right anterior temporal lobe. Neuroreport 14:2105–2109.

Belin P, Fecteau S, Bedard C (2004): Thinking the voice: neural correlates of voice perception. Trends Cogn Sci 8:129–135.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000): Human temporal lobe activation by speech and non speech sounds. Cereb Cortex 10:512–528.

Boemio A, Fromm S, Braun A, Poeppel D (2005): Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389–395.

Burton MW, Small SL, Blumstein SE (2000): The role of segmentation in phonological processing: an fMRI investigation. J Cogn Neurosci 12:679–690.

Caplan D, Gow D, Makris N (1995): Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. Neurology 45:293–298.

Caramazza A, Chialant D, Capasso R, Miceli G (2000): Separable processing of consonants and vowels. Nature 403:428–430.

Church BA, Schacter DL (1994): Perceptual specificity of auditory priming: implicit memory for voice intonation and fundamental frequency. J Exp Psychol Learn Mem Cogn 20:521–533.

Cohen L, Jobert A, Le Bihan D, Dehaene S (2004): Distinct unimodal and multimodal regions for word processing in the left temporal cortex. Neuroimage 23:1256–1270.

Davis MH, Johnsrude IS (2003): Hierarchical processing in spoken language comprehension. J Neurosci 23:3423–3431.

Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, van de Moortele PF, Lehéricy S, Le Bihan D (1997): Anatomical variability in the cortical representation of first and second languages. Neuroreport 8:3809–3815.

Dehaene S, Jobert A, Naccache L, Ciuciu P, Poline JB, Le Bihan D, Cohen L (2004): Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. Psychol Sci 15:307–313.

Dehaene-Lambertz G (2000): Cerebral specialization for speech and non-speech stimuli in infants. J Cogn Neurosci 12:449–460.

Dehaene-Lambertz G, Dehaene S. (1994): Speed and cerebral correlates of syllable discrimination in infants. Nature 370:292–295.

Dehaene-Lambertz G, Gliga T (2004): Common neural basis for phoneme processing in infants and adults. J Cogn Neurosci 16:1375–1387.

Dehaene-Lambertz G, Houston D (1998): Faster orientation latency toward native language in two-month-old infants. Lang Speech 41:21–43.

Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolle L, Dehaene S (2005): Neural correlates of switching from auditory to speech perception. Neuroimage 24:21–33.

Dronkers NF (1996): A new brain region for coordinating speech articulation. Nature 384:159–161.

Goldinger SD (1996): Words and voices: episodic traces in spoken word identification and recognition memory. J Exp Psychol Learn Mem Cogn 22:1166–1183.

Goldinger SD (1998): Echoes of echoes? An episodic theory of lexical access. Psychol Rev 105:251–279.

Grill-Spector K, Malach R (2001): fMR-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol (Amst) 107:293–321.

Grill-Spector K, Kushnir T, Edelman S, Itzchak Y, Malach R (1998): Cue-invariant activation in object-related areas of the human occipital lobe. Neuron 21:191–202.

Hahne A, Friederici AD (1999): Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. J Cogn Neurosci 11:194–205.

Hart J Jr, Berndt RS, Caramazza A (1985): Category-specific naming deficit following cerebral infarction. Nature 316:439–440.

Hashimoto R, Sakai KL (2002): Specialization in the left prefrontal cortex for sentence comprehension. Neuron 35:589–597.

Henson RN, Price CJ, Rugg MD, Turner R, Friston KJ (2002): Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. Neuroimage 15:83–97.

Hickok G, Poeppel D (2000): Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 4:131–138.

Kaan E, Swaab TY (2002): The brain circuitry of syntactic comprehension. Trends Cogn Sci 6:350–356.

Kisilevsky BS, Hains SM, Lee K, Xie X, Huang H, Ye HH, Zhang K, Wang Z (2003): Effects of experience on fetal voice recognition. Psychol Sci 14:220–224.

Kohler E, Keysers C, Umilta MA, Fogassi L, Gallese V, Rizzolatti G (2002): Hearing sounds, understanding actions: action representation in mirror neurons. Science 297:846–848.

◆ **Adaptation to Sentence Repetition** ◆

Kouider S, Dupoux E (2005): Subliminal speech priming. Psychol Sci 16:617–625.

Liberman AM, Mattingly IG (1985): The motor theory of speech perception revised. Cognition 21:1–36.

Luce PA, Lyons EA (1998): Specificity of memory representations for spoken words. Mem Cogn 26:708–715.

Mehler J, Bertoncini J, Barrière M, Jassik-Gerschenfeld D (1978): Infant recognition of mother's voice. Perception 7:491–497.

Mériaux S, Roche A, Dehaene-Lambertz G, Thirion B, Poline JB (2006): Combined permutation test and mixed-effect model for group average analysis in fMRI. Hum Brain Mapp 27:402–410.

Miller EK, Li L, Desimone R (1991): A neural mechanism for working and recognition memory in inferior temporal cortex. Science 254:1377–1379.

Naccache L, Dehaene S (2001): The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. Cereb Cortex 11:966–974.

Narain C, Scott SK, Wise RJ, Rosen S, Leff A, Iversen SD, Matthews PM (2003): Defining a left-lateralized response specific to intelligible speech using fMRI. Cereb Cortex 13:1362–1368.

Noppeney U, Price CJ (2004): An fMRI study of syntactic adaptation. J Cogn Neurosci 16:702–713.

Pallier C, Dehaene S, Poline JB, LeBihan D, Argenti AM, Dupoux E, Mehler J (2003): Brain imaging of language plasticity in adopted adults: can a second language replace the first? Cereb Cortex 13:155–161.

Petrides M, Pandya DN (2002): Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur J Neurosci 16:291–310.

Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S (2004): Tuning curves for approximate numerosity in the human intraparietal sulcus. Neuron 44:547–555.

Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M. (2003): Functional mapping of the primate auditory system. Science 299:568–572.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP (1999): Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–1136.

Sakai KL, Noguchi Y, Takeuchi T, Watanabe E (2002): Selective priming of syntactic processing by event-related transcranial magnetic stimulation of Broca's area. Neuron 35:1177–1182.

Schacter DL, Buckner RL, Koutstaal W, Dale AM, Rosen BR (1997): Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. Neuroimage 6:259–269.

Scott SK, Johnsrude IS (2003): The neuroanatomical and functional organization of speech perception. Trends Neurosci 26:100–107.

Thierry G, Boulanouar K, Kherif F, Ranjeva JP, Demonet JF (1999): Temporal sorting of neural components underlying phonological processing. Neuroreport 10:2599–2603.

Thierry G, Ibarrola D, Demonet JF, Cardebat D (2003): Demand on verbal working memory delays haemodynamic response in the inferior prefrontal cortex. Hum Brain Mapp 19:37–46.

Ulanovsky N, Las L, Nelken I (2003): Processing of low-probability sounds by cortical neurons. Nat Neurosci 6:391–398.

von Kriegstein K, Giraud AL (2004): Distinct functional substrates along the right superior temporal sulcus for the processing of voices. Neuroimage 22:948–955.

Vuilleumier P, Henson RN, Driver J, Dolan RJ (2002): Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. Nat Neurosci 5:491–499.

Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. Science 256:846–849.

CHAPITRE VIII

DÉTECTION-ESTIMATION CONJOINTE DE L'ACTIVITÉ CÉRÉBRALE EN IRMF

VIII.1　Étude de faisabilité

[a.8] S. Makni, P. Ciuciu, J. Idier et J.-B. Poline, « Joint detection-estimation of brain activity in functional MRI : a multichannel deconvolution solution », *IEEE Trans. Sig. Proc.*, vol. 53, 9, pp. 3488–3502, septembre 2005.

Joint Detection-Estimation of Brain Activity in Functional MRI: A Multichannel Deconvolution Solution

Salima Makni, Philippe Ciuciu, *Member, IEEE*, Jérôme Idier, and Jean-Baptiste Poline

Abstract—Analysis of functional magnetic resonance imaging (fMRI) data focuses essentially on two questions: first, a detection problem that studies which parts of the brain are activated by a given stimulus and, second, an estimation problem that investigates the temporal dynamic of the brain response during activations. Up to now, these questions have been addressed independently. However, the activated areas need to be known prior to the analysis of the temporal dynamic of the response. Similarly, a typical shape of the response has to be assumed *a priori* for detection purpose. This situation motivates the need for new methods in neuroimaging data analysis that are able to go beyond this unsatisfactory tradeoff. The present paper raises a novel detection-estimation approach to perform these two tasks simultaneously in region-based analysis. In the Bayesian framework, the detection of brain activity is achieved using a mixture of two Gaussian distributions as a prior model on the "neural" response levels, whereas the hemodynamic impulse response is constrained to be smooth enough in the time domain with a Gaussian prior. All parameters of interest, as well as hyperparameters, are estimated from the posterior distribution using Gibbs sampling and posterior mean estimates.

Results obtained both on simulated and real fMRI data demonstrate first that our approach can segregate activated and nonactivated voxels in a given region of interest (ROI) and, second, that it can provide spatial activation maps without any assumption on the exact shape of the Hemodynamic Response Function (HRF), in contrast to standard model-based analysis.

Index Terms—Bayesian analysis, detection-estimation, event-related fMRI, Gibbs sampling, HRF modeling, semi-blind deconvolution.

I. INTRODUCTION

THE overall aim of functional magnetic resonance imaging (fMRI) is to advance in the understanding of the relation between functions (cognitive or sensori-motor ones) and structure in the humain brain. To this end, current fMRI paradigms consist of various stimulus types or conditions (visual, auditory, etc.) presented to the subject while brain volumes are acquired (typically every few seconds). Each stimulus induces a neuronal activation, which itself is responsible for some changes

of physiological parameters (blood flow, blood volume, deoxy-hemoglobin concentration, etc.) leading to a local increase of blood oxygenation, which is referred to as the blood oxygenated level-dependent (BOLD) effect [1].

The fMRI data can be analyzed using exploratory methods (Cluster analysis [2], [3], principal component analysis (PCA) [4], [5], and independent component analysis (ICA) [6], [7]), but the most popular approach is based on massively univariate (voxelwise) regression techniques implemented in many paradigms such as statistical parametric mapping (SPM) [8], the FMRIB Software Library (FSL), and Analysis of Functional Neuroimages (AFNI). In this latter framework, once a model has been specified, Student-t or Fisher statistics are calculated over the whole brain in order to identify the regions that are activated in response to a given contrast of experimental conditions. However, methods to estimate the temporal dynamic in brain regions have received less attention and have yet to be further developed in neuroimaging. In this paper, we propose to merge detection and estimation of the dynamic in one step in a Bayesian framework.

For detection purposes, a general linear time invariant (LTI) model, coding for the so-called *design matrix*, is built to assess the link between the fMRI data and the expected BOLD signal (regressors) in any voxel of the brain. Typically, each column of this matrix defines a *regressor*, which is computed as the convolution of a specific stimulus sequence with an *a priori* Hemodynamic Response Function (HRF), modeling the impulse response of the neurovascular system.

Once the brain activity has been well localized for every contrast of interest, analysis of the HRF shape may help to understand the dynamic of the physiological process in terms of activation delay (time to peak), undershoot, and putative initial dip [9]. Over the last few years, a great deal of attention has been paid to the development of voxel-wise methods for HRF estimation. Recently, a nonlinear differential equation system (i.e., the Balloon model) has been proposed to explain the hemodynamics changes based on the mechanically compelling model of an expandable venous compartment [10] and the standard Windkessel theory [11]. Several studies have proposed a state-space equation framework first to identify the parameters of such nonlinear models and, second, to predict nonlinearities in the BOLD response [12], [13]. Assuming a nonlinear model between the fMRI time course and the stimulus sequence is necessary to account specially for a *refractory* period that may occur when the interstimulus interval (ISI) is shorter than 1 s. As shown in [12], other hemodynamic nonlinearities may also

S. Makni, P. Ciuciu, and J.-B. Poline are with Service Hospitalier Frédéric Joliot (CEA), 91406 Orsay, Cedex, France and with IFR 49, Paris, France (e-mail: name@shfj.cea.fr).

J. Idier is with Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN/CNRS), 44321 Nantes, Cedex 3, France (e-mail: Jerome.Idier@irccyn.ec-nantes.fr).

Digital Object Identifier 10.1109/TSP.2005.853303

appear at about 8 sec poststimulus, which are attributed to a high deoxyhemoglobin concentration during the flow undershoot following the first stimulus. Nonetheless, most of fMRI studies rely on an experimental design that makes the exploration of BOLD nonlinearities unreachable. In such cases, linear modeling constitutes a simpler and satisfactory approximation of the BOLD dynamics, especially when the ISI is more than 2 s.

Pioneering works have embedded parametric models of the HRF shape in a linear framework [14]–[16], [21]. Fitting the parameters in the least square sense has provided an easy way to get the dynamic of the response [16], [17], [21]. Nonetheless, fMRI modeling requires flexibility since the HRF may vary from region to region, task to task, and subject to subject [18]–[20]. Flexibility in parametric linear models has been achieved through the use of a function basis [21], [22]. However, the function basis set imposes a hard constraint on the detectable HRF shapes. To overcome these problems, more adaptive parametric models have been introduced first for fixed epoch fMRI experiments [23], [24]. Recently, these works have been extended to any kind of fMRI experiment in [25]. In this contribution, Woolrich and colleagues have generalized the HRF parameterization of [23] using a half-cosine basis, allowing isolation of different shape characteristics of the HRF. Other contributors have considered nonparametric formulations [26]–[29] in the Bayesian setting and, more specifically, in a state-space framework [30], [31]. All these studies impose a temporal constraint on the regularity of the HRF since the underlying physiological process is slow-varying in time.

A strong limitation that arises in the above-mentioned methods is their lack of robustness, mainly due to the low signal-to-noise ratio (SNR). HRF estimation turns out to be reliable only in high SNR voxels. To increase this ratio, one solution consists of pooling voxels. However, this procedure still requires to select voxels with high SNR values. This can be achieved using testing procedures with constraint models for HRF. On the other hand, two noticeable features have been pointed out [28], [29]. First, for a given stimulus type, the shape of the HRF tends to be spatially homogeneous. Second, the fluctuation of the HRF estimate from one condition to another in a given voxel essentially relies on a *magnitude modulation*. To account for these features in an efficient and robust way, we have derived a region-based formulation of the HRF estimation problem in [32]. In this latter work, we have characterized the vasculature of a given region of interest (ROI) by a single HRF shape but allowed for variation in the magnitude of the response. We have also introduced stimulus-dependent magnitude coefficients for each voxel of a ROI to model the *space-varying* response level to a given stimulus type. This coefficient defines a pseudo-neural response and may better represent the actual "neural" response level (NRL), as seen by deconvolution of the BOLD response [13]. A noticeable limitation of [32] is that the average activity is computed from all the voxels in a ROI. However, for a given region, some voxels may not be activated for one or several conditions. Therefore, in this paper, we develop a new formulation that allows us to perform the detection and estimation steps at the same time. As explained further, this model, which is embodied in a Bayesian approach, is able to segregate activated voxels within a region from nonactivated ones.

Following [27] and [28], the HRF is assumed to be a slowly time-varying function using a Gaussian prior distribution. The NRLs are supposed to be statistically independent from one stimulus type to another. In this work, no spatial correlation is introduced between pseudo-neural responses in nearby voxels because such a model should be based on measured activity, and until now, it is not clear what correlation will be observed. In any case, such a spatial model would preferably be considered on cortical geodesic distance rather than on voxel Euclidean distance. Akin to [33] and [34], in order to solve for the detection of activated voxels in response to a given condition, we consider a two-class Gaussian mixture as a prior probability density function (pdf) on the NRLs. For any condition of interest, the role of the mixture is to efficiently discriminate activating areas from not activating ones in the ROI. Modeling the deactivation process is beyond the scope of this paper, but it could be addressed in the future using a third class in the mixture (see, for instance, [33]).

All parameters of interest (the HRF and the NRLs), as well as hyperparameters, which govern the prior laws, and space-varying parameters (the deterministic trend due to physiological artifacts and the noise variances) should be estimated in an appropriate way. To this end, a Bayesian joint detection-estimation approach is proposed, as explained in the following.

The rest of the paper is organized as follows. Section II revisits the LTI voxelwise modeling of the HRF proposed in [28]. Some notations that will be used throughout the paper are also introduced in this section. Section III focuses on the modeling of fMRI signal across voxels in a given ROI. In Section IV, we present how detection-estimation can be jointly performed. Section V provides the main steps of our Gibbs sampling algorithm. Section VI illustrates the performances of our approach on synthetic data in comparison with previous works [27], [28]. In Section VII, the behavior of our method is analyzed on experimental fMRI data. In Section VIII, we discuss the possible extensions of the proposed method.

II. Voxelwise Formulation

A. Notations

The notations used in this paper are summarized in Table I. In what follows, all vectors are considered as columns by default. Moreover, we will use the notation $\ell = m : n$ as a shorthand for $\ell = m, \ldots, n$.

B. LTI System

Let us define $\boldsymbol{y}_j = (y_{j,t_n})_{n=1:N}$ as the BOLD fMRI time course measured in voxel V_j at (not necessarily uniformly sampled) times t_n, and define $(x_{t_n}^m)_{n=1:N}$ as the corresponding binary vector for the mth condition: $x_{t_n}^m = 1$ if t_n is an onset (i.e., an arrival time) for the mth condition. In classical voxel-dependent approaches [28], [35], a convolution model is assumed between the stimuli and the data

$$y_{j,t_n} = \sum_{m=1}^{M} \sum_{d=0}^{D} h_{j,d\Delta t}^m x_{t_n-d\Delta t}^m + \sum_{q=1}^{Q} p_{t_n,q} \ell_{q,j} + b_{j,t_n}.$$

TABLE I
LIST OF NOTATIONS

$\boldsymbol{y}_j \in \mathbb{R}^N$	Data fMRI time series acquired in voxel V_j
$\boldsymbol{y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_J]$	Matrix of size $N \times J$ representing fMRI time series acquired in all voxels of the ROI
$\boldsymbol{b}_j \in \mathbb{R}^N$	Gaussian noise vector in voxel V_j
$D+1$	Number of HRF coefficients
$\boldsymbol{X}^m$	Binary "onsets matrix" of the mth condition of size $N \times (D+1)$
$\boldsymbol{h}$	HRF to be estimated
M	Number of different conditions in the experiment
a_j^m	NRL for voxel V_j and condition m
Q	Number of drift parameters
$\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_Q]$	Low frequency orthogonal matrix of size $N \times Q$
$\boldsymbol{\ell}_j \in \mathbb{R}^Q$	Nuisance parameter vector for voxel V_j
ϵ_j^2	Noise variance for voxel V_j
$\boldsymbol{\lambda} = (\lambda_{1,m})_{m=1:M}$	Vector of the weighting probabilities
$\mu_{1,m}$	Mean value of NRLs for activated voxels in condition m
$\sigma_{1,m}^2$	Variance of NRLs for activated voxels in condition m
$\sigma_{0,m}^2$	Variance of NRLs for nonactivated voxels in condition m
$C_{1,m}$	$J_{1,m}$-dimensional vector of activated voxel indexes in condition m
$C_{0,m}$	$J_{0,m}$-dimensional vector of nonactivated voxel indexes in condition m

Vector $\boldsymbol{h}_j^m = (h_{j,d\Delta t}^m)_{d=0:D}$ represents the unknown HRF in voxel V_j corresponding to the mth condition. In matrix form, this relation reads

$$\boldsymbol{y}_j = \sum_{m=1}^{M} \boldsymbol{X}^m \boldsymbol{h}_j^m + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j \tag{1}$$

where $\boldsymbol{X}^m = (x_{t_n-d\Delta t}^m)_{n=1:N,d=0:D}$ is a $N \times (D+1)$ binary matrix corresponding to the arrival times for the mth condition. Equation (1) is the sum of three terms.

- The first one represents the stimulus-induced BOLD signal. More precisely, each term $\boldsymbol{X}^m \boldsymbol{h}_j^m$ represents the activation response in voxel V_j to condition m.
- $\boldsymbol{P}\boldsymbol{\ell}_j$ is the confounds part (deterministic trends).
- $\boldsymbol{b}_j$ models the noise term.

The goal of our method is to extract the first term from the other two. In the next paragraphs, we describe our assumptions for each of these two terms.

C. Drift Model

In neuroimaging experiments, the fMRI data are contaminated by a low-frequency drift, mainly due to physiological artifacts [36]. Breathing and cardiac pulses are aliased since the sampling frequency of the data is below Nyquist's bound. A highpass filter is generally used to remove these trends before estimating the HRF. These baseline techniques are merely preprocessing steps to eliminate drifts in the fMRI data. A relevant alternative was also proposed in a semi-parametric framework [37] and is adopted in this paper to model the trend. This approach relies on matrix $\boldsymbol{P} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_Q]$ (Q depends on the lowest frequency attributable to the drift term), which consists of an orthonormal basis of functions $\boldsymbol{p}_q = (p_q(t_n))_{n=1:N}$. To each voxel is attached an unknown weighting vector $\boldsymbol{\ell}_j$.

D. Noise Model

Several temporal noise models may be considered. For a given ROI $\mathcal{R}$, the simplest noise model is a zero-mean Gaussian white process $\boldsymbol{b}$ of unknown variance ϵ^2, independent of $\boldsymbol{h}$ ($\boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0}, \epsilon^2 \boldsymbol{I})$). Nonetheless, it is well known that fMRI time series are correlated in time [38]. Several authors have proposed to estimate the temporal covariance matrix of the noise using an autoregressive model [39], [40]. In [41], a *spatially varying first-order* AR model has been considered. In the present paper, we will only consider a *spatially varying* white noise model since it was actually demonstrated in [42] that various noise correlation structures have little influence on the performances of the HRF estimation. This means that a specific noise variance ϵ_j^2 is assigned to each voxel V_j, allowing for spatially varying artifacts (such as the partial volume effect) to be treated in an appropriate way. Hence, we will need to estimate vector $\boldsymbol{\epsilon}^2 = (\epsilon_j^2)_{j=1:J}$ of noise variances.

III. PROPOSED REGIONAL MODEL

A. Motivation

In [28], the authors have described a method for voxel-specific HRF estimation or, more generally, from a given time series in the context of event-related paradigms. This approach only focused on the temporal aspects, i.e., spatial features were not considered. Results of this method on actual fMRI data showed that HRF estimates, which are computed in neighbor voxels, have approximately the same shape up to a magnitude factor. This property of shape similarity was also observed, in numerous cases, for HRF estimates computed in the same voxel but for different stimulus types [28], [29]. However, in some experimental paradigms such as repetition-suppression ones, because of the habituation phenomenon, estimated HRFs in neighbor voxels may show differences with respect to (w.r.t.) the activation delay or the peak intensity.

Brain regions activated by a given stimulus usually spread over a number of contiguous voxels. A study of the response over an homogeneous region seems therefore reasonable. This region could be defined anatomically or functionally by some brain parcellation technique [43] or clustering algorithm [44].
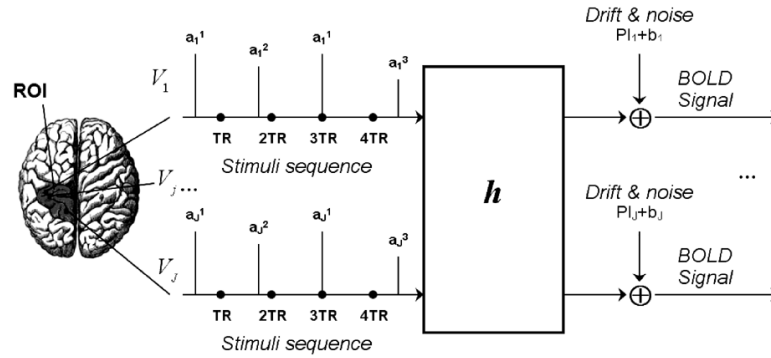
Fig. 1.   ROI-based model. $\boldsymbol{h}$ is the single HRF for all the voxels of the ROI. a_j^m defines the NRL for voxel V_j and stimulus type (or condition) m.

B. Assumptions

($\mathcal{H}1$) We consider a LTI model.

($\mathcal{H}2$) We assume that the fMRI time series $\boldsymbol{y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_J]$ are statistically independent and identically distributed (i.i.d.) in space. We also consider that there is no noise correlation in space.

($\mathcal{H}3$) We do not model any spatial correlation between the NRLs of neighbor voxels. In the recent litterature, several contributions have introduced spatial priors either on the *activation height* [24], [45] or on the labels of the classes associated with the mixture model [46], [47]. This latter idea seems relevant to remove isolated false positives since it encodes the prior belief that neighboring voxels are likely to belong to the same class. We could also extend our model to spatial mixtures that should be preferably defined on the cortical surface using a two-dimensional (2-D) Markov random field.

($\mathcal{H}4$) Generally, fMRI experiments consist of several conditions (visual, auditory, etc.). We assume that the responses to different stimuli combine in a *linear* way.

C. Problem Statement

Once an homogeneous ROI is defined, our purpose is to estimate its *canonical* time response. More precisely, we characterize the ROI by a single HRF shape and a magnitude coefficient for each voxel and stimulus type. It is likely that this coefficient may better represent the NRL. To this end, we introduce a special case of (1) that accounts for stimulus/voxel-dependent signal and voxel-dependent noise levels but assumes a single HRF shape over the region. Letting $\mathcal{R} = (V_j)_{j=1:J}$ be the ROI, then (1) reads

$$\boldsymbol{y}_j = \sum_{m=1}^{M} a_j^m \boldsymbol{X}^m \boldsymbol{h} + \boldsymbol{P}\boldsymbol{\ell}_j + \boldsymbol{b}_j, \quad \forall j = 1 : J \qquad (2)$$

where a_j^m stands for the NRL in voxel V_j for condition m. Fig. 1 summarizes the main features of model (2).

IV. *A Priori* Model and the Detection/Estimation Problem

Assuming that a given region has homogeneous vasculature properties, we propose to estimate the corresponding HRF and to simultaneously classify voxels in the ROI as either activating or not activating. This classification task has already been applied to fMRI data either using a nonspatial mixture model [34] or a spatial one [33], [46], [47]. Following [34], we introduce a two-class mixture model: one for activated voxels (class 1) and the other for nonactivated ones (class 0). In this framework, we need to estimate the label attached to each voxel of the ROI and the corresponding NRL. Since this model stands for a given stimulus type, we consider M two-class mixtures, each of them being coupled to a single stimulus type.

The estimation of the NRLs and the HRF may be thought of as a multichannel, semi-blind deconvolution problem. Indeed, the different channels correspond to the available voxels and "semi-blind" refers to the prior knowledge of the arrival times of the "neural" responses. In contrast, arrival times have to be identified as well in some other applications of sparse spike deconvolution such as geophysics [48]. For all parameters of interest, the respective *a priori* models are described in the following.

A. HRF

As physiologically advocated in [49], the HRF can be characterized as follows [28], [35]:

i) Its temporal variations are smooth. Quantification is achieved by setting Gaussian prior for the norm of the second derivative of the HRF, whose variance is adjusted by hyperparameter $\sigma_{\boldsymbol{h}}^2$. More precisely, we assume $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \|\partial^2 \boldsymbol{h}\|^2 / (2\sigma_{\boldsymbol{h}}^2))$. Following usual practice, $\|\partial^2 \boldsymbol{h}\|^2$ can be discretized as

$$(\partial^2 \boldsymbol{h})_{d\Delta t} \approx \frac{h_{(d+1)\Delta t} - 2h_{d\Delta t} + h_{(d-1)\Delta t}}{(\Delta t)^2}$$
$$\forall d = 1 : D - 1.$$

Taking into account (*i*), we obtain in matrix form $\partial^2 \boldsymbol{h} = \boldsymbol{D}_2 \boldsymbol{h}$, where $\boldsymbol{D}_2$ is the *truncated* second-order finite difference matrix of size $(D - 1) \times (D - 1)$ that depends on Δt [35]. Hence, $\|\partial^2 \boldsymbol{h}\|^2 = \boldsymbol{h}^{\mathrm{t}} \boldsymbol{R}^{-1} \boldsymbol{h}$, where $\boldsymbol{R} = (\boldsymbol{D}_2^{\mathrm{t}} \boldsymbol{D}_2)^{-1}$ is symmetrical positive definitive.

ii) It is causal, and its amplitude vanishes at the first and the end time points ($h_0 = h_{D\Delta t} = 0$), which corresponds to a duration of about 25 s.

Furthermore, akin to [50], we impose a normalization constraint i.e., $\boldsymbol{h} = 1$, to remove the scale ambiguity encountered in blind deconvolution problems [48].

B. "Neural" Response Levels

According to ($\mathcal{H}$3), we have $p(\boldsymbol{a}; \boldsymbol{\theta_a}) = \prod p(\boldsymbol{a}^m; \boldsymbol{\theta}_m)$ with $\boldsymbol{a} = [\boldsymbol{a}^1, \dots, \boldsymbol{a}^M]$, $\boldsymbol{a}^m = (a_j^m)_{j=1:J}$, and $\boldsymbol{\theta_a} = (\boldsymbol{\theta}_m)_{m=1:M}$, where vector $\boldsymbol{\theta}_m$ gathers the hyperparameters related to the prior pdf of the NRLs for the mth stimulus type. These hyperparameters are unknown and will be further specified.

In some cases, a few voxels of the ROI are activated by a given condition, whereas others may not. We therefore consider couples of random variables $z_j^m = (q_j^m, a_j^m), j = 1 : J, m = 1 : M$. Parameter q_j^m is a binary random variable that indicates whether voxel V_j is activated or not by condition m (respectively, $q_j^m = 1$ or $q_j^m = 0$). Conditional on q_j^m, a_j^m is a Gaussian random variable that represents the NRL for voxel V_j and condition m. To model this prior knowledge about the activation process, we introduce a two-class Gaussian mixture prior distribution

$$\lambda_{1,m}\mathcal{N}\left(\mu_{1,m}, \sigma_{1,m}^2\right) + \lambda_{0,m}\mathcal{N}\left(0, \sigma_{0,m}^2\right) \qquad (3)$$

where $\forall j \in \{1, \dots, J\}, \lambda_{1,m} = \Pr(q_j^m = 1)$, and $\lambda_{0,m} = 1 - \lambda_{1,m}$. According to ($\mathcal{H}$3), $\lambda_{1,m}$ is a condition-dependent factor, but it is not voxel-dependent.

As implied by (3), we attribute a Gaussian prior model to the NRLs

$$p\left(a_j^m \mid q_j^m = 1\right) \sim \mathcal{N}\left(\mu_{1,m}, \sigma_{1,m}^2\right)$$
$$p\left(a_j^m \mid q_j^m = 0\right) \sim \mathcal{N}\left(0, \sigma_{0,m}^2\right).$$

Note that we set $\mu_{0,m} = 0$ since this parameter represents the mean of the amplitude levels for nonactivated voxels, which, by definition, should be zero. Thus, four hyperparameters describe this mixture prior model for each condition: $\boldsymbol{\theta}_m = [\lambda_{1,m}, \mu_{1,m}, \sigma_{1,m}^2, \sigma_{0,m}^2]$. Such a model is a generalization of the Gaussian prior distribution for the NRL's advocated in [32].

C. Low-Frequency Drift

Vector $\boldsymbol{\ell} = [\ell_1, \dots, \ell_J]$ gathers the unknown regression coefficients of the function basis $\boldsymbol{P}$. We assume that $\boldsymbol{\ell}$ is a random vector independent of $\boldsymbol{h}$, such that $p(\boldsymbol{\ell}; \boldsymbol{\tau}_l^2) = \prod_j p(\ell_j; \tau_l^2)$ and $\ell_j \sim \mathcal{N}(\boldsymbol{0}, \tau_l^2 \boldsymbol{I}_Q)$. In this paper, the calculations will be provided in the noninformative case, that is, when $\tau_l^2 \to +\infty$. In

this case, the prior pdf becomes flat, which simplifies the expression of the full posterior pdf after integrating the drift parameters $\boldsymbol{\ell}$ out (see Section V).

D. Hyperparameters

The complete set of hyperparameters to be estimated is denoted $\Theta = [\epsilon^2, \sigma_{\boldsymbol{h}}^2, \boldsymbol{\theta_a}]$. For all these parameters, we resort to noninformative Jeffreys priors that meet the *reparametrization-invariance* requirement [51, p. 132]

$$p(\epsilon^2) = \prod_j p\left(\epsilon_j^2\right) = \prod_j \epsilon_j^{-1}, \quad p\left(\sigma_{\boldsymbol{h}}^2\right) = \sigma_{\boldsymbol{h}}^{-1}$$

$$p(\boldsymbol{\theta_a}) = \prod_m p(\boldsymbol{\theta}_m)$$

$$= \prod_m p(\lambda_{1,m}) p\left(\mu_{1,m}, \sigma_{1,m}^2\right) p\left(\sigma_{0,m}^2\right),$$

$$p(\lambda_{1,m}) \propto \mathcal{B}(1/2, 1/2)$$

$$p\left(\mu_{1,m}, \sigma_{1,m}^2\right) = \sigma_{1,m}^{-1}, \quad p\left(\sigma_{0,m}^2\right) = \sigma_{0,m}^{-1}.$$

V. ALGORITHM

A. Likelihood

According to ($\mathcal{H}$1)–($\mathcal{H}$2), the likelihood function reads

$$p(\boldsymbol{y} \mid \boldsymbol{h}, \boldsymbol{a}, \boldsymbol{\ell}, \epsilon^2) = \prod_j p\left(\boldsymbol{y}_j \mid \boldsymbol{h}, \boldsymbol{a}, \boldsymbol{\ell}_j, \epsilon_j^2\right)$$

$$\propto \left(\prod_j \epsilon_j^{-N} \exp\left(-\frac{1}{2\epsilon_j^2}\|\tilde{\boldsymbol{y}}_j - \boldsymbol{P}\boldsymbol{\ell}_j\|^2\right)\right)$$

where $\tilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h}$. Maximum likelihood (ML) estimation of $(\boldsymbol{h}, \boldsymbol{a})$ is a bilinear inverse problem since (2) is linear w.r.t. $\boldsymbol{h}$ when $\boldsymbol{a}$ is fixed and *vice-versa*. In addition, the ML solution $(\boldsymbol{h}^*, \boldsymbol{a}^*)$ is not unique. For instance, every couple $(\boldsymbol{h}^*/C, \boldsymbol{a}^* \times C)$ defines another pair of solutions in the ML sense. By contrast, structural prior information is available both on $\boldsymbol{h}$ and $\boldsymbol{a}$. Therefore, we resort to the Bayesian framework.

B. Joint Posterior Distribution

Considering the constructed model [cf. $\mathcal{H}$1)–$\mathcal{H}$4)] and assuming no further prior dependence between parameters, formal application of the chain rule yields (4), shown at the bottom of the page.

Since (2) is linear w.r.t. vectors $\boldsymbol{\ell}_j$ of nuisance variables, they can be easily integrated out. We therefore focus on the marginal distribution $p(\boldsymbol{h}, \boldsymbol{a}, \Theta \mid \boldsymbol{y}) = \int p(\boldsymbol{h}, \boldsymbol{a}, \boldsymbol{\ell}, \Theta \mid \boldsymbol{y})d\boldsymbol{\ell}$ in order to get

$$p(\boldsymbol{h},\boldsymbol{a},\boldsymbol{\ell},\Theta \mid \boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{h},\boldsymbol{a},\boldsymbol{\ell},\epsilon^2)p(\boldsymbol{h})p(\boldsymbol{\ell})\left(\prod_m \prod_j \sum_{i=0}^1 p(a_j^m|q_j^m = i, \boldsymbol{\theta_a})p(q_j^m = i)\right) p(\epsilon^2)\,p(\sigma_{\boldsymbol{h}}^2)\,p(\boldsymbol{\theta_a})$$

$$\propto \left(\prod_j \epsilon_j^{-N-1}\right)\sigma_{\boldsymbol{h}}^{-D}\exp\left(\frac{\boldsymbol{h}^{\mathbf{t}}R^{-1}\boldsymbol{h}}{2\sigma_{\boldsymbol{h}}^2} - \sum_j \frac{\|\tilde{\boldsymbol{y}}_j - \boldsymbol{P}\boldsymbol{\ell}_j\|^2}{2\epsilon_j^2}\right)\prod_m \left(p(\boldsymbol{\theta}_m)\prod_j \sum_{i=0}^1 \frac{\lambda_{i,m}}{\sigma_{i,m}}\exp\left(\frac{-\left(a_j^m - \mu_{i,m}\right)^2}{2\sigma_{i,m}^2}\right)\right). \qquad (4)$$

fewer parameters to sample. Moreover, due to the Rao-Black-wellization property [52], the marginalization step reduces the variance of the remaining parameters to be sampled and yields

$$
p(\boldsymbol{h}, \boldsymbol{a}, \Theta \,|\, \boldsymbol{y})
$$

$$
\propto \sigma_{\boldsymbol{h}}^{-D} \exp\left(-\frac{1}{2\sigma_{\boldsymbol{h}}^2}\boldsymbol{h}^{\mathbf{t}}R^{-1}\boldsymbol{h} - \frac{1}{2}\sum_j \tilde{\boldsymbol{y}}_j^{\mathbf{t}}\boldsymbol{Q}_j\tilde{\boldsymbol{y}}_j\right)
$$

$$
\times \prod_j \left(\epsilon_j^{Q-N-1}\prod_m\sum_{i=0}^1 \frac{\lambda_{i,m}^2\lambda_{1-i,m}}{\sigma_{i,m}^2\sigma_{1-i,m}}\right.
$$

$$
\times \left. \exp\left(-\frac{(a_j^m - \mu_{i,m})^2}{2\sigma_{i,m}^2}\right)\right)
$$

where $\boldsymbol{Q}_j = \epsilon_j^{-2}(\boldsymbol{I}_N - \boldsymbol{P}\boldsymbol{P}^{\mathbf{t}})$, and the orthogonality of $\boldsymbol{P}$ (i.e., $\boldsymbol{P}^{\mathbf{t}}\boldsymbol{P} = \boldsymbol{I}_N$) has been accounted for. As mentioned earlier in Section IV-C, choosing an informative prior for $\boldsymbol{\ell}$ would have modified the term $\prod_j \epsilon_j^{Q-N-1}$ in $p(\boldsymbol{h}, \boldsymbol{a}, \Theta \,|\, \boldsymbol{y})$, making its sampling w.r.t. noise variances more tricky.

Due to the bilinearity of model (2) w.r.t. $\boldsymbol{h}$ and $\boldsymbol{a}$, we do not consider the maximum *a posteriori* (MAP) estimate of $(\boldsymbol{h}, \boldsymbol{a})$. Indeed, its computation would require stochastic optimization since $p(\boldsymbol{h}, \boldsymbol{a}, \Theta \,|\, \boldsymbol{y})$ may possess many local maxima. "Joint MAP" estimation has already been addressed in blind deconvolution problems with fixed hyperparameters [53], [54], but the proposed method suffers from shortcomings w.r.t. local minima and nonuniqueness of the solution. More importantly, the issue of hyperparameter tuning would not be easily addressed using the MAP approach. In fact, it was shown that simultaneous signal estimation and hyperparameter identification is achieved in this framework by maximizing the generalized likelihood whose maximizer is not statistically consistent [55]. Considering all aspects of this problem, we rather choose to compute an estimate of the posterior mean of the unknown parameters as well as the hyperparameters. Direct sampling from the joint posterior pdf is impossible. Therefore, we resort to Monte Carlo Markov chain methods [52] and, more specifically, to Gibbs sampling for which the full conditional pdfs needs to be derived.

C. Numerical Inference Using Gibbs Sampling

Gibbs sampling has already been applied in the context of blind deconvolution problems [48]. It consists of starting with a seed vector and sequentially modifying one (scalar or multidimensional) component at a time by sampling according to the full conditional pdf of that component given the remaining variables (denoted as rest in the following) and the data. Samples are composed of the set of all vectors whose components have been updated an equal amount of iterations.

A key issue with Gibbs sampling is to partition the vector of parameters into blocks whose full conditional sampling can be performed easily. In our case, we exploit the bilinearity of model (2) w.r.t. $\boldsymbol{h}$ and $\boldsymbol{a}$ that makes the full conditional pdfs $p(\boldsymbol{h} \,|\, \boldsymbol{y}, \boldsymbol{a}, \Theta)$ and $p(\boldsymbol{a} \,|\, \boldsymbol{y}, \boldsymbol{h}, \Theta)$ Gaussian, given the noise model and the priors. Then, the derivation of the full conditional pdfs corresponding to the chosen clustering is straightforward. As detailed in Table II, the updating

TABLE II
STEPS OF THE GIBBS SAMPLING ALGORITHM

- Setting up: choose $\boldsymbol{h}^0, \boldsymbol{a}^0, \boldsymbol{\lambda}^0, (\epsilon^2)^0, \boldsymbol{\theta}_{\boldsymbol{a}}^0$.
- Iteration k: draw samples $\boldsymbol{h}^k, \boldsymbol{a}^k, \boldsymbol{\lambda}^k, (\epsilon^2)^k, \boldsymbol{\theta}_{\boldsymbol{a}}^k$ from the full conditional posterior pdfs:
  - HRF: $\boldsymbol{h}^k \sim p(\boldsymbol{h}^k \,|\, \boldsymbol{y}, \boldsymbol{a}^{k-1}, (\epsilon^2)^{k-1})$, *cf.* Eq. (7).
  - HRF scale: $(\sigma_{\boldsymbol{h}}^2)^k \sim p(\sigma_{\boldsymbol{h}}^2)^k \,|\, \boldsymbol{h}^k)$, *cf.* Eq. (8).
  - NRLs: for every condition m and every voxel j,
    * $(u_j^m)^k \sim \mathcal{U}[0,1]$; if $(u_{1,j}^m)^k \leq (\lambda_{1,j}^m)^k$, then $(q_j^m)^k = 1$ otherwise $(q_j^m)^k = 0$.
    * $(a_j^m)^k \,|\, (q_j^m)^k = i \sim \mathcal{N}((\mu_{i,j}^m)^k, ((\sigma_{i,j}^m)^2)^k)$, $(i = 0, 1)$, *cf.* Eq. (9).
  - Noise variances: for every voxel j, $(\epsilon_j^2)^k \sim p((\epsilon_j^2)^k \,|\, \boldsymbol{y}_j, \boldsymbol{h}^k, \boldsymbol{a}_j^k)$, *cf.* Eq. (11).
  - Mixture parameters: for every condition m,
    * Weighting probability: $\lambda_{1,m}^k \sim p(\lambda_{1,m}^k \,|\, \boldsymbol{y}, \boldsymbol{h}^k, \boldsymbol{a}^k, (\epsilon^2)^k)$, *cf.* Eq. (12).
    * Variance of NRLs for nonactivated voxels: $(\sigma_{0,m}^2)^k \sim p((\sigma_{0,m}^2)^k \,|\, \eta_{0,m}^k, \nu_{0,m}^k)$, *cf.* Eq. (13).
    * Variance of NRLs for activated voxels: $(\sigma_{1,m}^2)^k \sim p((\sigma_{1,m}^2)^k \,|\, \eta_{1,m}^k, \nu_{1,m}^k)$, *cf.* Eq. (14).
    * Mean of NRLs for activated voxels: $(\mu_{1,m})^k \sim p((\mu_{1,m})^k \,|\, (\sigma_{1,m}^2)^k, \eta_{1,m}^k, \nu_{1,m}^k)$, *cf.* Eq. (15).
- Iterate until maximum iteration number K_0 is achieved. Finally, we compute posterior mean estimates using the following approximation (Samples $\theta^1, \ldots, \theta^I$ are discarded: they correspond to the so-called *burn-in* period):

$$
\widehat{\theta}^{\mathrm{PM}} = \frac{1}{K_1}\sum_{k=I}^{K_0}\theta^k, \quad K_1 = K_0 - I + 1, \quad \forall \theta \in \{\boldsymbol{h}, \boldsymbol{a}, \Theta\}
$$

steps are performed on the following unknown variables: $\boldsymbol{h}, \sigma_{\boldsymbol{h}}^2, z_j^m = (q_j^m, a_j^m), \epsilon_j^2, \lambda_m, \sigma_{0,m}^2, \sigma_{1,m}^2$, and $\mu_{1,m}$. We therefore need access to the corresponding full conditional pdfs whose closed forms are given in Appendix A.

Once the Gibbs sampler has converged, these quantities are approximated by their sample counterparts, as emphasized in Table II. To check the convergence of our algorithm, we have plotted the values of several scalar parameters (a noise variance as well as other hyperparameters) w.r.t. iterations. Once the relative norm of these parameters is lower than a given threshold ($\alpha = 10^{-3}$), it was decided that the sampling scheme could stop. Following [29], we could implement more sophisticated convergence diagnosis if necessary. However, in practice, we did not experience convergence problems.

VI. SIMULATION RESULTS

A. Generation of Synthetic Datasets

We simulated a *random-intermixed* sequence of indexes coding for two different stimulus types ($M = 2$). Each index m corresponded to a specific stimulus. The timing of the trials was random since the retained inter-stimulus intervals (ISIs) were uniformly distributed on [1.5, 2.5] sec. For such ISI values, the linear approximation of the response is valid. For the simulation, we also took one session of $N = 100$ time points (e.g., scan number). The sampling interval δt of the trial onsets was set to 0.5 s. The onsets arrival times are put in the $n\delta t$ grid by moving them to the nearer time points on this grid.

The ROI $\mathcal{R}$ consisted of $J = 60$ voxels. As indicated in Table I, $C_{1,m}$ and $C_{0,m} = \mathcal{R}\backslash C_{1,m}$ are the sets of activated and nonactivated voxels in $\mathcal{R}$ for condition m, respectively. For
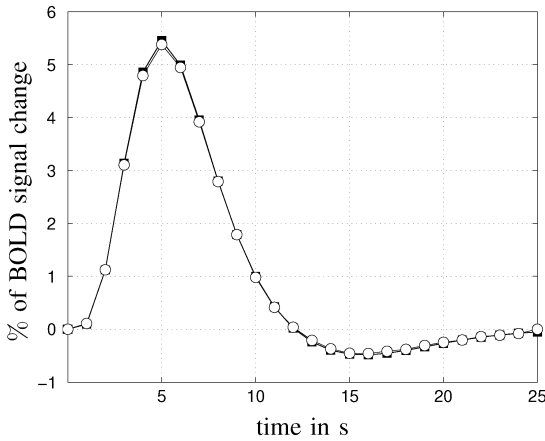
Fig. 2. Symbols ■ and ○ represent the true $\boldsymbol{h}_c$ and the corresponding HRF estimate, respectively.

the two conditions, we chose the following distributions for the NRLs:

$$\boldsymbol{a}_{j \in C_{1,1}}^1 \sim \mathcal{N}\left(\mu_{1,1} = 10, \sigma_{1,1}^2 = 3\right)$$
$$\boldsymbol{a}_{j \in C_{0,1}}^1 \sim \mathcal{N}\left(0, \sigma_{0,1}^2 = 1\right)$$
$$\boldsymbol{a}_{j \in C_{1,2}}^2 \sim \mathcal{N}\left(\mu_{1,2} = 2, \sigma_{1,2}^2 = 0.3\right)$$
$$\boldsymbol{a}_{j \in C_{0,2}}^2 \sim \mathcal{N}\left(0, \sigma_{0,2}^2 = 0.4\right)$$

For Condition 1, there were $J_{1,m} = 24$ activated voxels and $J_{0,m} = 38$ nonactivated ones. For condition 2, there were as many as activated voxels as nonactivated ones ($J_{0,m} = J_{1,m} = 30$). For all voxels, the binary stimulus sequence was convolved with the canonical HRF $\boldsymbol{h}_c$,[1] whose exact shape appears in Fig. 2 as the ■-line. A white Gaussian noise $\boldsymbol{b}_j$ was then added to the stimulus-induced signal $\sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h}$ in every voxel V_j. The noise variances $(\epsilon_j^2)_{j=1:J}$ were set as follows. Let us introduce the definition of space-varying contrast-to-noise ratio (CNR), based on the L_1-norm of the signal:

$$\text{CNR}_j \triangleq \frac{\sum_{d,m} \left| a_j^m h_{d\Delta t} \right|}{(D-1)\epsilon_j}$$

We restricted our simulations to constant CNR_j in space: $\text{CNR}_j = \text{CNR}, \forall j$. Since the NRLs were sampled from a target distribution (e.g., $p(\boldsymbol{a}_{j \in C_{i,m}}^m)$ for condition m and class i), the signal norm was different from one voxel to another. To achieve constant CNR, we adapted the noise variances ϵ_j accordingly. Two values of CNR were investigated (CNR = 1.3 and CNR = 0.3), leading first to a simpler synthetic case and, second, to a more realistic one.

Space-varying low-frequency drifts $\boldsymbol{P}\boldsymbol{\ell}_j$ were added to the fMRI time courses. They were generated from a cosine transform basis whose coefficients $\boldsymbol{\ell}_j$ were drawn from a normal distribution. The amount of low-frequency signal was tuned to be significant: We checked that the ratio between the quadratic norm of the drift components $\boldsymbol{P}\boldsymbol{\ell}_j$ and the quadratic norm of the drift-free data $\sum_m a_j^m \boldsymbol{X}^m \boldsymbol{h} + \boldsymbol{b}_j$ was no less than 50%. The trials of the two conditions were well distributed over time such that collinearity with the low-frequency signal is unlikely.

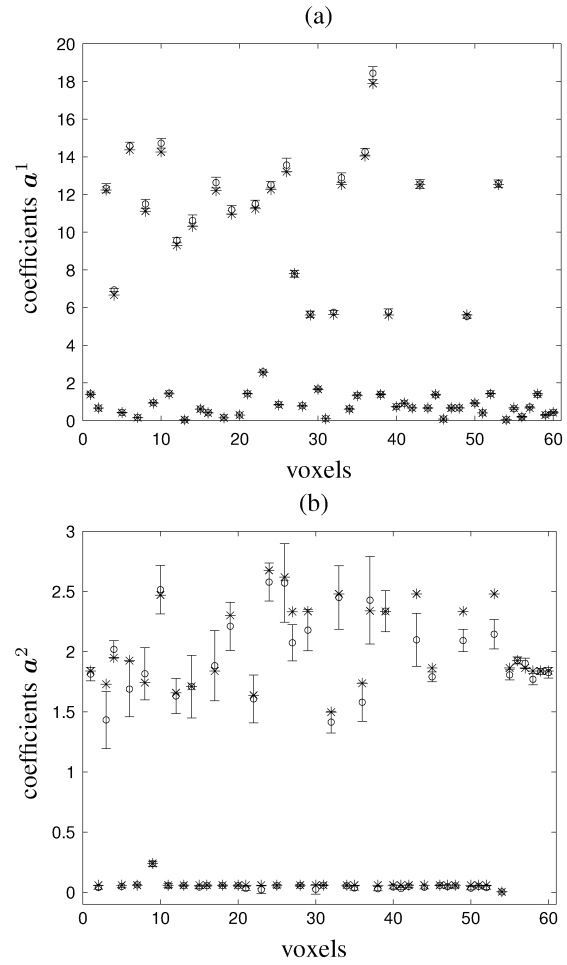[1]used in the SPM2 software www.fil.ion.ucl.ac.uk/spm/spm2.html

Fig. 3. (a) and (b) provide the NRLs for the first ($m = 1$) and second ($m = 2$) conditions, respectively. Symbols * and ○ represent true and estimated NRLs, respectively. The error bars are derived from the sampled posterior variances given by (5).

The chosen cut-off period (COP) related to the drift term was 70 s. corresponding to $Q = 4$. Synthetic data were then obtained after undersampling the sequences at a $[\text{TR}/\delta t]$ rate, the inter-scan interval being TR = 2 s.

Akin to [28], the sampling rate of the HRF estimate was set to $\Delta t = \delta t$ to avoid estimation bias due to instant-matching error.

B. High CNR ($CNR = 1.3$)

As shown in Fig. 2, the HRF estimate matches the canonical time course $\boldsymbol{h}_c$ very well.

Fig. 3(a) and (b) shows the NRL estimates for both conditions. Since the simulated average activity is stronger for $m = 1(\mu_{1,1} = 10)$, Fig. 3(a) demonstrates that the response magnitudes of activated voxels are estimated accurately. Indeed, the error bars v_j^m attached to the estimated NRLs and computed as the sampled variances of the NRLs

$$v_j^m = \frac{1}{K_1} \sum_{k=I}^{K_0} \left(\left(a_j^m\right)^k - \bar{a}_j^m \right)^2 \quad \forall j, m, \text{with}$$

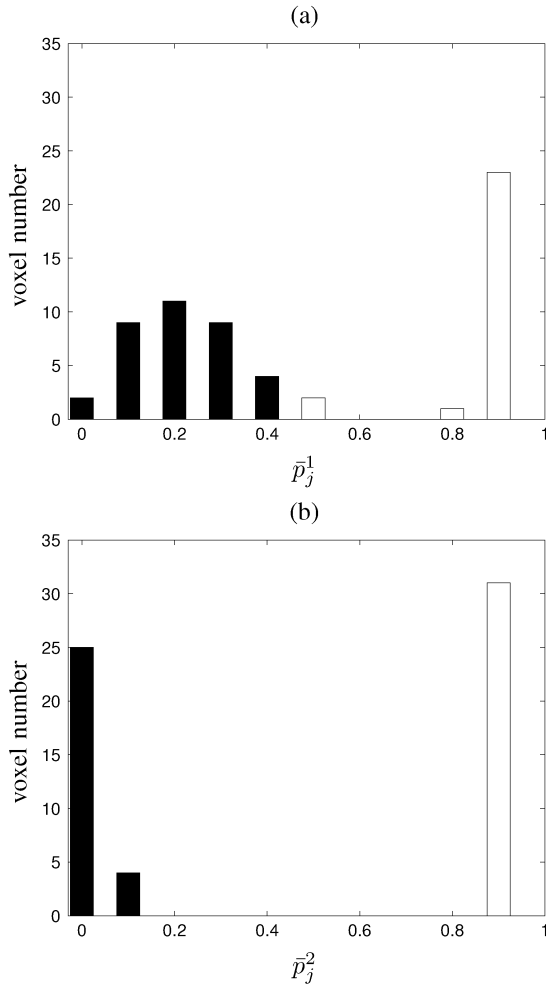$$\bar{a}_j^m = \frac{1}{K_1} \sum_{k=I}^{K_0} \left(a_j^m\right)^k \tag{5}$$

(a)

(b)

Fig. 4. (a) and (b) describe the histograms of activation probabilities $\bar{p}_j^m$ for $m = 1, 2$, respectively. Filled bars represent nonactivated voxels belonging to $C_{0,m}(\bar{p}_j^m < 0.5)$, whereas white bars represent activated voxels (i.e., belonging to $C_{1,m}$).

Fig. 5. Symbols ■ and ○ represent true $\boldsymbol{h}_c$ and estimated HRF $\boldsymbol{h}$, respectively.

because of $\sigma_{0,1}^2 = 1$. The middle part of this graph shows that $\bar{p}_j^1 = 0.5$ for two voxels, indicating that our algorithm hesitates to classify these voxels in class $C_{0,1}$. The right part of the histogram demonstrates that 24 voxels have been assigned to class $C_{1,1}$, as expected. Fig. 4(b) corresponds to an easier case where activated and nonactivated voxels for the second condition are clearly separated and well classified (30 voxels within each class). This results from our variance choice $\sigma_{0,2}^2 = 0.4$.

C. Low CNR $(CNR = 0.3)$

The HRF estimate plotted in Fig. 5 is a slightly oversmoothed version of $\boldsymbol{h}_c$ about its peak because of the regularization constraint.

Fig. 6(a) and (b) shows the NRL estimates for both conditions. For activated voxels, the estimated magnitudes are more biased for lower CNR. The noise level also has an influence on the confidence intervals, which appear broader for both conditions but, more significantly, for the second one. In contrast, the results for nonactivated voxels show no significant difference in comparison with the first simulation.

Fig. 7 shows histograms of activation probabilities for both conditions. It appears that lower CNR values induce detection of false positives for condition 1, whereas we found 30 voxels in both classes for condition 2. As expected, the number of misclassified voxels is larger for the first condition because of $\sigma_{0,1}^2 = 1$. In Fig. 7(b), it appears that the distribution of nonactivated voxels has a broader dispersion for this CNR in comparison with the first simulation.

VII. RESULTS ON REAL fMRI DATA

A. MRI Parameters

The experiment was performed on a 3-T whole-body system (Bruker, Germany) equipped with a quadrature birdcage radio frequency (RF) coil and a head-gradient coil insert designed for echoplanar imaging. Functional images were obtained with a T2*-weighted gradient echo, echo planar imaging sequence (TR = 2.4 s, TE = 30 ms). A 3-D volume is composed of $64 \times 64 \times 32$ voxels. A high-resolution ($1 \times 1 \times 1.2$ mm) anatomical image using a 3-D gradient-echo inversion-recovery sequence was also acquired for each subject.

are small enough to embed the true magnitudes depicted in (∗). In contrast, Fig. 3(b) illustrates the influence of a lower mean activity ($\mu_{1,2} = 2$) both on the estimated NRLs and on the corresponding error bars: A small amount of bias, as well as larger sampled variances v_j^m, are observed because of a lower CNR for the second condition. On the other hand, our method behaves normally for nonactivated voxels since estimated and true NRLs overlap.

Fig. 4 shows histograms of activation probabilities $\bar{p}_j^m$ that are computed by averaging the labels q_j^m over iterations

$$\bar{p}_j^m = \frac{1}{K_1} \sum_{k=I}^{K_0} \left(q_j^m\right)^k, \quad \forall j, m. \tag{6}$$

Each bar depicted in Fig. 4(a) and (b) returns the number of voxels corresponding to a given value of $\bar{p}_j^1$ and $\bar{p}_j^2$, respectively. Filled bars show nonactivated voxels (i.e., belonging to $C_{0,m}$ because of $(\hat{q}_j^m)^{\mathrm{PM}} = 0$) for which $\bar{p}_j^m < 0.5$, whereas white bars represent activated voxels belonging to $C_{1,m}$.

As expected, the histogram depicted in Fig. 4(a) for nonactivated voxels spreads over the left part of the probability axis
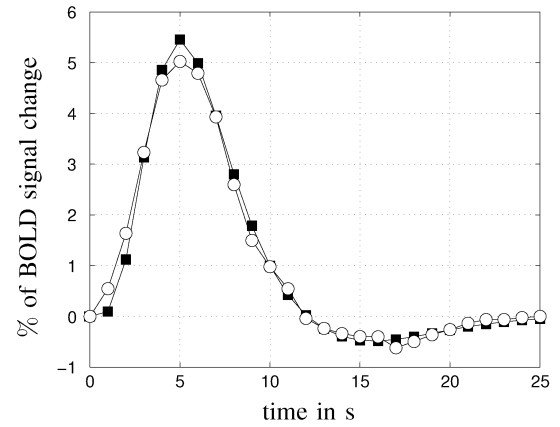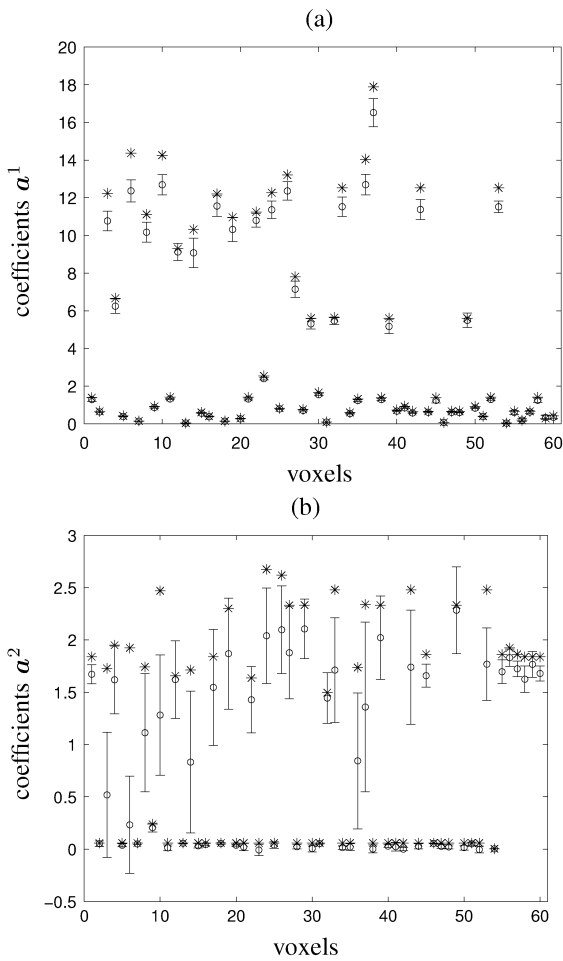
Fig. 6. (a) and (b) provide the NRLs for the first ($m = 1$) and second ($m = 2$) conditions, respectively. Symbols $*$ and $\circ$ represent true and estimated NRLs, respectively. The error bars are derived from the sampled posterior variances from (5).
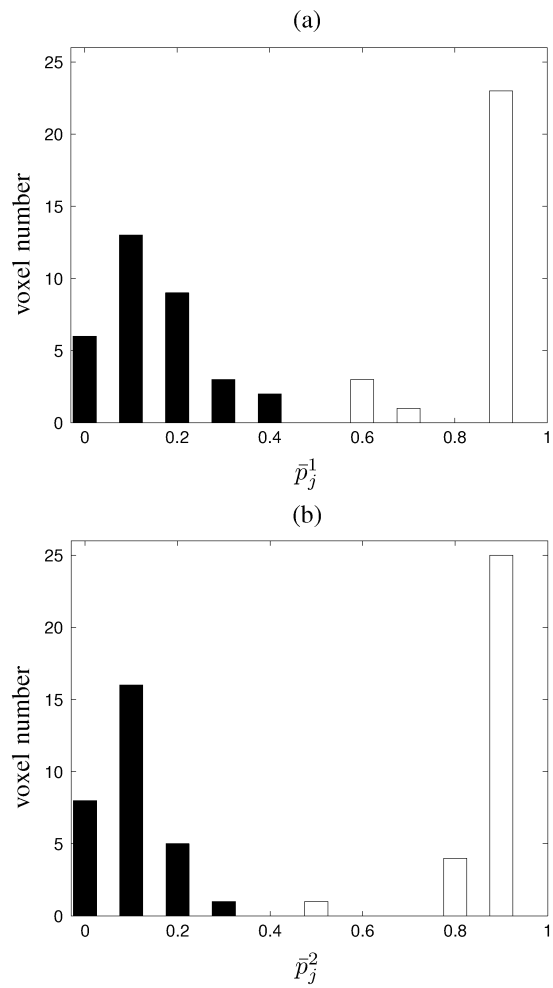

Fig. 7. (a) and (b) describe the histograms of activation probability $\bar{p}_j^m$ for $m = 1, 2$, respectively. Filled bars represent nonactivated voxels belonging to $C_{0,m}(\bar{p}_j^m < 0.5)$, whereas white bars represent activated voxels (i.e., belonging to $C_{1,m}$).

B. Experimental Paradigm

The experiment is a fast event-related paradigm. It consisted of a single session of 125 trials ($N = 125$) lasting TR = 2.4 sec each. The main goal of this experiment was to quickly map several brain functions such as motor, visual, and auditory responses, as well as higher cognitive functions like computation, but here, we will focus only on four stimuli: right-hand button click, left-hand button click, audio, and video.

The chosen ROIs are SPM clusters obtained from t maps (thresholded at $P = 0.001$ corrected for multiple comparisons) based on standard SPM activation detection using a canonical model, least squares estimation, and inference on relevant contrasts.

In the following, we choose to study two contrasts: (right click minus left click) contrast and (audio minus video) contrast. ROIs $\mathcal{R}_1$ and $\mathcal{R}_2$ were defined using the MARSBAR toolbox.[2] The 178 voxels of the (right click minus left click) $\mathcal{R}_1$ are located around the voxel of Talairach coordinates in millimeters: ($X = -36, Y = -21, Z = 48$). The 632 voxels of the (audio minus video) $\mathcal{R}_2$ are located around the voxel of Talairach coordinates in millimeters: ($X = -63, Y = -21, Z = 3$).

[2]www.sourceforge.net/projects/marsbar

C. Results

In Figs. 8(a) and 9(a), the HRF estimates corresponding to the two selected ROIs ($\mathcal{R}_1$ and $\mathcal{R}_2$) are plotted. These HRF estimates have a regular shape, as enforced by the prior model. Figs. 8(b)–(c) and 9(b)–(c) show the maps of the NRL estimates computed for the right and left clicks in $\mathcal{R}_1$ and for the audio and video conditions in $\mathcal{R}_2$, respectively. Note the different scaling in these figures, and as expected, high values are found for the right click condition in $\mathcal{R}_1$ defined with a (right click minus left click) contrast, whereas positive or negative values close to 0 are found for the left click. The same apply for the (audio minus video) contrast.

Figs. 8(d)–(e) and 9(d)–(e) show the voxels classification in class 0 (red) and in class 1 (white). Classification was based on the maximum likelihood principle (e.g., voxels with greater chance to be in class 1 than in class 0 are classified in 1). It is crucial to note that voxels classified in class 1 [for instance, for the left click in Fig. 8(c)] *are not necessarily "activated"* in the usual sense used in neuroimaging results but can only be said to be more likely in class 1. Whether they are "activated" depends on whether class 1 can be differentiated from class 0. Indeed, as shown in
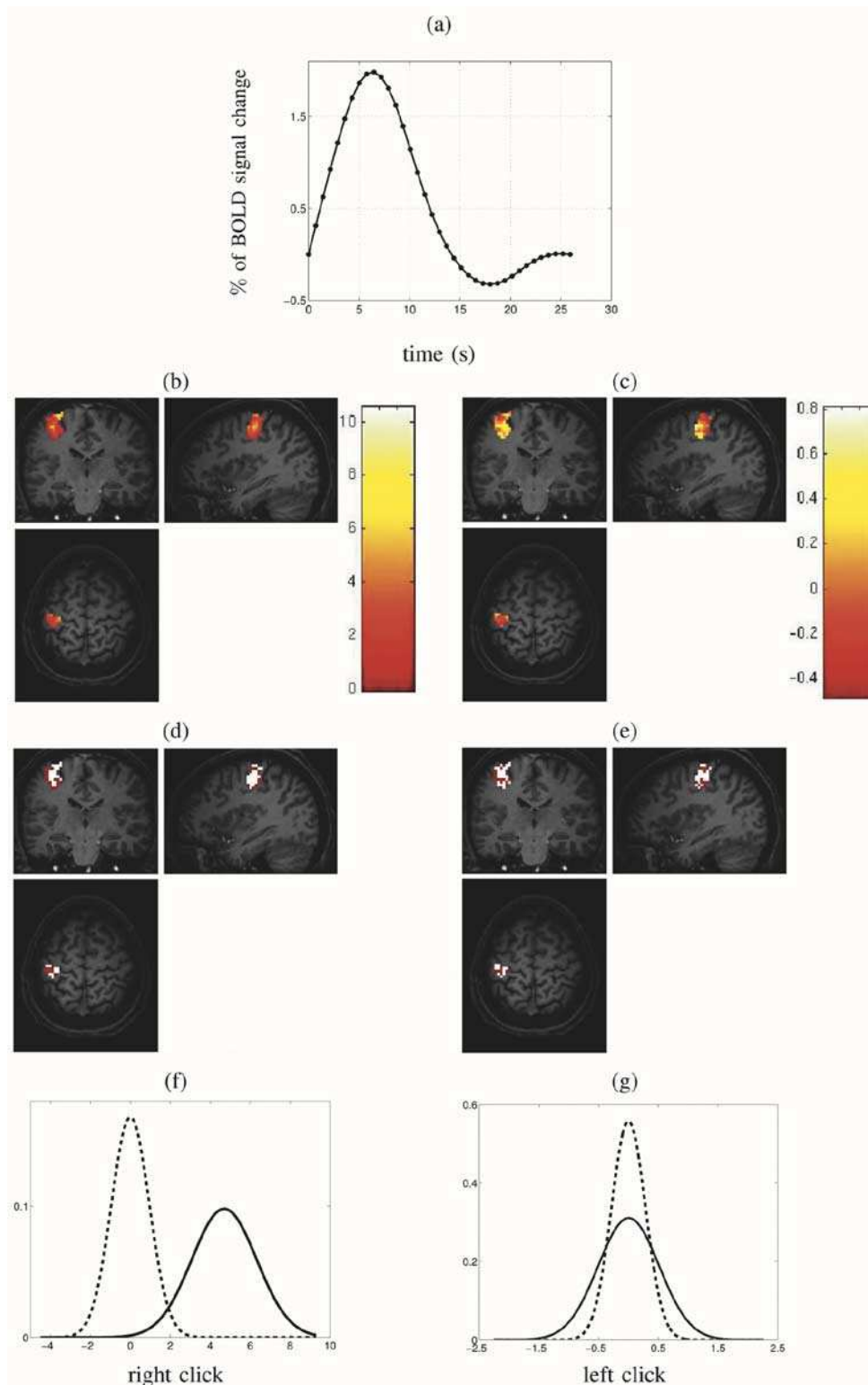
Fig. 8.   Estimation results for the (right click minus left click) contrast using raw (unsmoothed) data. From top to bottom and from left to right. (a) HRF estimate for $\mathcal{R}_1$ defined from the (right click minus left click) contrast. (b) Right click voxels NRLs estimates. (c) Left click voxels NRLs estimates. (d) and (e) Classification results for the right and left click stimulus, respectively (red: class 0, white: class 1). (f) and (g) Classes distributions of the right and left click stimuli, respectively. Dashed lines: class 0. Solid lines: class 1. Values are only plotted for voxels in the chosen ROI and not for all voxels in the brain.

Fig. 8(f)–(g), the estimated distributions of the two classes may be too close to each other to make their discrimination from the neural response levels efficient. To cope with this issue, we have implemented a classical test of the null hypothesis to keep voxels that survive a 5% risk of false positive threshold, as illustrated in Fig. 9(f)–(g). Clearly, the region of interest is almost entirely sig-
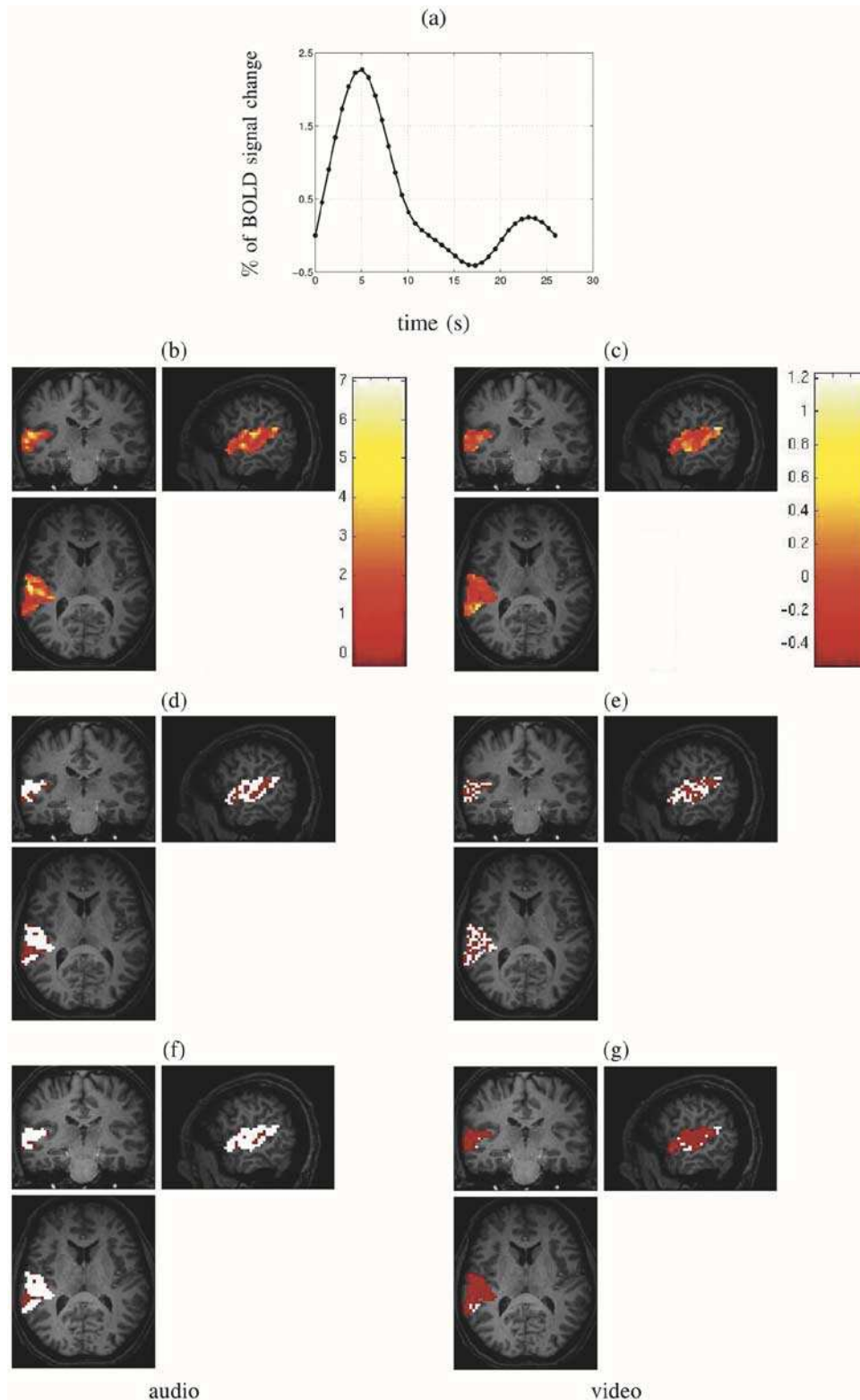
Fig. 9. Estimation results for the (audio minus video) contrast using raw (unsmoothed) data. From top to bottom and from left to right. (a) HRF estimate for $\mathcal{R}_2$ defined from the (audio minus video) contrast. (b) and (c) Audio (resp. video) voxels NRLs estimates. (d) and (e) Classification results for the audio and video stimulus, respectively (red: class 0, white: class 1). (f) and (g) Statistical test results for the audio (resp. video) stimulus after at the 5% risk of error. Significantly activated voxels are in white, and nonsignificant voxels are in red. Values are only plotted for voxels in the ROI and not for all voxels in the brain.

nificantly activated for the audio condition, whereas few voxels survive this threshold for the video condition.

Our approach was also tested using data smoothed with a 6-mm isotropic Gaussian kernel. Essential results conveyed the

same information, with a greater spatial homogeneity observed in all results (voxels NRL activity, voxels classifications, significant voxels). Results are not reported here but are available from the authors.

The *a priori* classification in two classes may seem as a constraint. Since no formal model testing is performed to choose between a model including only one class with the model that includes two classes, the interpretation of the data classified in class 1 is not straightforward when, for a given condition, the ROI is actually made up of nonactivated voxels. In this case, a model with only one class would be more suitable. However, we showed that a simple hypothesis testing procedure can be used for the detection of "activated" voxels. Besides, the implementation of a Bayesian information criterion would also easily raise this limitation. Note also that a threshold corrected for multiple comparison can be used if a strong specificity is needed from the analysis. In our experimental results, classes 1 and 0 were straightforwardly close or different because of our ROI choice. This may not be the case in other experimental conditions or if ROIs are selected on an anatomical basis.

The knowledge of the alternative hypothesis estimation (class 1 estimated model) is information that is to date not normally used in neuroimaging analyses but allows one to ask some different classes of questions to the data that cannot be answered in the null hypothesis testing framework. For instance, one may ask whether a given voxel activity is indeed not activated (its value is close to zero within a given confidence interval); this is information that is not obtainable from the "nonrejection" of the null hypothesis. Furthermore, one would also be able to control for the risk of false *negative* by setting a threshold based on the alternative distribution.

VIII. Conclusion

In this paper, we have proposed an original method for semi-blind deconvolution of impulse pseudo-neural response in functional neuroimaging. This method extends a previous work [28] to deal with regional HRF estimation in an efficient way while modeling the spatial variability of the pseudo-neural response for each stimulus type. Mixture modeling provides us a way of estimating the distributions of activated and nonactivated voxels from the data itself. Indeed, detection is performed on the estimated class 0 (centered on 0) and class 1 distribution parameters. To our knowledge, it is the first time that a Bayesian joint detection-estimation approach is proposed for the analysis of fMRI data. We have validated this approach on both synthetic and real fMRI data. The method is general enough to deal with all specific features of fMRI data (several sessions per subject, voxel specific noise variance, asynchronous timing between event onsets and scanning time, physiological artifacts, and so on). Our optimization scheme is also efficient enough to allow for large fMRI time series to be processed. Furthermore, this approach provides a tool to compare estimated activity between conditions, regions, or subjects.

Compared with standard detection techniques (like SPM), this paper introduces the joint estimation of the shape of the HRF and the associated spatial map and, therefore, should yield more precise estimations of voxel activity. Our modeling also

permits the use of the estimated distribution of activity under the "alternative" hypothesis in a Bayesian framework. In this regard, this approach provides the neuroscientist with a method that is able to answer questions that are not tractable in the classical null hypothesis framework [21]. Another advantage of estimating the "alternative" distribution is that it allows some adjustement for modeling assumption violations. In the case of imperfect modeling of the noise characteristics, for instance, the null distribution may have a different form from what is currently assumed in the null-hypothesis testing framework [21]. However, mixture modeling-based approaches may fail when the data does not meet the distributional assumptions, whatever the source of discrepancy (non Gaussianity, three-class mixture model, ...). This point is critical in our approach for stimulus types that do not elicit activation in any voxel of the selected ROI. This could be solved using a model selection step that requires the estimation of the parameters of the simpler one-class model described in [32].

The method can be extended in several ways. First, spatial regularization could be introduced on the spatial map of classification labels using 2–D (based on the cortical surface) or three–dimensional (3-D) (based on the volume) Markov random fields. The specific use of Pickard random fields could be particularly interesting to automatically tune the amount of regularization from the data without facing the estimation of the *partition function* [56]. Another solution has been recently proposed to cope with the same issue [47].

Second, the model presented here assumes that the NRLs are constant in time. However, phenomena such as adaptation or learning may be better modeled by including the time dimension in more sophisticated models. This is the subject of future research. Third, anatomical information (grey/white matter) have yet to be considered. However, we know that activation should be localized in the grey matter of the brain, and sophisticated segmentation tools of anatomical data are now available (see, for instance, the Brainvisa software[3]). If good registration between the anatomical and functional data can be achieved, this information should be taken into account in the model formulation as a prior.

Last, white noise is assumed in this paper. We have started to develop an extension that accounts for serial correlation of fMRI time series using a spatially varying first-order AR model of errors [41]. First, this extension is more computationally demanding since AR, as well as drift parameters, now need to be sampled. Our first tests on simulations have shown that this modeling has an influence over the posterior error bars of the NRLs. More precisely, for correlated synthetic datasets, we get smaller error bars when modeling this correlation. Nonetheless, our first tests on real fMRI data yielded no major improvement compared with the results described in Section VII.

Appendix

A. Computational Details

We will use specific notations to make reference to usual distributions (see Table III). In the following, we derive the poste-

---

[3]http://brainvisa.info/index.html

TABLE III
LIST OF DISTRIBUTION NOTATIONS

$\mathcal{U}$	uniform distribution
$\mathcal{N}$	Gaussian distribution
$\mathcal{IG}$	Inverse Gamma distribution
$\mathcal{B}$	Beta distribution

rior full conditional distributions from which we sample in our algorithm.

1) Hemodynamic Response Function: The full conditional posterior pdf $p(\boldsymbol{h}\,|\,\boldsymbol{y},\boldsymbol{a},\boldsymbol{\Theta})$ can be identified from the full posterior pdf $p(\boldsymbol{h},\boldsymbol{a},\boldsymbol{\Theta}\,|\,\boldsymbol{y})$, assuming that $(\boldsymbol{a},\boldsymbol{\Theta},\boldsymbol{y})$ are constant, while $\boldsymbol{h}$ may still vary: $p(\boldsymbol{h}\,|\,\boldsymbol{y},\boldsymbol{a},\boldsymbol{\Theta}) \propto p(\boldsymbol{y}\,|\,\boldsymbol{h},\boldsymbol{a},\epsilon^2)\,p(\boldsymbol{h})$. Since the likelihood and the prior are Gaussian, $\boldsymbol{h}$ also follows a Gaussian conditional distribution $p(\boldsymbol{h}\,|\,\mathrm{rest}) \sim \mathcal{N}(\boldsymbol{m_h},\boldsymbol{\Sigma_h})$, with

$$\boldsymbol{\Sigma_h}^{-1} = \sigma_{\boldsymbol{h}}^{-2}\boldsymbol{R}^{-1} + \sum_j \boldsymbol{S}_j^{\mathbf{t}}\boldsymbol{Q}_j\boldsymbol{S}_j,$$

$$\boldsymbol{m_h} = \boldsymbol{\Sigma_h}\sum_j \boldsymbol{S}_j^{\mathbf{t}}\boldsymbol{Q}_j\boldsymbol{y}_j \tag{7}$$

where $\boldsymbol{S}_j = \sum_m a_j^m \boldsymbol{X}^m$. In practice, drawing realizations from $p(\boldsymbol{h}\,|\,\mathrm{rest})$ is a three-step procedure [51].
1) Compute the square root of matrix $\boldsymbol{\Sigma_h}$. This can be achieved using Cholesky factorization such that $\boldsymbol{\Sigma_h} = \boldsymbol{LL}^{\mathbf{t}}$, where $\boldsymbol{L}$ is a lower triangular matrix.
2) Generate a Gaussian vector $\boldsymbol{k} \sim \mathcal{N}(\boldsymbol{0},\boldsymbol{I}_{D-1})$.
3) Compute $\boldsymbol{h} = \boldsymbol{m_h}+\boldsymbol{Lk}$ to get a sample from $\mathcal{N}(\boldsymbol{m_h},\boldsymbol{\Sigma_h})$.

2) HRF Scale: Sampling the hyperparameter $\sigma_{\boldsymbol{h}}^2$ amounts to simulating according to

$$\left(\sigma_{\boldsymbol{h}}^2\big|\,\mathrm{rest}\right) = p\left(\sigma_{\boldsymbol{h}}^2\big|\,\boldsymbol{h}\right) \sim \mathcal{IG}(D/2,\boldsymbol{h}^{\mathbf{t}}\boldsymbol{R}^{-1}\boldsymbol{h}/2). \tag{8}$$

3) "Neural" Response Levels: The sampling scheme of the NRLs relies on two nested loops, where the inner corresponds to the stimulus types (e.g., index m) and the outer to voxels (e.g., index j). The basic operation is to sample each a_j^m from the conditional distribution

$$p\left(a_j^m\big|\,\mathrm{rest}\right) = p\left(a_j^m\,\Big|\,\boldsymbol{y}_j,\boldsymbol{h},\boldsymbol{\theta}_m,\epsilon_j^2,z_j^{m'\neq m}\right).$$

This pdf is the marginal distribution of an *a posteriori* Gaussian mixture:

$$p\left(a_j^m\,\Big|\,\boldsymbol{y}_j,\boldsymbol{h},\boldsymbol{\theta}_m,\epsilon_j^2,z_j^{m'\neq m}\right)$$
$$= \sum_{i=0,1} p\left(a_j^m\,\Big|\,q_j^m = i,\boldsymbol{y}_j,\boldsymbol{h},\epsilon_j^2,z_j^{m'\neq m}\right)\Pr(q_j^m = i)$$
$$\sim \sum_{i=0,1} \lambda_{i,j}^m\mathcal{N}\left(\mu_{i,j}^m,\left(\sigma_{i,j}^m\right)^2\right). \tag{9}$$

After some straightforward calculations, we get the expression of the mixture parameters

$$\left(\sigma_{i,j}^m\right)^{-2} = \sigma_{i,m}^{-2} + \boldsymbol{g}_m^{\mathbf{t}}\boldsymbol{Q}_j\boldsymbol{g}_m,$$
$$\mu_{i,j}^m = \left(\sigma_{i,j}^m\right)^2\left(\boldsymbol{g}_m^{\mathbf{t}}\boldsymbol{Q}_j\boldsymbol{e}_{m,j} + i\,\mu_{i,m}\sigma_{i,m}^{-2}\right)$$

where $\boldsymbol{g}_m = \boldsymbol{X}^m\boldsymbol{h}$, and $\boldsymbol{e}_{m,j} = \boldsymbol{y}_j - \sum_{m'\neq m} a_j^{m'}\boldsymbol{g}_{m'} = \tilde{\boldsymbol{y}}_j + a_j^m\boldsymbol{g}_m$. The weighting probability $\lambda_{i,j}^m$ of the two distributions can be expressed as follows:

$$\lambda_{i,j}^m = \left(1 + \tilde{\lambda}_{1-i,j}^m\Big/\tilde{\lambda}_{i,j}^m\right)^{-1}$$
$$\tilde{\lambda}_{i,j}^m = \frac{\lambda_{i,m}^2\lambda_{1-i,m}}{\sigma_{i,m}}\sigma_{i,j}^m$$
$$\times \exp\left(\left(\mu_{i,j}^m\right)^2\Big/2\left(\sigma_{i,j}^m\right)^2\right), \quad i = 0,1.$$

As indicated in (9), we first need to sample the binary label q_j^m. For doing so, we generate u_j^m from $\mathcal{U}([0,1])$, and the following rule applies:

$$u_j^m \leq \lambda_{1,j}^m \Longrightarrow q_j^m = 1, \quad \text{otherwise } q_j^m = 0.$$

Then, the sampling of the NRL a_j^m is based on the full conditional pdf $p(a_j^m\,|\,q_j^m = i,\boldsymbol{y}_j,\boldsymbol{h},\epsilon_j^2,z_j^{m'\neq m})$:

$$p\left(a_j^m\,\Big|\,q_j^m = i,\boldsymbol{y}_j,\boldsymbol{h},\boldsymbol{\theta}_m,\epsilon_j^2\right) \sim \mathcal{N}\left(\mu_{i,j}^m,\left(\sigma_{i,j}^m\right)^2\right), \quad i = 0,1. \tag{10}$$

4) Noise Variances: Sampling the noise variances $\sigma_{\boldsymbol{\epsilon}}^2$ can be performed in parallel. Let us denote $\boldsymbol{a}_j = (a_j^m)_{m=1:M}$. Then, drawing a sample from $p(\sigma_{\varepsilon_j}^2\,|\,\mathrm{rest})$ is straightforward, according to

$$p\left(\epsilon_j^2\big|\,\mathrm{rest}\right) = p\left(\epsilon_j^2\big|\,\boldsymbol{y}_j,\boldsymbol{h},\boldsymbol{a}_j\right) \propto p\left(\boldsymbol{y}_j\,\big|\,\boldsymbol{h},\boldsymbol{a}_j,\epsilon_j^2\right)$$
$$\sim \mathcal{IG}\left((N+1-Q)/2,\|\tilde{\boldsymbol{y}}_j\|_{\boldsymbol{I}_N-\boldsymbol{PP}^{\mathbf{t}}}^2/2\right) \tag{11}$$

5) Weighting Probabilities: Sampling the weighting probabilities $\boldsymbol{\lambda}$ can be parallelized as well since λ_m is independent of $m' \neq m$. For simplicity, drawing a realization of λ_m consists of sampling from a Beta distribution

$$p(\lambda_m\,|\,\mathrm{rest}) = p(\lambda_m\,|\,\boldsymbol{q}^m) \propto p(\boldsymbol{q}^m\,|\,\lambda_m)p(\lambda_m)$$
$$\sim \mathcal{B}(J_{1,m} + 3/2, J_{0,m} + 3/2) \tag{12}$$

where $\boldsymbol{q}^m = (q_j^m)_{j=1:J}$. $J_{i,m}$ stands for the number of voxels in class $C_{i,m}$ for condition m. Note that $J_{1-i,m} = J - J_{i,m}$ for $i = 0,1$.

6) Mixture Parameters: Here, we focus on the generation of realizations of hyperparameters $\sigma_{0,m}^2$ and $(\mu_{1,m},\sigma_{1,m}^2)$ attached to classes $C_{0,m}$ and $C_{1,m}$, respectively:

$$p\left(\sigma_{0,m}^2\big|\,\mathrm{rest}\right) = p\left(\sigma_{0,m}^2\big|\,\boldsymbol{z}^m\right)$$
$$p\left(\mu_{1,m},\sigma_{1,m}^2\big|\,\mathrm{rest}\right) = p\left(\mu_{1,m},\sigma_{1,m}^2\big|\,\boldsymbol{z}^m\right)$$

where $\boldsymbol{z}^m = (\boldsymbol{a}^m,\boldsymbol{q}^m)$. Note that both the labels $\boldsymbol{q}^m$ and the corresponding NRLs $\boldsymbol{a}^m$ have to be known to be sampled from

the right posterior distribution for condition m. Bayes' rule enables us to derive the closed-form expression of these pdfs:

$$
p\left(\sigma_{0,m}^2 \,\middle|\, \boldsymbol{z}^m\right)
$$
$$
\propto p\left(\boldsymbol{a}^m \,\middle|\, \boldsymbol{q}^m, \boldsymbol{y}, \sigma_{0,m}^2\right) \Pr(\boldsymbol{q}^m) p\left(\sigma_{0,m}^2\right)
$$
$$
\propto \sigma_{0,m}^{-1} \prod_{j \in C_{0,m}} p\left(a_j^m \,\middle|\, q_j^m = 0, \boldsymbol{y}, \sigma_{0,m}^2\right)
$$
$$
p\left(\mu_{1,m}, \sigma_{1,m}^2 \,\middle|\, \boldsymbol{z}^m\right)
$$
$$
\propto p\left(\boldsymbol{a}^m \,\middle|\, \boldsymbol{q}^m, \boldsymbol{y}, \mu_{1,m}, \sigma_{1,m}^2\right) \Pr(\boldsymbol{q}^m) p\left(\mu_{1,m}, \sigma_{1,m}^2\right)
$$
$$
\propto \sigma_{1,m}^{-1} \prod_{j \in C_{1,m}} p\left(a_j^m \,\middle|\, q_j^m = 1, \boldsymbol{y}, \mu_{1,m}, \sigma_{1,m}^2\right).
$$

For $i = 0, 1$, let

$$
\nu_{i,m} = \sum_{j \in C_{i,m}} \left(a_j^m - \eta_{i,m}\right)^2 \quad \text{and}
$$
$$
\eta_{i,m} = \frac{1}{J_{i,m}} \sum_{j \in C_{i,m}} a_j^m.
$$

According to (10), $\{a_j^m\}_{j \in C_{0,m}}$ are i.i.d. variables, each being distributed according to $\mathcal{N}(0, \sigma_{0,m}^2)$-distributed. The ensuing posterior pdf of $\sigma_{0,m}^2$ therefore reads

$$
\left(\sigma_{0,m}^2 \,\middle|\, \boldsymbol{z}^m\right) \sim \mathcal{IG}\left(\frac{J_{0,m} - 1}{2}, \frac{\nu_{0,m}}{2}\right). \tag{13}
$$

For class $C_{1,m}$, we need to sample from the joint pdf of $\left(\mu_{1,m}, \sigma_{1,m}^2\right)$. This can be decomposed in two steps using ([51, Prop. 4.4.1, p 187]).

1) Draw a sample for the variance from $p(\sigma_{1,m}^2 \,|\, \boldsymbol{z}^m)$

$$
\left(\sigma_{1,m}^2 \,\middle|\, \boldsymbol{z}^m\right) \sim \mathcal{IG}\left(\frac{J_{1,m} - 1}{2}, \frac{\nu_{1,m}}{2}\right). \tag{14}
$$

2) Draw a sample for the mean from the conditional posterior pdf

$$
p\left(\mu_{1,m} \,\middle|\, \sigma_{1,m}^2, \boldsymbol{z}^m\right) \sim \mathcal{N}\left(\eta_{1,m}, \frac{\sigma_{1,m}^2}{J_{1,m}}\right). \tag{15}
$$

ACKNOWLEDGMENT

REFERENCES

[1] S. Ogawa, T. Lee, A. Kay, and D. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation," *Proc. Nat. Acad. Sci. USA*, vol. 87, no. 24, pp. 9868–9872, 1990.

[2] C. Goutte, P. Toft, E. Rostrup, F. A. Nielsen, and L. K. Hansen, "On clustering fMRI time series," *Neuroimag.*, vol. 9, no. 3, pp. 298–310, 1999.

[3] M. Fadili, S. Ruan, D. Bloyet, and B. Mazoyer, "A multistep unsupervised fuzzy clustering analysis of fMRI time series," *Hum. Brain Mapp.*, vol. 10, no. 4, pp. 160–178, 2000.

[4] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai, "Comparison of two exploratory data analysis methods for fMRI: Fuzzy clustering vs. principal component analysis," *Magn. Reson. Imaging*, vol. 18, no. 1, pp. 89–94, 2000.

[5] A. Anderson, D. Grash, and M. Avison, "Principal component analysis of the dynamic response measured by fMRI: A generalized linear systems framework," *J. Magn. Reson. Imag.*, vol. 17, pp. 795–815, 1999.

[6] M. McKeown, "Detection of consitently task-related activation in fMRI data with hybrid independent component analysis," *Neuroimag.*, vol. 11, pp. 24–35, 2000.

[7] V. Calhoun, T. Adali, G. Pearlson, and J. Pekar, "Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms," *Hum. Brain Mapp.*, vol. 13, pp. 43–53, 2001.

[8] R. S. J. Frackowiak, K. J. Friston, C. D. Dolan, and J. C. Maziotta, *Human Brain Function*. New York: Academic, 1997.

[9] D. Malonek and A. Grinvald, "Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: Implications for functional brain mapping," *Science*, vol. 272, pp. 551–554, 1996.

[10] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: The balloon model," *Magn. Reson. Med.*, vol. 39, pp. 855–864, Jun. 1998.

[11] J. Mandeville, J. Marota, C. Ayata, C. Zaharchuck, M. Moskowitz, B. Rosen, and R. Weisskoff, "Evidence of a cerebro-vascular post-arteriole windkessel with delayed compliance," *J. Cereb. Blood Flow Metab.*, vol. 19, pp. 679–689, 1999.

[12] K. J. Friston, A. Mechelli, R. Turner, and C. J. Price, "Nonlinear responses in fMRI: the balloon model, Volterra kernels, and other hemodynamics," *Neuroimag.*, vol. 12, pp. 466–477, 2000.

[13] J. Riera, J. Watanabe, I. Kazuki, M. Naoki, E. Aubert, T. Ozaki, and R. Kawashima, "A state-space model of the hemodynamic approach: Nonlinear filtering of BOLD signal," *Neuroimag.*, vol. 21, pp. 547–567, 2004.

[14] N. Lange, "Empirical and substantive models, the Bayesian paradigm, and meta-analysis in functional brain imaging," *Hum. Brain Mapp.*, vol. 5, pp. 259–263, 1997.

[15] M. S. Cohen, "Parametric analysis of MRI data using linear systems methods," *Neuroimag.*, vol. 6, pp. 93–103, 1997.

[16] J. C. Rajapakse, F. Kruggel, J. M. Maisog, and D. Von Cramon, "Modeling hemodynamic response for analysis of functional MRI time-series," *Hum. Brain Mapp.*, vol. 6, pp. 283–300, 1998.

[17] G. H. Glover, "Deconvolution of impulse response in event-related BOLD fMRI," *Neuroimag.*, vol. 9, pp. 416–429, 1999.

[18] G. K. Aguirre, E. Zarahn, and M. D'Esposito, "The variability of humain bold hemodynamic responses," *Neuroimag.*, vol. 7, p. 574, 1998.

[19] R. L. Buckner, J. Goodman, M. Burock, M. Rotte, W. Koutstaal, D. L. Schachter, B. R. Rosen, and A. M. Dale, "Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI," *Neuron*, vol. 20, pp. 285–296, 1998.

[20] F. M. Miezin, L. Maccotta, J. M. Ollinger, S. E. Petersen, and R. L. Buckner, "Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing," *Neuroimag.*, vol. 11, pp. 735–759, 2000.

[21] K. J. Friston, A. P. Holmes, K. Worsley, J.-B. Poline, C. Frith, and R. Frackowiak, "Statistical parametric maps in functional neuroimaging: A general linear approach," *Hum. Brain Mapp.*, vol. 2, pp. 189–210, 1995.

[22] O. Joseph and R. N. Henson, "Event-related functional magnetic resonance imaging: Modeling, inference and optimization," *Philos. Trans. R. Soc. Lond. B: Biol.*, vol. 354, no. 1387, pp. 1215–1228, 1999.

[23] C. Genovese, "A Bayesian time-course model for functional magnetic resonance imaging data (with discussion)," *J. Acoust. Soc. Amer.*, vol. 95, pp. 691–719, 2000.

[24] C. Gössl, L. Fahrmeir, and D. P. Auer, "Bayesian modeling of the hemodynamic response function in BOLD fMRI," *Neuroimag.*, vol. 14, pp. 140–148, 2001.

[25] M. Woolrich, M. Jenkinson, J. Brady, and S. Smith, "Fully Bayesian spatio-temporal modeling of fMRI data," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 213–231, Feb. 2004.

[26] C. Goutte, F. A. Nielsen, and L. K. Hansen, "Modeling the haemodynamic response in fMRI using smooth filters," *IEEE Trans. Med. Imag.*, vol. 19, no. 12, pp. 1188–1201, Dec. 2000.

[27] G. Marrelec, H. Benali, P. Ciuciu, M. Pélégrini-Issac, and J.-B. Poline, "Robust Bayesian estimation of the hemodynamic response function in event-related BOLD MRI using basic physiological information," *Hum. Brain Mapp.*, vol. 19, no. 1, pp. 1–17, 2003.

[28] P. Ciuciu, J.-B. Poline, G. Marrelec, J. Idier, C. Pallier, and H. Benali, "Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment," *IEEE Trans. Med. Imag.*, vol. 22, no. 10, pp. 1235–1251, Oct. 2003.

[29] P. Ciuciu, J. Idier, A. Roche, and C. Pallier, "Outlier detection for robust region-based estimation of the hemodynamic response function in event-related fMRI," in *Proc. 2th Proc. IEEE ISBI*, Arlington, VA, 2004, pp. 392–395.

[30] C. Gössl, D. P. Auer, and L. Fahrmeir, "Dynamic models in fMRI," *Magn. Reson. Med.*, vol. 43, pp. 72–81, 2000.

[31] ——, "Bayesian spatio-temporal modeling of the hemodynamic response function in BOLD fMRI," *Biometr.*, vol. 57, pp. 554–562, 2001.

[32] S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline, "Semi-blind deconvolution of neural impulse response in event-related fMRI using Gibbs sampler," in *Proc. 2th Proc. IEEE ISBI*, Arlington, VA, 2004, pp. 860–863.

[33] N. V. Hartvig and J. Jensen, "Spatial mixture modeling of fMRI data," *Hum. Brain Mapp.*, vol. 11, no. 4, pp. 233–248, 2000.

[34] B. S. Everitt and E. T. Bullmore, "Mixture model mapping of brain activation in functional magnetic resonance images," *Hum. Brain Mapp.*, vol. 7, pp. 1–14, 1999.

[35] G. Marrelec, P. Ciuciu, M. Pélégrini-Issac, and H. Benali, "Estimation of the hemodynamic response function in event-related functional MRI: Bayesian networks as a framework for efficient Bayesian modeling and inference," *IEEE Trans. Med. Imag.*, vol. 23, no. 8, pp. 959–967, Aug. 2004.

[36] E. Zarahn, G. K. Aguirre, and M. D'Esposito, "Empirical analysis of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions," *Neuroimag.*, vol. 5, pp. 179–197, 1997.

[37] F. G. Meyer, "Wavelet-based estimation of a semi-parametric generalized linear model of fMRI time series," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 315–322, Mar. 2003.

[38] E. T. Bullmore, M. Brammer, S. C. Williams, S. Rabe-Hesketh, N. Janot, A. David, J. Mellers, R. Howard, and P. Sham, "Statistical methods of estimation and inference for functional MR image analysis," *Magn. Reson. Med.*, vol. 35, pp. 261–277, 1996.

[39] M. Woolrich, B. Ripley, M. Brady, and S. Smith, "Temporal autocorrelation in univariate linear modeling of fMRI data," *Neuroimag.*, vol. 14, no. 6, pp. 1370–1386, 2001.

[40] K. Worsley, C. Liao, J. Aston, V. Petre, G. Duncan, F. Morales, and A. Evans, "A general statistical analysis for fMRI data," *Neuroimag.*, vol. 15, no. 1, pp. 1–15, 2002.

[41] S. Makni, P. Ciuciu, J. Idier, and J.-B. Poline. (2005) Joint Detection-Estimation of Brain Activity in fMRI: An Extended Model Accounting for Serial Correlation. CEA/SHFJ, Orsay, France. [Online]. Available: http://www.madic.org/biblio/en/Year/2005.php

[42] G. Marrelec, H. Benali, P. Ciuciu, and J.-B. Poline, "Bayesian estimation of the hemodynamic response function in functional MRI," in *Proc. Bayesian Inference Maximum Entropy Methods Workshop*, R. Fry, Ed., 2001.

[43] G. Flandin, F. Kherif, X. Pennec, D. Rivière, N. Ayache, and J.-B. Poline, "Parcellation of brain images with anatomical and functional constraints for fMRI data analysis," in *Proc. First Proc. IEEE ISBI*, Washington, DC, 2002, pp. 907–910.

[44] B. Thirion and O. Faugeras, "Feature detection in fMRI data: The information bottleneck approach," in *Proc. 6th MICCAI*, Montréal, QC, Canada, 2003, pp. 83–91.

[45] W. Penny, N. Trujillo-Barreto, and K. J. Friston, "Bayesian fMRI time series analysis with spatial priors," *Neuroimag.*, vol. 23, no. 2, pp. 350–362, 2005.

[46] M. Smith, B. Pütz, D. Auer, and L. Fahrmeir, "Assessing brain activity through spatial Bayesian variable selection," *Neuroimag.*, vol. 20, pp. 802–815, 2003.

[47] M. Woolrich, T. Behrens, C. Beckmann, and S. Smith, "Mixture models with adpative spatial regularization for segmentation with an application to fMRI data," *IEEE Trans. Med. Imag.*, vol. 24, no. 1, pp. 1–11, Jan. 2005.

[48] Q. Cheng, R. Chen, and T.-H. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 377–384, Feb. 1996.

[49] R. Buxton and L. Frank, "A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation," *J. Cereb. Blood Flow Metab.*, vol. 17, no. 1, pp. 64–72, 1997.

[50] M. Woolrich, M. Jenkinson, J. M. Brady, and S. Smith, "Constrained linear basis set for HRF modeling using variational Bayes," *Neuroimag.*, vol. 21, pp. 1748–1761, 2004.

[51] C. P. Robert, *The Bayesian Choice*, Second ed, ser. Springer Texts in Statistics. New York: Springer Verlag, 2001.

[52] T. Liu, L. Frank, E. C. Wong, and R. B. Buxton, "Detection power, estimation efficiency, and predictability in event-related fMRI," *Neuroimag.*, vol. 13, no. 4, pp. 759–773, 2001.

[53] S. Gautier, J. Idier, A. Mohammad-Djafari, and B. Lavayssière, "Traitement d'échogrammes ultrasonores par déconvolution aveugle," in *Actes du 16e Colloq. GRETSI*, Grenoble, France, Sep. 1997, pp. 1431–1434.

[54] Y.-L. You and M. Kaveh, "A regularization approach to joint blur identification and image restoration," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 416–428, May 1996.

[55] E. Gassiat, F. Monfront, and Y. Goussard, "On simultaneous signal estimation and parameter identification using a generalized likelihood approach," *IEEE Trans. Inf. Theory*, vol. 38, no. 1, pp. 157–162, Feb. 1992.

[56] J. Idier, Y. Goussard, and A. Ridolfi, "Unsupervised image segmentation using a telegraph parameterization of Pickard random fields," in *Spatial Statistics. Methodological Aspects and Some Applications*, vol. 159, M. Moore, Ed.. New York, 2001, pp. 115–140.

Salima Makni was born in Tunisia in 1979. She graduated from the École Supérieure des Communications, Tunis, Tunisia, in 2002 and received the D.E.A. degree in signal processing from the Université de Paris-sud, Orsay, France, in 2003. Since 2003, she has been pursuing the Ph.D. degree with the Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique, Orsay.

Her general interests are signal processing and biomedical imaging. She is mainly concerned with functional magnetic resonance imaging (fMRI), and she focuses on new approaches that allow for a better comprehension of brain functions.

Philippe Ciuciu (M'02) was born in France in 1973. He graduated from the École Supérieure d'Informatique Électronique Automatique, Paris, France, in 1996 and received the D.E.A. and Ph.D. degrees in signal processing from the Université de Paris-sud, Orsay, France, in 1996 and 2000, respectively.

Since November 2000, he has held a postdoctoral position with the Service Hospitalier Frédéric Joliot, Commissariat à l'Énergie Atomique, Orsay, where, since November 2001, he has held a full research scientist position with the functional and anatomical neuroimaging unit. His research interests include spectral analysis and optimization, and presently, he focuses on the application of statistical methods and regularized approaches to functional brain imaging (fMRI).

Jérôme Idier was born in France in 1966. He received the diploma degree in electrical engineering from École Supérieure d'Électricité, Gif-sur-Yvette, France, in 1988 and the Ph.D. degree in physics from University of Paris-Sud, Orsay, France, in 1991.

Starting in 1991, he was with the Centre National de la Recherche Scientifique. He is currently with the Institut de Recherche en Communications et Cybernétique, Nantes, France. His major scientific interest are in probabilistic approaches to inverse problems for signal and image processing.

Jean-Baptiste Poline received the degree in biomathematics in 1991 and the Ph.D. degree in 1994.

Since then, he has been interested in functional brain imaging analyses. He is now with the Service Hospitalier Fridiric Joliot, Orsay, France, where he is responsible for the functional data analysis group.

VIII.2 Extension cerveau entier

[**?**] S. Makni, J. Idier, T. Vincent B. Thirion, G. Dehaene-Lambertz and P. Ciuciu, « A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI », accepté pour publication dans *NeuroImage*, janvier 2008.

A fully Bayesian approach to the parcel-based detection-estimation of brain activity in fMRI

Salima Makni,[a] Jérôme Idier,[b] Thomas Vincent,[c,d] Bertrand Thirion,[e]
Ghislaine Dehaene-Lambertz,[f,d] and Philippe Ciuciu[c,d,*]

[a]*Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, John Radcliffe Hospital, Oxford, UK*
[b]*IRCCyN (CNRS), Nantes, France*
[c]*CEA, NeuroSpin, Gif-sur-Yvette, France*
[d]*IFR 49, Institut d'Imagerie Neurofonctionnelle, Paris, France*
[e]*INRIA Futurs, Orsay, France*
[f]*INSERM U562, NeuroSpin, Gif-sur-Yvette, France*

Within-subject analysis in fMRI essentially addresses two problems, i.e., the *detection* of activated brain regions in response to an experimental task and the *estimation* of the underlying dynamics, also known as the characterisation of Hemodynamic response function (HRF). So far, both issues have been treated sequentially while it is known that the HRF model has a dramatic impact on the localisation of activations and that the HRF shape may vary from one region to another. In this paper, we conciliate both issues in a region-based joint detection-estimation framework that we develop in the Bayesian formalism. Instead of considering function basis to account for spatial variability, spatially adaptive General Linear Models are built upon region-based non-parametric estimation of brain dynamics. Regions are first identified as functionally homogeneous *parcels* in the mask of the grey matter using a specific procedure [Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J.-B., August 2006. Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets. Hum. Brain Mapp. 27 (8), 678–693.]. Then, in each parcel, prior information is embedded to constrain this estimation. Detection is achieved by modelling activating, deactivating and non-activating voxels through mixture models within each parcel. From the posterior distribution, we infer upon the model parameters using Markov Chain Monte Carlo (MCMC) techniques. Bayesian model comparison allows us to emphasize on artificial datasets first that inhomogeneous gamma-Gaussian mixture models outperform Gaussian mixtures in terms of sensitivity/specificity trade-off and second that it is worthwhile modelling serial correlation through an AR(1) noise process at low signal-to-noise (SNR) ratio. Our approach is then validated on an fMRI experiment that studies habituation to auditory sentence repetition. This phenomenon is clearly recovered as well as the

hierarchical temporal organisation of the superior temporal sulcus, which is directly derived from the parcel-based HRF estimates.
© 2008 Elsevier Inc. All rights reserved.

Introduction

Since the first report of the BOLD effect in human (Ogawa et al., 1990), functional magnetic resonance imaging (fMRI) has represented a powerful tool to non-invasively study the relation between cognitive stimulus and the hemodynamic (BOLD) response. Within-subject analysis in fMRI is usually addressed using a hypothesis-driven approach that actually postulates a model for the HRF and enable voxelwise inference in the General Linear Model (GLM) framework. In this formulation, the modelling of the BOLD response i.e., the definition of the *design matrix* is crucial. In its simplest form, this matrix relies on a spatially invariant temporal model of the BOLD signal across the brain meaning that the expected response to each stimulus is modelled by a single regressor. Assuming the neurovascular system as *linear* and *time-invariant* (LTI), this regressor is built as the convolution of a sparse spike train representing the stimulation signal and the canonical HRF, i.e., a composition of two gamma functions which reflects the BOLD signal best in the visual and motor cortices (Glover, 1999).

Intra-individual differences in the characteristics of the HRF have been exhibited between cortical areas in (Aguirre et al., 1998; Miezin et al., 2000; Neumann and Lohmann, 2003; Handwerker

* Corresponding author. Tel.: +33 1 6908 7785; fax: +33 1 6908 7855.
E-mail address: philippe.ciuciu@cea.fr (P. Ciuciu).
Available online on ScienceDirect (www.sciencedirect.com).

et al., 2004). Although smaller than inter-individual fluctuations, this regional variability is large enough to be regarded with care. To account for these spatial fluctuations at the voxel level, one usually resorts to hemodynamic function basis. For instance, the canonical HRF can be supplemented with its first and second derivatives to model differences in time (Friston, 1998; Henson et al., 2002). To make the basis spatially adaptive, Woolrich et al. (2004a) have proposed a half-cosine parameterisation in combination to the selection of the best basis set. Although powerful and elegant, the price to be paid for such a flexible modelling lies in a loss of sensitivity of detection: the larger the number of regressors in the basis, the smaller the number of effective degrees of freedom in any subsequent statistical test. Crucially, in a GLM involving several regressors per condition, the Student-t statistic can no longer be used to infer on differences between experimental conditions. Rather, an *unsigned* Fisher statistic has to be computed, making direct interpretation of activation maps more difficult. Indeed, the null hypothesis is actually rejected whenever any of the contrast components deviates from zero and not specifically when the difference of the response magnitudes is far from zero.

In this paper, to facilitate cognitive interpretations, we argue in favour of a spatially adaptive GLM in which a local estimation of the HRF is performed. This allows us to factorise the expected BOLD response with a *single* regressor attached to each experimental condition and to enforce direct statistical comparisons based on response magnitudes. However, to conduct the analysis in an efficient and reliable manner, local estimation is performed at the scale of several voxels.

As mentioned earlier, the localisation of brain activation strongly depends on the modelling of the brain response and thus of its estimation. Of course, the converse also holds: HRF estimation is only relevant in voxels that elicit signal fluctuations correlated with the paradigm. Hence, detection and estimation are intrinsically linked to each other. The key point is therefore to tackle the two problems in a common setting, *i.e.,* to set up a formulation in which *detection* and *estimation* enter naturally and simultaneously. This setting cannot be the classical hypothesis testing framework. Indeed, the sequential procedure which first consists in estimating the HRF on a given dataset and then building a specific GLM upon this estimate for detecting activations in the same dataset, entails statistical problems in terms of sensitivity and specificity: the control of the false positive rate actually becomes hazardous due to the use of an erroneous number of degrees of freedom. We rather propose a Bayesian approach that provides an appropriate framework to address both detection and estimation issues in the same formalism.

The literature on Bayesian fMRI methods offers several approaches to adequately choose priors for detection. As introduced in (Everitt and Bullmore, 1999; Vaever Hartvig and Jensen, 2000; Penny and Friston, 2003) and further developed in (Smith et al., 2003; Woolrich et al., 2005; Ou and Golland, 2005; Woolrich and Behrens, 2006; Flandin and Penny, 2007), prior mixture models define an appropriate way to perform the classification or the segmentation of statistical parametric maps into activating, non-activating or deactivating brain regions. The pioneering contributions related to mixture modelling in fMRI (Everitt and Bullmore, 1999; Vaever Hartvig and Jensen, 2000) have proposed a voxel by voxel classification to decide whether the estimated effect is analogous to signal or noise in each voxel. Yet, the use of mixture modelling in a joint detection-estimation problem introduces specific concerns in comparison to the usual "hypothesis testing framework". Indeed, our data are *not* the voxelwise *z*-statistics but rather the raw fMRI time courses, which are required for the estimation step.

As regards HRF estimation, various priors may be thought of depending on the underlying HRF model. Basically, three classes of models coexist. Parametric models appeared first in the literature (Friston, 1994; Lange, 1997; Cohen, 1997; Rajapakse et al., 1998; Kruggel and Von Crammon, 1999). In this setting, the estimation problem consists in minimising some criterion with respect to (w.r.t.) some parameters of a precise function (*e.g.,* Gaussian, gamma ,…). However, parametric models tend to introduce some bias in the HRF estimate, since it is unlikely that they capture the true shape variations of the brain dynamics. Moreover, the objective function to be minimized is often non-convex making the parameter estimates unreliable and dependent of the initialisation. Hence, more flexible *semi-parametric* approaches have been proposed later to capture these variations (Genovese, 2000; Gossl et al., 2001; Woolrich et al., 2004a). In a semi-parametric framework, the HRF time course is decomposed into different periods (initial dip, attack, rise, decay, fall, …), each of them being described by specific parameters. At the same time, *non-parametric* approaches or Finite Impulse Response (FIR) models have emerged in the fMRI literature as a powerful tool to infer on the HRF shape (Nielsen et al., 1997; Goutte et al., 2000; Marrelec et al., 2003, 2004; Ciuciu et al., 2003). Most of these works take place in the Bayesian setting and constrain the HRF to be temporally smooth, which warrants a stable estimation in case of ill-posed identification.

Whatever the model in use, most methods are massively univariate and therefore neglect the spatial structure of the BOLD signal. Early investigations have shown that estimating the HRF using regularised FIR models over a functionally homogeneous region of interest provides more reliable results (Gössl et al., 2001; Ciuciu et al., 2004). In the following, a region-based non-parametric model of the HRF is therefore adopted. Then, the critical issue arising consists in exhibiting a functionally homogeneous clustering of the fMRI datasets over the whole brain. To that end, the grey matter's mask is segregated into a few hundreds of connected Regions of Interest (ROIs), called *parcels*. The parcels are derived using the parcellation procedure proposed by Thirion et al. (2006), according to a compound criterion balancing spatial and functional homogeneity. The second step of our analysis solves for the detection-estimation problem over each parcel.

The rest of this paper is organised as follows. Section 2 details how anatomical information is handled and how parcels are built up. Then, the forward parcel-based model of the BOLD signal is derived and priors over the unknown parameters are specified. In Section 3, we explain the key steps of our inferential procedure based on MCMC methods, *posterior mean* (PM) HRF estimation and marginal Maximum *A Posteriori* (MAP) classification for detection. On artificial datasets, Section 4 reports the performance of our approach in terms of sensitivity-specificity trade-off depending on the mixture prior and the noise modelling. In Section 5, our joint detection-estimation approach is tested on real fMRI data acquired during an habituation study to auditory sentence repetition. On the same datasets, we also performed a classical GLM analysis employing the widely used Statistical parametric mapping (SPM) software[1]. The two approaches are then compared and the main differences are exhibited. The pros and cons of the proposed method are discussed in Section 6 and some future extensions are envisaged.

---

[1] http://www.fil.ion.ucl.ac.uk/spm/.

Methodology

Definition of functionnally homogeneous brain regions

Anatomical representation

The segmentation of the grey-white matter interface is performed on an anatomical T_1-weighted MRI image using the BrainVisa software[2] (Mangin et al., 1995). It provides us with the anatomical representation of the cortex. To accommodate the coarser spatial resolution of fMRI data (typically, 3.5 mm along each direction), the grey matter mask $\mathcal{M}_a$ is dilated using a sphere as structural element, with a radius equal to the resolution of functional images.

Concurrently, a functional mask $\mathcal{M}_f$ was computed from the motion-corrected[3] BOLD EPI volumes. Also, an average EPI volume was created. Then, we carried out a histogram analysis of this volume: a Gaussian density $\mathcal{N}(\mu, \sigma^2)$ was fitted on the main mode of the EPI signal of interest. A threshold defined as $\mu - 3\sigma$ was used to obtain the functional mask. Finally, the mask of interest where activation most likely occurs was built as $\mathcal{M}_s = \mathcal{M}_a \cap \mathcal{M}_f$.

Parcellation of the grey-matter

The volume in mask $\mathcal{M}_s$ was then divided in K functionally homogeneous parcels or ROIs using the parcellation technique proposed in (Thirion et al., 2006). The goal of this procedure is to segregate the brain into connected and functionally homogeneous components. For doing so, the parcellation algorithm relies on the minimisation of a compound criterion reflecting both the spatial and functional structures and hence the topology of the dataset. The spatial similarity measure favours the closeness in the Talairach coordinates system. The functional part of this criterion is computed on parameters that characterise the functional properties of the voxels. These parameters can be chosen either as the fMRI time series themselves or as the β-parameters estimated during a first-level SPM analysis. The latter choice is nothing but a projection onto a subspace of reduced dimension, *i.e.,* the feature space. Typically, the feature space is defined from a F-contrast in a SPM analysis.

The number of parcels K needs to be set by hand. The larger the number of parcels, the higher the degree of within-parcel homogeneity. Of course, there exists a trade-off between the within-parcel homogeneity and the signal-to-noise ratio (SNR). If the number of voxels is too small in a given parcel, the HRF estimation may become inaccurate, specifically in regions where no voxel elicits a specific response to any experimental condition. To objectively choose an adequate number of parcels, Thyreau et al. (2006) have used the Bayesian information criterion (BIC) and cross validation techniques on an fMRI study of ten subjects. They have shown converging evidence for $K \approx 500$ for a whole brain analysis and recommend $K = 200$ as a fair setting for a restricted analysis to the grey matter's mask leading to typical parcel sizes around a few hundreds voxels.

Parcel-based modelling of the BOLD signal

Vectors and matrices are displayed in lower and upper cases, respectively, both in bold font (*e.g., y* and *P*). Unless stated otherwise, subscripts *i*, *j*, *m* and *n* are respectively indexes over mixture components, voxels, stimulus types and time points. We

refer the reader to Appendix A for the definitions of the non-standard probability density functions (pdf). Also, the pdf families are denoted using calligraphic letters (*e.g.,* $\mathcal{N}$ and $\mathcal{G}$ for the Gaussian and gamma densities).

The regional forward model of the BOLD signal introduced in (Makni et al., 2005) is used to account for voxel-dependent and task-related fluctuations of the magnitudes of the BOLD response. Hereafter, the latter magnitudes are called *Neural Response Levels* (NRLs). In short, this time-invariant model characterises each and every parcel by a single HRF shape and a NRL for each voxel and stimulus type. As shown in Fig. 1, this means that although the HRF shape is assumed constant within a parcel, the magnitude of the activation can vary in space and across experimental conditions. Let $\mathcal{P} = (V_j)_{j=1:J}$ be the current parcel and V_j a voxel in $\mathcal{P}$. Then, the generative BOLD model reads:

$$y_j = \sum_{m=1}^{M} a_j^m X^m h + P \ell_j + b_j, \qquad (1)$$

where

- $y_j = (y_{j,t_n})_{n=1:N}$ denotes the BOLD fMRI time course measured in voxel V_j at times $(t_n)_{n=1:N}$ (N is the number of scans) with $t_n = n$TR and TR is the time of repetition,
- $X^m = \left(x^m_{t_n - d\Delta t} \right)_{n=1:N, d=0:D}$ is a $N \times (D+1)$ binary matrix corresponding to the arrival times for the *m*th condition. Δt is the sampling period of the HRF, usually smaller than TR. The onsets of the stimuli are put on the Δt-sampled grid by moving them to the nearest time points on this grid. Note that X^m can be adapted to paradigms having trial varying stimulus magnitudes or durations.
- Vector $h = (h_{d\Delta t})_{d=0:D}$ represents the unknown HRF shape in parcel $\mathcal{P}$ ($D+1$ is the number of HRF coefficients). It actually seems reasonable to assume a single HRF shape in homogeneous parcels.
- a_j^m stands for the NRL in voxel V_j for condition m (M is the number of experimental conditions in the paradigm). Hence, the activation profile associated to the *m*th stimulus type in voxel V_j is computed as the product $h \times a_j^m$.
- $P = [p_1, \dots, p_Q]$ is the low frequency orthogonal matrix of size $N \times Q$. It consists of an orthonormal basis of functions $p_q = (p_q(t_n))_{n=1:N}$. To each voxel is attached an unknown weighting vector ℓ_j that has to be regressed in order to estimate the trend in V_j. We denote $\mathbb{l} = (\ell_j)_{j=1:J}$ the set of low frequency drifts involved in $\mathcal{P}$.
- $b_j \in \mathbb{R}^N$ is the noise vector in voxel V_j. In (Woolrich et al., 2001; Worsley et al., 2002) an autoregressive (AR) noise model has been introduced to account for the serial correlation of the fMRI time series in the detection analysis. Importantly, when this temporal correlation is correctly estimated, the number of false positives decreases, yielding more conservative detection results. Similarly, in a joint detection-estimation framework, Makni et al. (2006b) have shown that the introduction of a spatially-varying first-order AR noise in model (1) provides a lower false positive rate. Hence, b_j is defined by $b_{j,t_n} = \rho_j b_{j,t_{n-1}} + \varepsilon_{j,t_n}, \forall j, t$, with $\varepsilon_j \sim \mathcal{N}(0_N, \sigma^2_{\varepsilon_j} I_N)$,, where 0_N is a null vector of length N, and I_N is the identity matrix of size N.

Although the noise structure is correlated in space (Woolrich et al., 2004b) and non-stationary across tissues (Worsley et al., 2002), we do not essentially account for this correlation for two
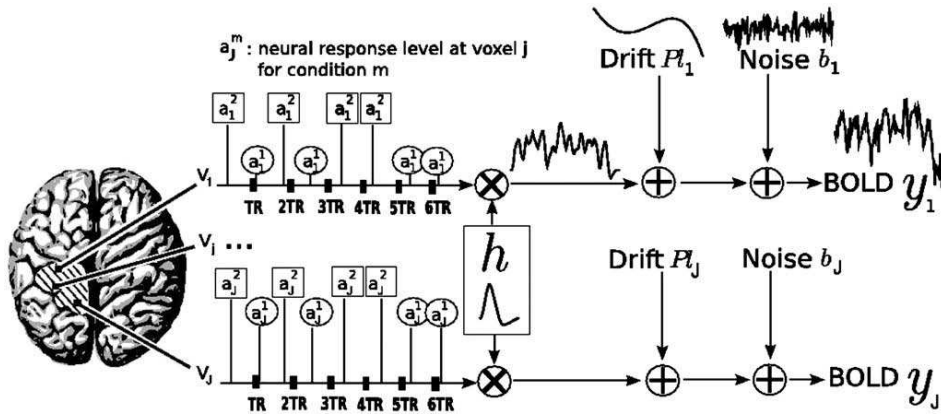

Fig. 1. Summary of the proposed regional BOLD model. The size of each parcel $\mathcal{P}$ varies typically between by a few tens and a few hundreds: $80 \leqslant J \leqslant 350$. The number M of experimental conditions involved in the model usually varies from 1 to 5. In our example, $M=2$, the NRLs $(\boldsymbol{a}_j^1, \boldsymbol{a}_j^2)$ corresponding to the first and the second conditions are surrounded by circles and squares, respectively. Note that our model accounts for asynchronous paradigms in which the onsets do not necessarily match acquisition time points. As illustrated, the NRLs take different values from one voxel to another. The HRF $\boldsymbol{h}$ can be sampled at a period of 1 s and estimated on a range of 20 to 25 s (e.g., $D=25$). Most often, the LFD coefficients ℓ_j are estimated on a few components ($Q=4$).

reasons. First, it is likely that a large part of the noise may be due to misspecification of the HRF. Second, we actually assume that the spatial correlation of the signal of interest is more important. Hence, the fMRI time series $\mathbb{y} = (\boldsymbol{y}_j)_{j=1:J}$ are supposed to be statistically conditionally independent. The likelihood then factors over voxels:

$$p\big(\mathbb{y}|\boldsymbol{h}, \mathbb{a}, \mathbb{l}, \boldsymbol{\theta}_0\big) = \prod_{j=1}^{J} p\Big(\boldsymbol{y}_j|\boldsymbol{h}, \boldsymbol{a}_j, l_j, \theta_{0,j}\Big) \qquad (2)$$
$$\propto \prod_{j=1}^{J} |\Lambda_j|^{1/2} \sigma_{\varepsilon_j}^{-N} \exp\left(-\sum_{j=1}^{J} \frac{\widetilde{\boldsymbol{y}}_j^t \Lambda_j \widetilde{\boldsymbol{y}}_j}{2\sigma_{\varepsilon_j}^2}\right)$$

where $\theta_{0,j} = (\rho_j, \sigma_{\varepsilon_j}^2)$, $\boldsymbol{\theta}_0 = (\theta_{0,j})_{j=1:J}$ and $\widetilde{\boldsymbol{y}}_j = \boldsymbol{y}_j - \Sigma_m a_j^m \boldsymbol{X}^m \boldsymbol{h} - \boldsymbol{P}\boldsymbol{\ell}_j$. Note that $\sigma_{\varepsilon_j}^{-2} \Lambda_j$ defines the inverse of the autocorrelation matrix of $\boldsymbol{b}_j$. According to Kay (1988, Chap VI, p. 177), Λ_j is tridiagonal, with $(\Lambda_j)_{1,1} = (\Lambda_j)_{N,N} = 1$, $(\Lambda_j)_{n,n} = 1 + \rho_j^2$ and $(\Lambda_j)_{n+1,n} = (\Lambda_j)_{n,n+1} = -\rho_j \ \forall n = 2:N-1$. Its determinant is given by $|\Lambda_j| = 1 - \rho_j^2$. In what follows, we do not approximate Eq. (2) by dropping the term $|\Lambda_j|^{1/2}$, as done in previous works (Roberts and Penny, 2002; Penny et al., 2003; Woolrich et al., 2004b). Indeed, when the AR parameter ρ significantly departs from zero (e.g., $\rho \geq 0.4$), this approximation is biased and potentially far from the exact likelihood.

On the sole basis of the likelihood function (2), it seems impractical to identify the pair $(\boldsymbol{h}, \mathbb{a})$. Indeed, Maximum likelihood (ML) estimation of $(\boldsymbol{h}, \mathbb{a})$ is a bilinear inverse problem since (1) is linear w.r.t. $\boldsymbol{h}$ when $\mathbb{a}$ is fixed and vice-versa. Therefore, the ML solution $(\boldsymbol{h}^*, \mathbb{a}^*)$ is not unique. For instance, every couple $(\boldsymbol{h}^*/s, \mathbb{a}^* \times s)$ defines another pair of solutions in the ML sense whatever the scale parameter $s > 0$. To get rid of identifiability problems and reach a more reliable estimation, in the Bayesian formalism we introduce suitable prior distributions attached to the unknown quantities $(\boldsymbol{h}, \mathbb{a})$.

Priors

The Hemodynamic response function

Akin to (Buxton and Frank, 1997; Goutte et al., 2000; Marrelec et al., 2003), the HRF is characterised as follows: *(i)* its variations are smooth; *(ii)* it is causal and returns to a baseline after a given

time interval T ($h_t = 0$, $\forall \ t < 0$ and $t > T$); T is fixed by the user according to the experimental paradigm (usually 25 seconds).

Condition *(i)* may be fulfilled using an approximation of the second-order derivative $\|\partial^2 \boldsymbol{h}\|^2$:

$$\big(\partial^2 \boldsymbol{h}\big)_{d\tau} \approx \big(h_{(d+1)\tau} - 2h_{d\tau} + h_{(d-1)\tau}\big)/\tau^2, \ \forall d = 1:D-1.$$

In matrix form, we get $\partial^2 \boldsymbol{h} = \boldsymbol{D}_2 \boldsymbol{h}$.

Condition *(ii)* is ensured with a HRF $\boldsymbol{h}$ whose magnitude vanishes at first and last time points ($h_0 = h_D = 0$). Hence, $\boldsymbol{D}_2$ is the *truncated* second-order finite difference matrix of size $(D-1) \times (D-1)$ and $\|\partial^2 \boldsymbol{h}\|^2 = \boldsymbol{h}^t \boldsymbol{R}^{-1} \boldsymbol{h}$ with $\boldsymbol{R} = (\boldsymbol{D}_2^t \ \boldsymbol{D}_2)^{-1}$ a symmetric positive definite matrix. The prior on $\boldsymbol{h}$ thus reads $\boldsymbol{h} \sim \mathcal{N}(\boldsymbol{0}, \sigma_h^2 \boldsymbol{R})$. To overcome the scale ambiguity problem mentioned earlier, we constrain the HRF to be of unitary norm ($\|\boldsymbol{h}\| = 1$). Alternative constraints such as setting the value of the peak could be considered.

The "neural" response levels

Mixture models are often used as a second stage to segment the SPMs (i.e., the statistical maps) resulting from a first-level temporal analysis of fMRI time series (Vaever Hartvig and Jensen, 2000; Everitt and Bullmore, 1999; Woolrich et al., 2005). This means that the data to be classified correspond to some normalised effect $c^t \hat{\boldsymbol{\beta}}$ / std ($c^t \hat{\boldsymbol{\beta}}$), where vector $\boldsymbol{c}$ defines a contrast of interest (typically a comparison between two experimental conditions) and $\hat{\boldsymbol{\beta}}$ is the vector of parameter estimates after fitting a GLM against the fMRI data.

In the present paper, as well as in (Makni et al., 2005), prior mixture models are used in a different way, closer to that proposed by Svensen et al. (2000). In the same spirit, a mixture model is introduced on the NRLs for every experimental condition m and not specifically on the linear combination $c^t \hat{\boldsymbol{\beta}}$. In (Makni et al., 2006a), it was stressed that a two-class Gaussian mixture model (GMM) may be inadequate for segregating noise from true activations. In particular, it can be shown that this kind of independent mixture may degenerate in the sense that the two probability density functions (pdf) overlap almost entirely if there are not enough activating voxels in the current parcel (see (Makni et al., 2005, §VII.)). For this reason, we have rather adopted an inhomogeneous prior mixture

Table 1
Model definition and notations

	GaGMM	GMM	GaGGaMM
AR(1) noise	$\mathcal{M}_1$	$\mathcal{M}_3$	
white noise	$\mathcal{M}_2$	$\mathcal{M}_4$	$\mathcal{M}_5$

Here, AR(1) stands for a first-order autoregressive noise in time whose parameters vary in space. In this respect, it is referenced as a spatially-varying AR(1) noise. The second noise model under study is a spatially-varying white noise. The columns describe the different NRLs priors: GMM stands for a *two-class Gaussian* mixture model (a zero-mean Gaussian density (G) for non-activating voxels and a Gaussian density (G) for activating voxels). GaGMM stands for a *two-class gamma-Gaussian* mixture model (a centred Gaussian density (G) for non-activating voxels and a gamma density (Ga) for activating voxels). GaGGaMM stands for a *three-class mixture* model composed of a zero-mean Gaussian density and two gamma densities (a gamma density (Ga) for activating voxels and a flipped gamma density (Ga) for deactivating voxels).

model. Among several possibilities (Gaussian-lognormal MM, Gaussian-truncated Gaussian MM, …), a gamma-Gaussian mixture model (GaGMM) has been retained for technical reasons that will become clearer in what follows. The non-activating voxels are still modelled using a zero-mean Gaussian pdf while a gamma distribution is used to enforce positivity of activating voxels. Akin to (Vaever Hartvig and Jensen, 2000; Woolrich et al., 2005), a three-class mixture prior model is actually considered to account for deactivating voxels. Since we assume that deactivation corresponds to a negative BOLD response, we use a flipped gamma density defined on the left part of the real line leading to define the GaGGaMM extension (see Table 1).

In our model, different stimulus types are supposed to induce statistically independent hemodynamic magnitudes or NRLs *i.e.*, $p(\mathbb{a}|\boldsymbol{\theta}_\mathbb{a}) = \Pi_m p(\boldsymbol{a}^m|\boldsymbol{\theta}^m)$ with $\mathbb{a} = (\boldsymbol{a}^m)_{m=1:M}$, $\boldsymbol{a}^m = \left(a_j^m\right)_{j=1:J}$ and $\boldsymbol{\theta}_\mathbb{a} = \left\{\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^m\right\}$. Vector $\boldsymbol{\theta}^m$ denotes the set of unknown hyper-parameters related to the mth stimulus type. Because our mixture model is voxelwise, the prior pdf factors over voxels: $p(\boldsymbol{a}^m \mid \boldsymbol{\theta}^m) = \Pi_j p(a_j^m \mid \boldsymbol{\theta}^m)$. Importantly, the hyper-parameters are kept constant for all voxels in a given parcel because of the within-parcel homogeneity. These parameters may actually vary from one parcel to another. Let q_j^m be the *allocation variable* (the label) that indicates whether voxel V_j is activating ($q_j^m = 1$), deactivating ($q_j^m = -1$) or non-activating ($q_j^m = 0$) in condition m. The marginal density $p(a_j^m \mid \boldsymbol{\theta}^m)$ thus reads:

$$p\left(a_j^m|\theta^m\right) = \sum_{i=-1}^1 \Pr\left(q_j^m = i|\lambda_m\right) f\left(a_j^m|q_j^m = i, \boldsymbol{\theta}^m\right)$$
$$= \sum_{i=-1}^1 \lambda_{i,m} f_i\left(a_j^m|\theta^m\right), \qquad (3)$$

with $\boldsymbol{\lambda}_\mathrm{m} = (\lambda_{-1,m},\ \lambda_{0,m},\ \lambda_{1,m})$ and $f_0\left(a_j^m|\boldsymbol{\theta}^m\right) = \mathcal{N}\left(0, \upsilon_{0,m}\right)$, $f_{\pm 1}\left(a_j^m|\boldsymbol{\theta}^m\right) = \mathcal{G}\left(\alpha_{\pm 1,\ m}, \beta_{\pm 1,m}\right)$. The $\lambda_{i,m}$ parameters define the prior probabilities of the three-class mixture on the NRLs ($\Sigma_i \lambda_{i,m} = 1$). For instance, $\lambda_{1,m}$ gives us the prior probability of being activated in response to condition m. Since the mixture is independent in space, we have $\lambda_{1,m} = Pr(q_j^m = 1 \mid \boldsymbol{\lambda}_m)$, $\forall j$. Note that $q_j^m \mid \boldsymbol{\lambda}_m$ follows a multinomial distribution over the 3-dimensional probability simplex, *i.e.*, $q_j^m \sim \mathcal{MN}_3(1; \boldsymbol{\lambda}_m)$ (see Appendix A). Hence, seven hyper-parameters are necessary to describe the prior mixture for each experimental condition m:

$$\boldsymbol{\theta}^m = \left\{\lambda_{\pm 1,m}, \alpha_{\pm 1,m}, \beta_{\pm 1,m}, \upsilon_{0,m}\right\}.$$

Compared to (Woolrich et al., 2005), we set the mean of the non-activating class to zero ($\mu_{0,m} = 0$, $\forall\ m$), while we do not need to place restrictions on the mode of the activation and deactivation gamma classes.

The nuisance variables

We assume that $\mathbb{l}$ is a random process independent of $\boldsymbol{h}$ such that $p(\mathbb{l}; \sigma_\ell^2) = \Pi_j p(\ell_j; \sigma_\ell^2)$ and $\ell_j \sim \mathcal{N}\left(\boldsymbol{0},\ \sigma_\ell^2 \boldsymbol{I}_Q\right)$.

The hyper-parameters

All the hyper-parameters are concatenated into the overall parameter vector $\boldsymbol{\Theta} = \left\{\boldsymbol{\theta}_0, \sigma_h^2, \sigma_\ell^2, \boldsymbol{\theta}_\mathbb{a}\right\}$. Without informative prior knowledge, the following priors are retained for $(\sigma_h^2, \sigma_\ell^2, \boldsymbol{\theta}_0)$:

$$p\left(\sigma_h^2, \sigma_\ell^2\right) = (\sigma_h \sigma_\ell)^{-1}, p(\boldsymbol{\theta}_0) = \prod_{j=1}^J p\left(\rho_j \sigma_{\varepsilon_j}^2\right) = \prod_{j=1}^J \sigma_{\varepsilon_j}^{-1} u([-1,1])(\rho_j),$$
$$(4)$$

to ensure stability of the AR(1) noise process (Kay, 1988).

Mixture parameters. As regards variances $\upsilon_{0,m}$, an improper Jeffreys' prior $p(\upsilon_{0,m}) = \upsilon_{0,m}^{-1/2}$ is considered because we do expect non-activating voxels in a given parcel. Hence, class 0 should never be empty a priori. However, to avoid emptiness and subsequent degeneracy problems making the sampling of the posterior distribution of $\upsilon_{0,m}$ unfeasible, a conjugate prior could also be chosen, that is, an inverse gamma density $\mathcal{IG}\left(\upsilon_{0,m}, a_{\upsilon_0}, b_{\upsilon_0}\right)$, where $(b_{\upsilon_0}, c_{\upsilon_0})$ are fixed values chosen in an appropriate way to make the prior flat enough.

The non-negativity of parameters $\alpha_{i,m}$ is guaranteed through the use of an exponential density $\mathcal{E}\left(\alpha_{i,m}; s_i\right) \equiv \mathcal{G}\left(\alpha_{i,m}; 1, s_i\right)$ as prior distribution (see Appendix A). For parameters $\beta_{i,m}$ we resort to the conjugate prior, given by a gamma density $\mathcal{G}\left(\beta_{i,m}; b_i, c_i\right)$ for $i = \pm 1$.

Mixture probabilities. As regards mixture probabilities $\boldsymbol{\lambda}_m \in [0,1]^3$, a Dirichlet prior distribution is used as it is conjugate to the multinomial distribution used for labels, *i.e.*, $\mathcal{MN}_3\left(q_j^m|\boldsymbol{\lambda}_m\right)$. More exactly, a symmetric Dirichlet density $\mathcal{D}_3(\boldsymbol{\lambda}_m|\boldsymbol{\delta})$ is selected with $\boldsymbol{\delta} = \delta \boldsymbol{1}_3$ and $\delta > 0$ (see Appendix A).

The full prior density $p(\boldsymbol{\theta}^m)$ thus reads:

$$p(\boldsymbol{\theta}^m) = \upsilon_{0,m}^{-1/2} \frac{\Gamma(3\delta)}{3\Gamma(\delta)} \prod_{i=\pm 1} \lambda_{i,m}^{\delta-1} s_i \frac{c_i^{b_i}}{\Gamma(b_i)} \beta_{i,m}^{b_i-1} \exp\left(-s_i \alpha_{i,m} - c_i \beta_{i,m}\right).$$
$$(5)$$

Values of $(a_{\pm 1}, b_{\pm 1}, c_{\pm 1}, s_{\pm 1}, \delta)$ are fixed empirically but do not really influence the results in most cases[4]. These parameters make the sampling steps of $(\alpha_{\pm 1,m}, \beta_{\pm 1,m})$ always possible even when one of the two classes ± 1 is empty, because the hyper-prior densities have been chosen proper.

The full posterior distribution

Combining data-driven information in each parcel with prior knowledge using Bayes' rule, we get the full posterior distribution, which is the keystone both for localising activations and

---

[4] Except potentially when the corresponding class is empty: $J_{i,m} = 0$ for $i = \pm 1$.

deactivations as well as for estimating the corresponding parcel-based HRF:

$$p(\boldsymbol{h}, \mathbb{a}, \mathbb{l}, \Theta | \mathbb{y}) \propto p(\mathbb{y} | \boldsymbol{h}, \mathbb{a}, \mathbb{l}, \boldsymbol{\theta}_0) p(\mathbb{a} | \boldsymbol{\theta}_{\mathbb{a}}) p(\boldsymbol{h} | \sigma_{\boldsymbol{h}}^2) p(\mathbb{l} | \sigma_{\ell}^2) p(\Theta)$$

$$\propto \sigma_{\boldsymbol{h}}^{-D} \sigma_{\ell}^{-JQ} \prod_{j=1}^{J} \left(\frac{\left(1 - \rho_j^2\right)^{1/2}}{\sigma_{\varepsilon_j}^{N+1}} \mathbb{1}_{(-1,1)}(\rho_j) \right) \times \ldots$$

$$\times \exp\left(-\frac{\boldsymbol{h}^t \boldsymbol{R}^{-1} \boldsymbol{h}}{2\sigma_{\boldsymbol{h}}^2} - \sum_{j=1}^{J} \left(\frac{1}{2\sigma_{\varepsilon_j}^2} \widetilde{\boldsymbol{y}}_j^t \Lambda_j \widetilde{\boldsymbol{y}}_j + \frac{1}{2\sigma_{\ell}^2} ||\ell_j||^2 \right) \right)$$

$$\times \prod_{m=1}^{M} \left(p(\boldsymbol{\theta}^m) \prod_{j=1}^{J} p(\mathbb{a}_m^j | \boldsymbol{\theta}^m) \right) \tag{6}$$

where $p(a_j^m | \boldsymbol{\theta}^m)$ and $p(\boldsymbol{\theta}^m)$ are defined by (3) and (5), respectively.

Note that the parcel-based HRF $\boldsymbol{h}$ can be identified if at least one voxel elicits activation in response to one or several experimental conditions involved in model(1). In addition, other identifiability problems may occur on hyper-parameters such as the mean and variance parameters. It is necessary that at least two voxels belong to each class in order to properly estimate the variances attached to the mixture components. In practice, there is no numerical problem because of the choice of proper priors for the hyper-parameters; see Subsection 3.1.2 for practical details.

Inference scheme

Our objective is to obtain an estimate of the joint posterior distribution of all unknown parameters, given the observed data. Exact and analytical approaches are not feasible with non-Gaussian models such as (6). Several competing inferential schemes are possible. For instance, approximations to the full posterior distribution can be derived in the Variational Bayes (VB) framework or using Taylor series expansion. In our context, given the bilinear structure of the generative BOLD model (6), the VB formulation would be feasible only at the expense of separability assumptions between $\mathbb{a}$ and $\boldsymbol{h}$ in the approximation of the posterior distribution. Further work is required to decide whether or not this hypothesis is tenable. Instead, we resort to a more computationally demanding but exact approach to simulate realisations of the full posterior distribution.

Gibbs sampling algorithm

To draw realisations of the full posterior distribution, a Gibbs sampler is implemented. This consists in building a Markov chain, whose stationary distribution is the joint posterior pdf (6), by sequentially generating random samples from the full conditional pdfs of all the unknown parameters and hyper-parameters; see (Liu, 2001; Robert, 2001) for a general introduction to MCMC.

As shown in Appendix B, direct sampling according to the full conditional distributions is only feasible for the HRF $\boldsymbol{h}$, the labels $\mathbb{q}$, the nuisance variables $\mathbb{l}$, the noise variances $\boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}$, the mixture probabilities $\boldsymbol{\lambda}_m$, and part of the hyper-parameters (scales σ_h and σ_ℓ, class 0 variances υ_0 and shape parameters β_i for $i = \pm 1$). In contrast, direct simulation is not tractable for the other parameters, *i.e.,* the NRLs $\mathbb{a}$ corresponding to classes ± 1, the AR parameters $\boldsymbol{\rho}$ and the scale parameters $\boldsymbol{\alpha}_i$ of the gamma densities for $i = \pm 1$. Therefore, single-component Metropolis-Hastings jumps (Hastings, 1970) are specifically designed. More precisely, separate jumps are proposed

for each of the parameters in turn. To this end, suitable instrumental distributions regarding the parameters of interest are designed (see Appendix B for details).

Initialisation

Parameters are uniformly initialised. This means that we set up all voxels with the same noise statistical parameters ($\theta_{0,j} = \theta_0, \forall j$) and that we use the same starting values of mixture hyper-parameters ($\boldsymbol{\theta}^m = \boldsymbol{\theta}^*, \forall m$). In the first parcel, the HRF is initialised to the canonical shape (Glover, 1999). In the next ones, the HRF is set up using the mean of the estimates computed over the already processed neighbouring parcels. We resort to the same strategies for the labels and the corresponding NRLs when the parcel sizes match approximately. We have checked that this strategy provides shorter burn-in periods[5] and thus reduces the computation load.

Identifiability issues

To cope with these identifiability problems, we have carried out the following three steps procedure over the first iterations of our MCMC algorithm:

- initialise each parcel-specific HRF with a fixed shape in order to obtain a first estimate of labels $\hat{\mathbb{q}}$;
- check that the class of activating voxels is effectively not empty for at least one experimental condition in the current parcel $\mathcal{P}$:
- -- If $\exists\, m \in \mathbb{N}_M^* = \{1, \ldots, M\}$ such that $\exists\, j \in \mathcal{P} | q_j^m = 1$ then release the HRF constraint to estimate the complete model *i.e.,* including the HRF shape;
- -- otherwise, discard the current parcel: the HRF estimate is not reliable in $\mathcal{P}$. Since the corresponding NRLs are close to zero in that case there is no evoked activation due to the experimental paradigm.

Convergence diagnosis

We use a burn-in period of 500 iterations, followed by 1000 subsequent jumps and compute PM and MAP estimates every two jumps. Observations of the chain with different initial conditions confirmed that a burn-in of 500 jumps was sufficient. In addition, convergence has been checked by monitoring on-line the behaviour of the estimated values of some *scalar* parameters (*e.g.,* noise variances, AR parameters, …) from one iteration to another. These observations confirmed also that 1000 iterations were sufficient.

Computational load and parallel implementation

Our current implementation (PyHRF package) is in Python, while the most intensive computations (*e.g.,* computation of the inverse covariance matrix of $\boldsymbol{h}$) have been coded in C-language and interfaced with the Gnu Scientific Library (GSL)[6]. This allows us to take advantage of a parallel computing system available through the Seppo library (Simple Embarrassingly Parallel Python[7]) and the Pyro (Python Remote Object) server. Using such a system, all the parameter estimates are obtained in about 2 mn for a parcel of mean size (250 voxels) for two experimental conditions ($M = 2$).

---

[5] The burn-in period is the starting part of the Markov chain built by any MCMC algorithm which is used to ensure that the subsequent samples follow the equilibrium target distribution, *i.e.,* the posterior law.

[6] http://www.gnu.org/software/gsl.

[7] see http://www.its.caltech.edu/~astraw/seppo.html.

Since about 200 parcels are necessary to cover the grey matter's mask, a complete within-subject analysis takes about 2 hours when running four processes on a dual core bi-processors Pentium IV (2.7 GHz). PyHRF will be available in the next release of BrainVisa[8] in April 2008.

Derivation of parcel-based summaries

After convergence of the MCMC algorithm in each parcel $\mathcal{P}$, the samples of the quantities of interest are averaged over iterations to compute approximations of marginal posterior expectations:

$$\hat{x}_{\mathcal{P}}^{\mathrm{PM}} = \sum_{k=L_0}^{L_1} x^{(k)}/L, \quad L = L_1 - L_0 + 1, \quad \forall x \in \{h, \mathrm{a}, \mathbb{l}, \Theta\}, \qquad (7)$$

where L_0 stands for the length of the burn-in period and L the effective number of iterations. For classification purpose, we proceed in two steps:

1. Compute the PM estimates $(\overline{p}_j^m)_i$ of $Pr(q_j^m = i \,|\, \mathbf{y}_j)$ for $i = -1, 0$ using the following expression:

$$\left(\overline{p}_j^m\right)_i = \sum_{k=L_0}^{L_1} I\left[\left(q_j^m\right)^{(k)} = i\right]/L, \qquad (8)$$

where I stands for the identity function. Then, deduce $(\overline{p}_j^m)_1$ from the constraint of unitary probability mass: $(\overline{p}_j^m)_1 = 1 - (\overline{p}_j^m)_{-1} - (\overline{p}_j^m)_0$.

2. Sort the probabilities $(\overline{p}_j^m)_i$ and select the MAP estimate:

$$\left(\hat{q}_j^m\right)^{\mathrm{MAP}} = \arg\max_i Pr\left(q_j^m = i | \mathbf{y}_j\right) \approx \arg\max_i \left(\overline{p}_j^m\right)_i. \qquad (9)$$

whatever the number of components in the mixture. The MAP estimator is easily obtained in the two-class mixture case: V_j is non-activating $((\hat{q}_j^m)^{\mathrm{MAP}} = 0)$ for the mth condition if $(\overline{p}_j^m)_1 < 0.5$.

In combination with these PM estimates, one can attach uncertainty measures to the NRLs. More precisely, the error bars are derived as follows:

$$e_j^m = \sum_{k=L_1}^{L_2} \left(\sigma_{i,j}^m\right)^{(k)}/L \text{ with } i = \left(\hat{q}_j^m\right)^{MAP}. \qquad (10)$$

Interestingly, $\sigma_{0,j}^m$ is directly given by $\sqrt{v_{0,j}^m}$ since the full conditional posterior distribution of the zero class is Gaussian, *i.e.,* $\mathcal{N}\left(\mu_{0,j}^m, v_{0,j}^m\right)$. In contrast, the standard deviations (SD) $\sigma_{\pm 1,j}^m$ require further computation since these full conditional densities are gamma-Gaussian (see Section A.5). As derived in Eq. (A.13), the variance of a gamma-Gaussian density admits a closed form expression, which gives $\sigma_{\pm 1,j}^m$ after taking the square root. These SD estizmates are then plugged into (10) to get corresponding error bars e_j^m.

The stochastic algorithm is summarised in Table 2.

Table 2
Gibbs sampling algorithm in a given parcel $\mathcal{P}$

- Setting up: choose $h^0, \mathrm{a}^0, \mathbb{l}^0, \boldsymbol{\theta}^0, \boldsymbol{\theta}_a{}^0$.
- Iteration k: draw samples $h^k, \mathbf{a}^k, \boldsymbol{\lambda}^k, (\varepsilon^2)^k, \boldsymbol{\theta}_a^k$ from the conditional posterior pdfs:
  - -- HRF: $h^k \sim \mathcal{N}(\boldsymbol{\mu_h}, \boldsymbol{\Sigma_h})$.
  - -- HRF variance: $(\sigma_h^2)^k \sim \mathcal{IG}(D/2, h^t \mathbf{R}^{-1} h/2)$.
  - -- NRLs: for every condition m and every voxel j,
    - -- $(u_j^m)^k \sim \mathcal{U}[0, 1]$; if $(u_j^m)^k \leqslant \lambda_{-1,j}^m$, then $(q_j^m)^k = -1$ else if $(u_j^m)^k \leqslant \lambda_{-1,j}^m + \lambda_{0,j}^m$ then $(q_j^m)^k = 0$, otherwise $(q_j^m)^k = 1$.
    - -- $(a_j^m)^k | (q_j^m)^k = 0 \sim \mathcal{N}(\mu_{0,j}^m, v_{0,j})$.
      $(a_j^m)^k | (q_j^m)^k = \pm 1 \sim \mathcal{GN}(a_j^m | \alpha_{\pm 1,m}, \mu_{\pm 1,j}^m), v_{\pm 1,j})$.
  - -- drift coefficients: $\forall_j, (\ell_j)^k \sim \mathcal{N}\left(\boldsymbol{\mu}_{\ell_j}, \sum_{\ell_j}\right)$
  - -- Noise variances: $\forall_j, \left(\sigma_{\varepsilon_j}^2\right)^k \sim \mathcal{IG}\left((N+1)/2, \|\widetilde{\mathbf{y}}_j\|_{A_j}^2/2\right)$.
  - -- AR parameters: $\forall_j, (\rho_j)^k \sim \sqrt{1 - \rho_j^2} \exp\left(-\frac{A_j}{2\sigma_{\varepsilon_j}^2}\left(\rho_j - \frac{B_j}{A_j}\right)^2\right) \mathbb{1}_{(-1,1)}(\rho_j)$.
  - -- Mixture parameters: for every condition m,
    - -- Weighting probabilities λ_m:

      $$(\lambda_m)^k \sim \mathcal{D}(\delta'), \text{ with } \delta'_i = \delta + \underbrace{\mathrm{Card}\left[C_{i,m} = \left\{ j \in 1 : J | q_j^m = i\right\}\right]}_{= J_{i,m}},$$

      $\forall i = -1 : 1$.
    - -- Variance of NRLs for non-activating voxels: $(v_{0,m})^k \sim \mathcal{IG}(\eta_{0,m}^k, v_{0,m}^k)$.
    - -- Shape parameters: $(\alpha_{\pm 1,m})^k \sim \exp(J_{i,m}\tau_{i,m}\alpha_{i,m})/\Gamma(\alpha_{i,m})^{J_{i,m}} \mathbb{1}_{\mathbb{R}_+}(\alpha_{i,m})$.
    - -- Scale parameters: $(\beta_{\pm i,m})^k \sim \mathcal{G}\left(J_{i,m}\alpha_{i,m} + b_i + 1, \sum_{j \in C_{i,m}} a_j^m + c_i\right)$.
- Iterate until convergence is achieved. PMEs of $\{h, \mathrm{a}, \mathbb{l}, \boldsymbol{\theta}_\alpha\}$ are computed using (7).
- Classification is performed according to the MAP criterion using (8)-(9).

The parameters of the sampled distributions are derived in Appendix B.

Statistical comparisons for cognitive interpretation

Akin to the contrast definition in any GLM-based approach, statistical comparison between our task-related NRL estimates can be addressed in the proposed formalism. One might be interested in assessing *unsigned* or *signed* differences like using Fisher or Student-t tests, respectively in the classical hypothesis testing framework.

Let m and m' be the indexes of the conditions we plan to contrast across the brain. This contrast can be assessed by measuring how close the voxelwise marginal distributions $(p_j^m, p_j^{m'})$ of the NRLs $(a_j^m, a_j^{m'})$ are in every voxel V_j. Since these densities write as posterior mixtures, say $p_j^m = \Sigma_i \pi_i f_{i,j}^m$, we start with identifying the MAP estimates $(\hat{q}_j^m, \hat{q}_j^{m'})$ and then we compare the full conditional posterior densities $(f_{\hat{q}_j^m,j}^m, f_{\hat{q}_j^{m'},j}^{m'})$ instead of computing a distance between p_j^m and $p_j^{m'}$. Hence, three different (respectively, six) situations may arise depending on the mixture prior in use (two or three-class mixture, respectively). The different cases correspond to all possible combinations of the pair $(\hat{q}_j^m, \hat{q}_j^{m'})$:

a. if $\hat{q}_j^m = \hat{q}_j^{m'} = -1$, voxel V_j generates deactivations for both conditions. Comparing the NRLs $(a_j^m, a_j^{m'})$ is achievable by measuring how close $(f_{-1,j}^m, f_{-1,j}^{m'})$ are. This comparison therefore answers the question of deciding whether or not the deactivation is stronger for one condition w.r.t. the other (signed comparison) or if there is any difference between the two conditions (unsigned comparison).

b. if $\hat{q}_j^m = \hat{q}_j^{m'} = 0$, voxel V_j is non-activating for both conditions. Comparing the NRLs $(a_j^m, a_j^{m'})$ amounts to computing a criterion between $(f_{0,j}^m, f_{0,j}^{m'})$. The interesting comparison consists in

deciding whether or not there is some difference in the non-activating profile.

c. if $\hat{q}_j^m = \hat{q}_j^{m'} = 1$, both conditions elicit activations in V_j. By measuring how close $(f_{1,j}^m, f_{1,j}^{m'})$ are, we hope to know if activation occurring for condition m or m' is stronger or if there is any difference irrespective of its sign.

d. if $\hat{q}_j^m = -1$ and $\hat{q}_j^{m'} = 0$, V_j is deactivating in response to the mth stimulus type but is non-activating in response to the m'th condition. To quantify this decision, one can measure a signed or unsigned criterion between $(f_{-1,j}^m, f_{0,j}^{m'})$. By symmetry this case is equivalent to $\hat{q}_j^m = 0$ and $\hat{q}_j^{m'} = 1$.

e. if $\hat{q}_j^m = -1$ and $\hat{q}_j^{m'} = 1$ or vice-versa, V_j is activating in condition m' and deactivating in condition m. To quantify this decision, one can measure a signed or unsigned distance between $(f_{-1,j}^m, f_{1,j}^{m'})$.

f. if $\hat{q}_j^m = 0$ and $\hat{q}_j^{m'} = 1$ or vice-versa, V_j is activating in condition m' and non-activating in condition m. To quantify this decision, one can measure a signed or unsigned criterion between $(f_{0,j}^m, f_{1,j}^{m'})$.

Due to the use of mixture models, these comparisons can allow us to assess the null hypothesis ($H_0 : a_j^m = a_j^{m'}$) or the alternative one (*e.g.*, $H_1 : a_j^m \neq a_j^{m'}$ or $H_1 : a_j^m < a_j^{m'}$) depending on the computed criterion. The question is now to define what kind of signed or unsigned criteria we can implement to quantitatively discriminate the two underlying distributions $(f_{i,j}^m, f_{i',j}^{m'})$.

Unsigned task comparison

Unsigned comparison between $f_{i,j}^m$ and $f_{i',j}^{m'}$ can be computed using the Kullback-Leibler (KL) divergence *i.e.*,

$$D\left(f_{i,j}^m \| f_{i,j}^{m'}\right) = \int_{\mathbb{R}} f_{i,j}^m(a) \log \frac{f_{ij}^m(a)}{f_{ij}^{m'}(a)} da.$$

In the present case, its exact computation is only feasible when the two distributions are Gaussian *i.e.*, when $i = i' = 0$ (case b); see (A.3) in Appendix A for details. Otherwise, an approximation of $D(\cdot\|\cdot)$ has to be derived. For doing so, we proceed as follows. In cases (a, c, e), the sampling step of the NRLs ($a_j^m, a_j^{m'}$) relies on two Metropolis jumps, one for each NRL. The corresponding instrumental laws are truncated normal distributions (see (A.4) in Appendix B). Therefore, we approximate f_{ij}^m and $f_{i'j}^{m'}$ by these positive Gaussian distributions which mean and variance parameters are given in (A.5)-(A.6). We end up by applying the KL divergence formula (Eq. (A.3)) to these truncated Gaussian approximations. In cases (*d, f*), we proceed similarly for the single activating or deactivating component.

Signed task comparison

To go one step further and recover a sign information regarding the difference $d_j^{m-m'} = a_j^m - a_j^{m'}$, we need to estimate its posterior probability distribution $f_j^{m-m'}$ from a histogram H_j^B (.) with B time bins $(\beta_b)_{b=1:B}$ constructed over the last 500 iterations (*i.e.*, the generated values $(d_j^{m-m'})^{(k)} = (a_j^m)^{(k)} - (a_j^{m'})^{(k)}$ in any voxel of the mask $\mathcal{M}_f$. The posterior cumulative distribution function (cdf) $F(\cdot)$ can then be easily estimated from H_j^B (.). Contrast-based posterior probability maps (PPMs) are thus given by looking at differences $d_j^{m-m'}$ above a given threshold α:

$$P\left(d_j^{m-m'} > \alpha\right) = 1 - F\left(d_j^{m-m'} \leqslant \alpha\right) = 1 - \int_{-\infty}^{\alpha} f_j^{m-m'}(t)\, dt \quad (11)$$

$$\approx 1 - \sum_{n=1}^{d} H_j^B \left(\frac{\beta_n + \beta_{n+1}}{2}\right) \Delta\beta \text{ with } d < \alpha \leqslant d+1, \quad (12)$$

where $\Delta\beta = \beta_{n+1} - \beta_n$. Setting $\alpha = 0$, we actually find the voxels where $(a_j^m) > (a_j^{m'})$. Finally, we can threshold $P(d_j^{m-m'} > \alpha)$ at level η to retain the voxels which make the comparison significant at this level (*e.g.*, $\eta = 0.95$). Formally, the thresholded PPMs are given by $P(d_j^{m-m'} > \alpha) > \eta$. Note that this only provides uncorrected PPMs for multiple comparisons. The control of the familywise error is an open issue in the Bayesian formalism and is beyond the scope of this paper.

Results on synthetic data

Goal of the study

A comparison between two different prior mixture models has been done in (Makni et al., 2006a). In short, it has been shown that the gamma-Gaussian mixture model (*GaGMM*) introduced on the NRLs is more efficient than a two-class Gaussian mixture model (*GMM*) in terms of specificity: it provides a better control of the false positive rate. Similar conclusions have been drawn in (Makni et al., 2006b) when considering an AR(1) noise model instead of a white Gaussian one in combination with a *GMM* prior. As the two changes induce higher computation time, it is worth assessing which modelling effort is preferable i.e. leads to the more significant improvement: the introduction of an inhomogeneous prior mixture or the consideration of serial correlation. For doing so, the models described in Table 1 are tested on the same artificial fMRI dataset.

Artificial fMRI dataset

These data were obtained by first generating two sets of trials, each of them corresponding to a specific stimulus ($M = 2$). These binary time series were then multiplied by a stimulus-dependent scale factor. Here, the functionally homogeneous region $\mathcal{P}$ consisted of $J = 60$ voxels. The number of activating voxels $J_{1,m}$ was varied with the stimulus type m according to $(J_{1,1}, J_{1,2}) = (22, 30)$. Positive NRLs corresponding to activating voxels were simulated according to gamma pdfs:

activating voxels : $\boldsymbol{a}_j^1 \sim \mathcal{G}(\alpha_1 = 3, \beta_1 = 1)$, $\boldsymbol{a}_j^2 \sim \mathcal{G}(\alpha_2 = 10, \beta_2 = 2)$,
non-activating voxels : $\boldsymbol{a}_j^{1,2} \sim \mathcal{N}(0, v_{0,m} = 0.1)$.

Remark that the chosen gamma parameter values yields a lower SNR for condition 1 ((μ_1, v_1) = (3, 3) vs. (μ_2, v_2) = (5, 2.5)). For all voxels, the binary stimulus sequence was convolved with the canonical HRF $\boldsymbol{h}_c$, whose exact shape appears in Fig. 2(a) in ■-line. An AR(1) noise $\boldsymbol{b}_j$ was then added to the stimulus-induced signal $\Sigma_m a_j^m \boldsymbol{X}^m \boldsymbol{h}$ in every voxel V_j. All AR parameters were set to the same value: $(\rho_j)_{j=1:J} = 0.4$, which is compatible with the serial correlation observed on actual fMRI time series. Also, a low SNR (SNR = 0.3) was considered in our simulations, in conformity with the real situation. Space-varying low-frequency drifts $\boldsymbol{P}\ell_j$ (generated from a cosine transform basis with coefficients ℓ_j drawn from a normal distribution) were also added to the fMRI time courses according to (1).

General comments

As shown in Figs. 2–5(a), all HRF estimates obtained using the four different models match the canonical time course $\boldsymbol{h}_c$ pretty
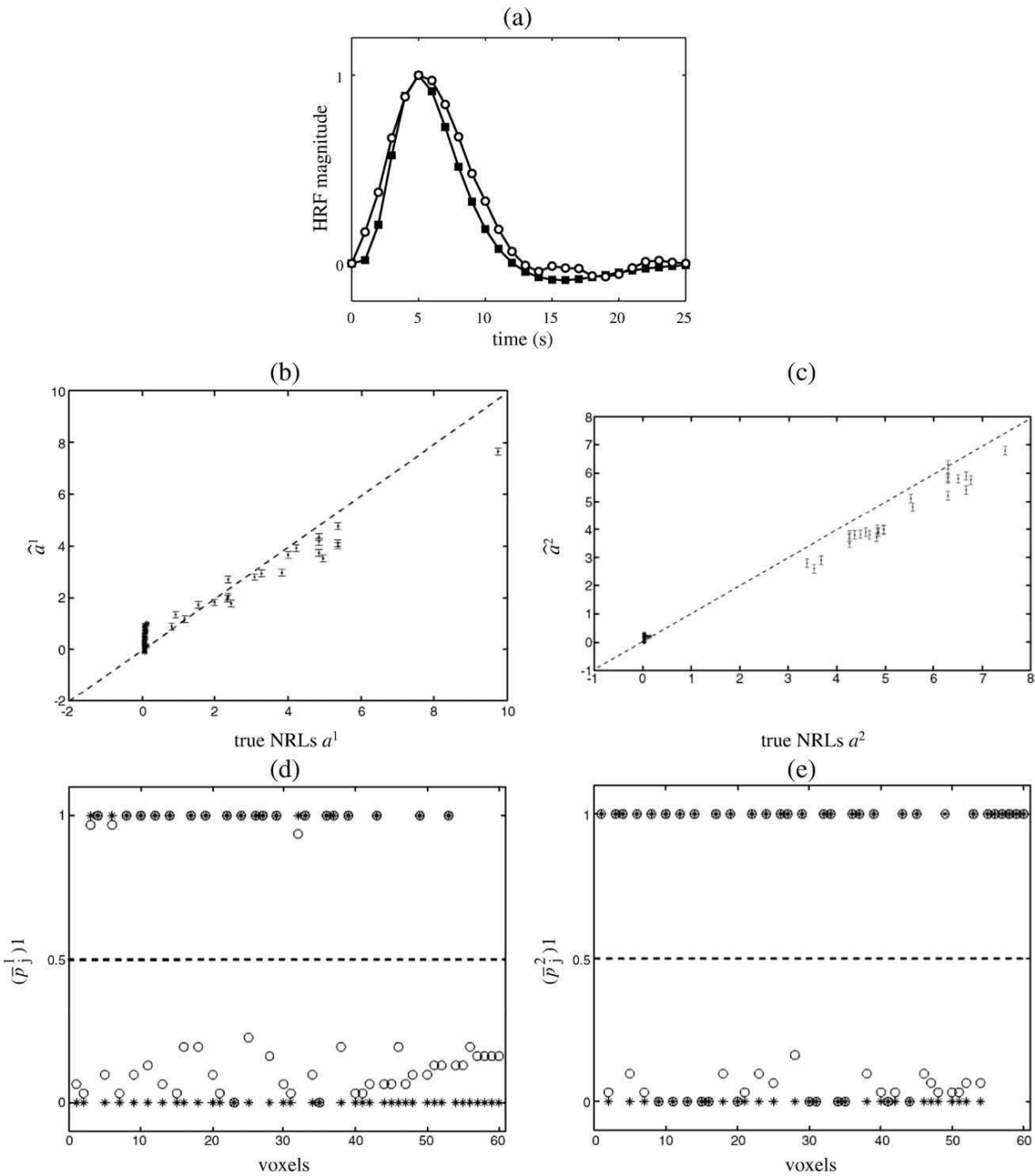
Fig. 2. Estimation results on the simulated data using model $\mathcal{M}_1$. (a) HRF results: Symbols ■ and ○ represent the true h_c and its corresponding HRF estimate, respectively. (b)-(c): NRL estimates for conditions 1 and 2, respectively. True values appear on the x-axis and estimated values on the y-axis. The error bars follow Eq. (10). (d)-(e): PM estimates of activation probabilities $\bar{p}_j^m$ (○ symbols) for the conditions 1 and 2, respectively. Symbols * depict the true class attached to each voxel.

well. Figs. 2–5(b) show the corresponding NRL estimates that we obtained from models $\mathcal{M}_1 - \mathcal{M}_4$, respectively in response to condition 1 while Figs. 2–5(c) summarize the same results for condition 2.

Since the artificial fMRI time courses were synthesised using a *GaGMM* prior and some correlated noise, it is not surprising that

the estimation performed under model $\mathcal{M}_1$ provides the most accurate NRL estimates. Let us remark that the NRL estimates have a small but not negligible amount of bias, which is due to the bias/variance trade-off arising in the Bayesian approach in the non-asymptotic case. Nonetheless, we have checked that the bias tends to zero when the SNR increases.

Influence of the noise model

Figs. 2–3(b)-(c) illustrate the impact of the noise model: a more accurate estimation of the NRLs, with smaller error bars and lower mean square error, is observed in Figs. 2(b)-(c) compared to Figs. 3(b)-(c), that is for model $\mathcal{M}_1$ compared to model $\mathcal{M}_2$. This is a direct consequence of accounting for serial correlation in $\mathcal{M}_1$. The same conclusion holds when looking at Figs. 4 and 5(b)-(c), so irrespective of the prior mixture type. As regards the HRF estimate (compare Figs. 2 and 3(a)), the noise

model has only little influence on the recovered shape, as already advocated in (Marrelec et al., 2003). As regards AR parameters, the estimated first order coefficients $(\rho_i)_i$ are close to the true values in every voxel for both models $\mathcal{M}_1 - \mathcal{M}_3$ (results not shown).

We also assessed the sensitivity and the specificity of the four models. Figs. 2–5(d)-(e) show the posterior mean estimates $(\bar{p}_j^m)_1$ of deciding that voxel V_j lies in class 1, *i.e.*, is activating for models $\mathcal{M}_1 - \mathcal{M}_4$ and conditions 1 and 2, respectively. These results confirm
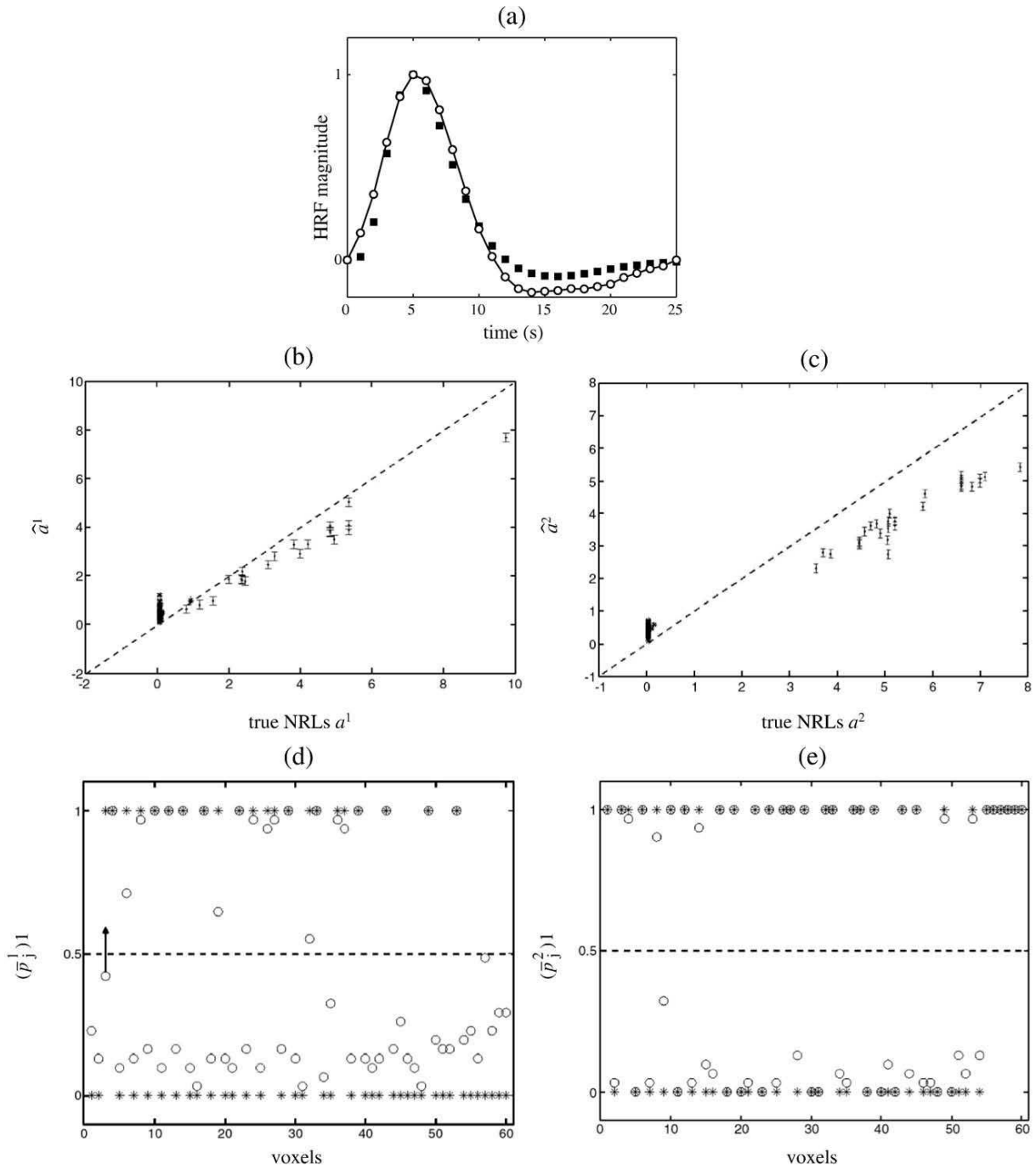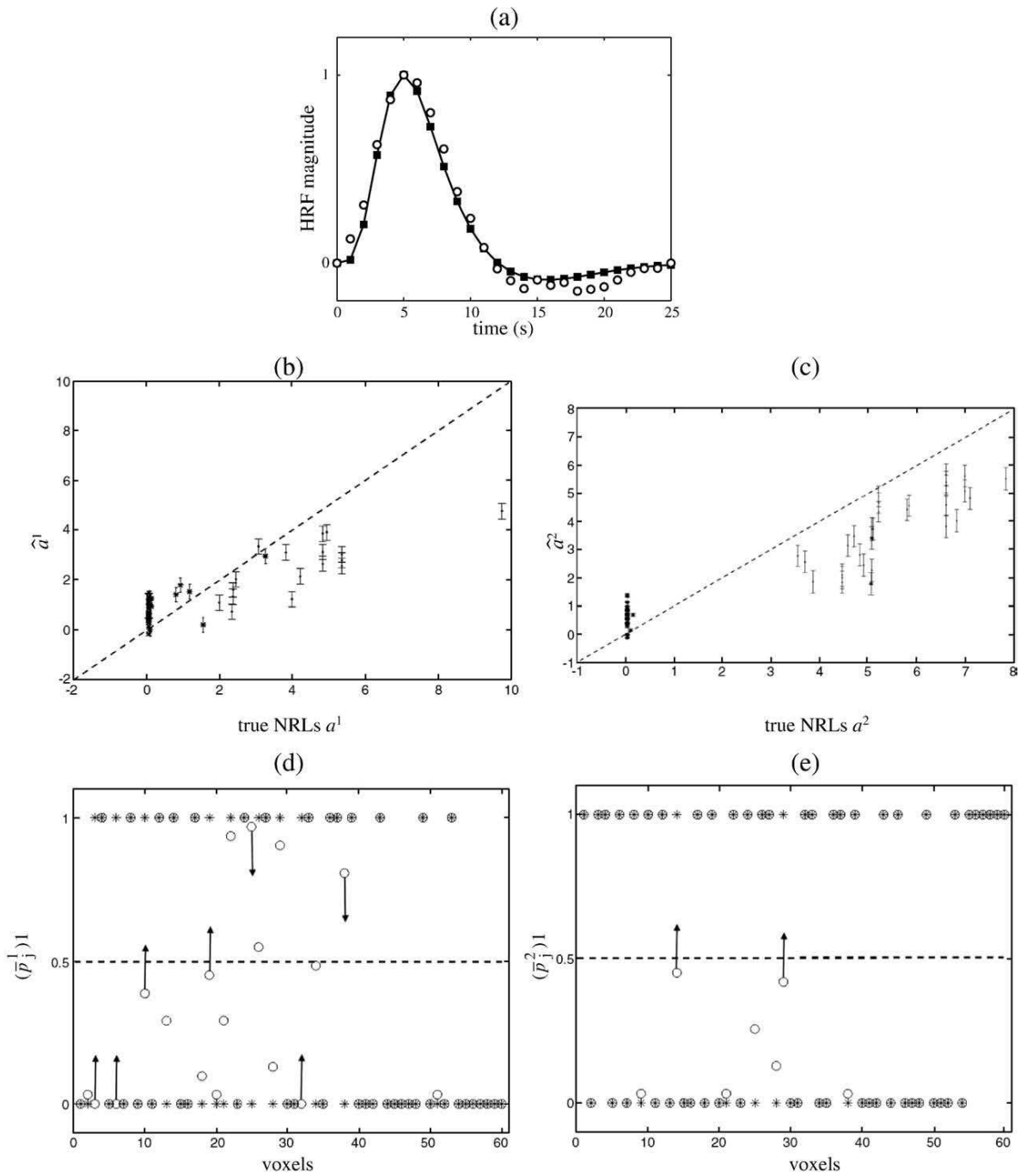

Fig. 3. Simulation results using model $\mathcal{M}_2$. The same legend as in Fig. 2 holds. Only one FN voxel is present, indicated with an upward arrow.

our expectations: the modelling of the temporal correlation significantly improves both the sensitivity and the specificity. A higher/lower value of $(\bar{p}_j^m)_1$ is obtained with $\mathcal{M}_1 - \mathcal{M}_3$ when V_j is truly activating/non-activating (compare Figs. 2 and 3(d)-(e) for *GaGMM* priors or Figs. 4 and 5(d)-(e) for *GMM* priors). This means that models $\mathcal{M}_1 - \mathcal{M}_3$ provide lower false positive (FP) and false negative (FN) rates than models $\mathcal{M}_2 - \mathcal{M}_4$, respectively. This effect is stronger in condition 1. This is in agreement with the idea that the precision of the noise model plays a more important role at a lower SNR.

Influence of the mixture prior

Not surprisingly, the estimated NRLs are recovered more accurately using the true prior mixture ($\mathcal{M}_1 - \mathcal{M}_2$): compare Figs. 2–4(b)-(c) one to another for an AR(1) noise model and observe the difference in Figs. 4 and 5(b)-(c) for a white noise model. This effect is much more important at low SNR, *i.e.,* for condition 1. However, we have checked that when the true NRLs of the activating voxels follow a Gaussian distribution, the estimated shape and scale parameters of the gamma density in the *GaGMM* mixture provide close estimates
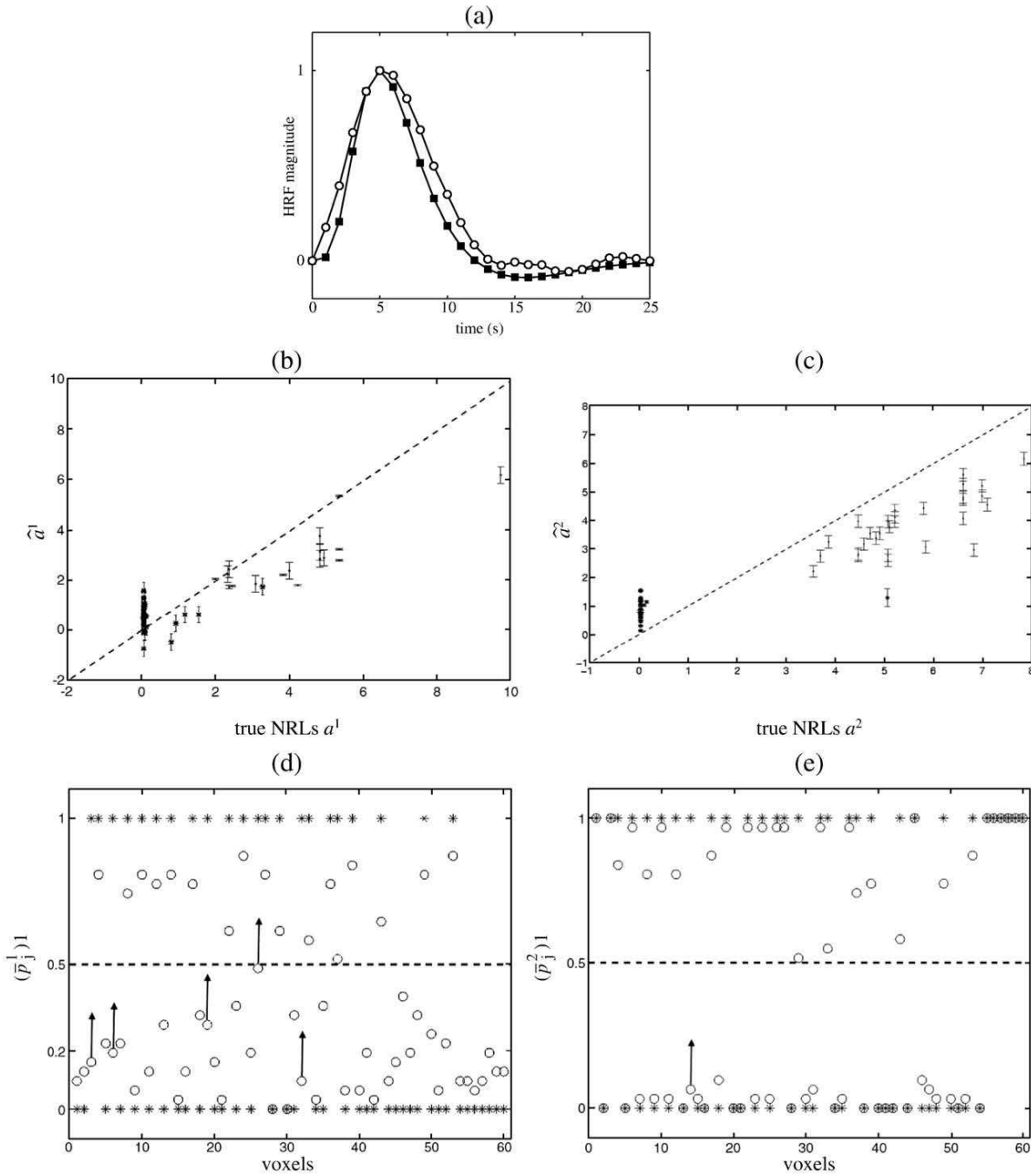
Fig. 4. Simulation results using model $\mathcal{M}_3$. The same legend as in Fig. 2 holds. FN voxels are indicated by upward arrows.

Fig. 5. Simulation results using model $\mathcal{M}_4$. The same legend as in Fig. 2 holds. FP and FN voxels are indicated by downward and upward arrows, respectively.

of the mean and variance parameters of the uncentered Gaussian distribution (results not shown).

We are now interested in assessing the differences between $\mathcal{M}_2$ and $\mathcal{M}_3$. The purpose of such a comparison is to decide whether or not a good mixture type provides more accurate and sensitive results than a precise noise modelling. Contrasting Figs. 3 and 4(b) allows us to note that $\mathcal{M}_2$ outperforms $\mathcal{M}_3$ in terms of accuracy of estimation for the first experimental condition. The NRLs attached to the non-activating voxels are over-estimated, leading to a much

larger bias. In case of high SNR arising for the second condition, the comparison of Figs. 3 and 4(b) is less clear. The small NRLs are still over-estimated but the large ones are better estimated using $\mathcal{M}_2$ in some cases (*e.g.,* voxels 27, 58, 60). In terms of detection, Fig. 3(d) shows that a single false negative (voxel 3) is retrieved by model $\mathcal{M}_2$ for condition 1, while five FNs are found by model $\mathcal{M}_2$, as shown in Fig. 4(d) (voxels 3, 6, 19, 26, 32). Hence, model $\mathcal{M}_2$ achieves better results in terms of sensitivity and specificity. Therefore, we conclude that introducing an inhomogeneous prior

mixture is more powerful than modelling the serial correlation as regards both estimation and detection.

Receiver-operator-characteristic (ROC) curves have been also computed to quantitatively evaluate the differences between models $\mathcal{M}_1 - \mathcal{M}_4$. Fig. 6 illustrates and confirms that model $\mathcal{M}_1$ provides the most sensitive detection when specificity is fixed and a better specificity at a given sensitivity. These ROC curves also validate that model $\mathcal{M}_4$ is the less sensitive and the less specific out of the four models. Finally, model $\mathcal{M}_2$ outperforms $\mathcal{M}_3$ and provides better results in terms of sensitivity and specificity, irrespective of the stimulus type. Figs. 6(a)-(b) allows us to claim again that the noise model has a stronger impact in detection at low SNR since the distance between continuous and dotted lines is larger in Fig. 6(a) than in Fig. 6(b), except at very low specificity (0.1). This holds whatever the mixture type.

Deactivation modelling

Our purpose was to compare an inhomogeneous two-class mixture model with its three class extension. In the latter case, a third class is used to account for putative deactivation phenomenon arising for instance during sustained bursts of interictal epileptiform activity (Bagshaw et al., 2005; Bénar et al., 2006).

Suitable artificial fMRI datasets were simulated accordingly. We considered a ROI consisting of $J = 60$ voxels. Let $J_{-1,m}$, $J_{0,m}$ $J_{1,m}$

be respectively the number of deactivating, non-activating and activating voxels in response to condition m. We set $(J_{1,1}, J_{-1,1}) = (28, 19)$ and $(J_{1,2}, J_{-1,2}) = (32, 12)$, so that $(J_{0,1}, J_{0,2}) = (13, 16)$. We simulated the NRLs as follows:

activating voxels : $a_j^1 \sim \mathcal{G}(3, 1)$, $a_j^2 \sim \mathcal{G}(5, 2)$
non-activating voxels : $a_j^{1,2} \sim \mathcal{N}(0, 0.1)$
deactivating voxels : $-a_j^1 \sim \mathcal{G}(5, 4)$, $-a_j^2 \sim \mathcal{G}(5, 4)$

The same procedure as before (see §4.1.1) was applied to simulate artificial fMRI time series. The only difference concerns the noise type, which is white, Gaussian and homogeneous in space to save computation time ($\forall j$, $\sigma_j^2 = 0.3$). Hence, model $\mathcal{M}_2$ and $\mathcal{M}_5$ (see Table 1) were tested and compared in terms of estimation, detection performance and evidence.

The HRF estimates corresponding to models $\mathcal{M}_2$-$\mathcal{M}_5$ are shown in Figs. 7 and 8(a), respectively. These estimated time courses appear very close to the true HRF shape. Figs. 7 and 8(b)-(c) show the NRL estimates related to conditions 1 and 2, computed for model $\mathcal{M}_2$ and $\mathcal{M}_5$, respectively. First, we observe that $\mathcal{M}_2$ provides under-estimated NRLs for activating voxels but over-estimated parameters for deactivating ones, irrespective of the stimulus type. The estimated error bars also appear significantly larger when deriving from $\mathcal{M}_2$. In contrast, model $\mathcal{M}_5$ provides more reliable NRL estimates with smaller error bars, as illustrated in Figs. 8-(b)(c). Also, the mean square error is decreased for the NRLs corresponding to deactivating and non-activating voxels.

Figs. 7(d)-(e) demonstrates that model $\mathcal{M}_2$ reports a few FN voxels (see upward arrows). All these voxels have small NRL coefficients, inducing their assignment to class 0. More importantly, we observe that the truly non-activating and deactivating voxels are mixed in class 0, irrespective of the condition. Figs. 8 (d)-(e) reports the posterior mean estimates $(\bar{p}_j^m)_i$ (see (8)), which are then combined to get the final classification according to the MAP criterion $(\hat{q}_j^m)^{\text{MAP}}$ (see (9)). As indicated on these graphs, model $\mathcal{M}_5$ produces an accurate classification. Figs. 8(d)-(e) respectively show the presence of three FN voxels for condition 1 and only two FNs for condition 2. These classification errors could be explained by the low values taken by the true NRL coefficients in these voxels, making likely the assignment to class 0.

Finally, note that modelling the third class induces a higher computation time. In our simulations, inferring the parameters of models $\mathcal{M}_2$ and $\mathcal{M}_5$ takes about 6 and 11 minutes, respectively. If the ROI is large or if the experimental paradigm involves numerous conditions, it seems reasonable to start with a careful analysis of the paradigm to anticipate potential deactivations before inferring upon parameters of $\mathcal{M}_5$ instead of $\mathcal{M}_2$.

Bayesian model comparison

More formally, from a statistical point of view we compare models $\mathcal{M}_1 - \mathcal{M}_5$ by computing sample-based approximations to the model evidence $p(\mathbb{y} | \mathcal{M}_m)$. That allows us to derive Bayes factors BF_{mn} as ratios of model evidence (see Appendix C for computational details). Bayes factor provides us with good statistical summary for model comparison. As reported in Table 3, there is a strong evidence in favour of Model $\mathcal{M}_1$. More interestingly, our conclusion drawn from the parameter estimates are also confirmed when comparing $\mathcal{M}_2$ with $\mathcal{M}_3$ using Bayes factor (line 2, Table 3). This also holds for the comparison between the two-class and the three-class mixtures, $\mathcal{M}_2$ and $\mathcal{M}_5$ respectively (line 5, Table 3).
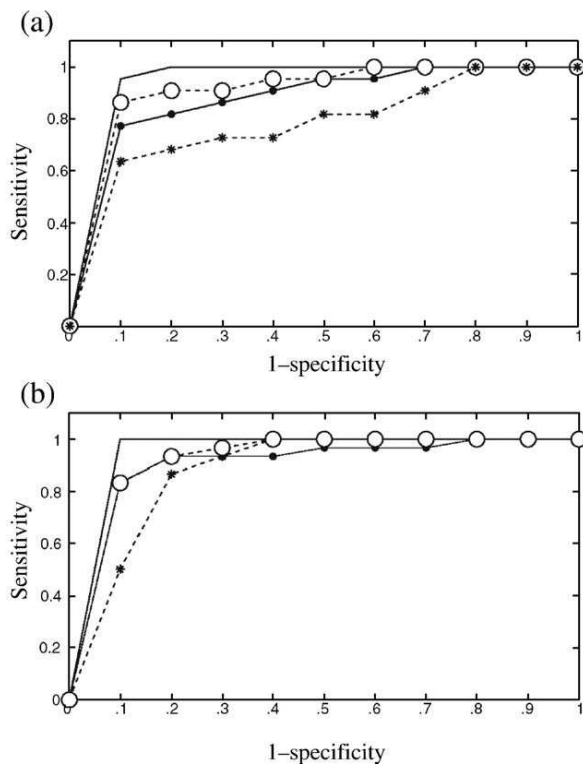
Fig. 6. (a)-(b): ROC curves associated to the four different models for condition 1 (a) and condition 2 (b), respectively. Continuous line, interrupted line with ○, continuous line with (a)-(b): ROC curves associated to the four different models for condition 1 (a) and condition 2 (b), respectively. Continuous line, interrupted line with ○, continuous line with ● and interrupted line with * represent the ROC curves for models $\mathcal{M}_1 - \mathcal{M}_4$, respectively.

Results on real fMRI data

fMRI experiment

MRI settings

The experiment was performed on a 3T whole-body system (Bruker, Germany) equipped with a quadrature birdcage radio frequency (RF) coil and a head-gradient coil insert designed for echo planar imaging (EPI). Functional images were obtained with a T2*-weighted GE-EPI sequence with an acquisition matrix at the 64×64 in-plane spatial resolution and 32 slices. A high-resolution ($1 \times 1 \times 1.2$ mm^3) anatomical image was also acquired for each subject using a 3-D gradient-echo inversion-recovery sequence.
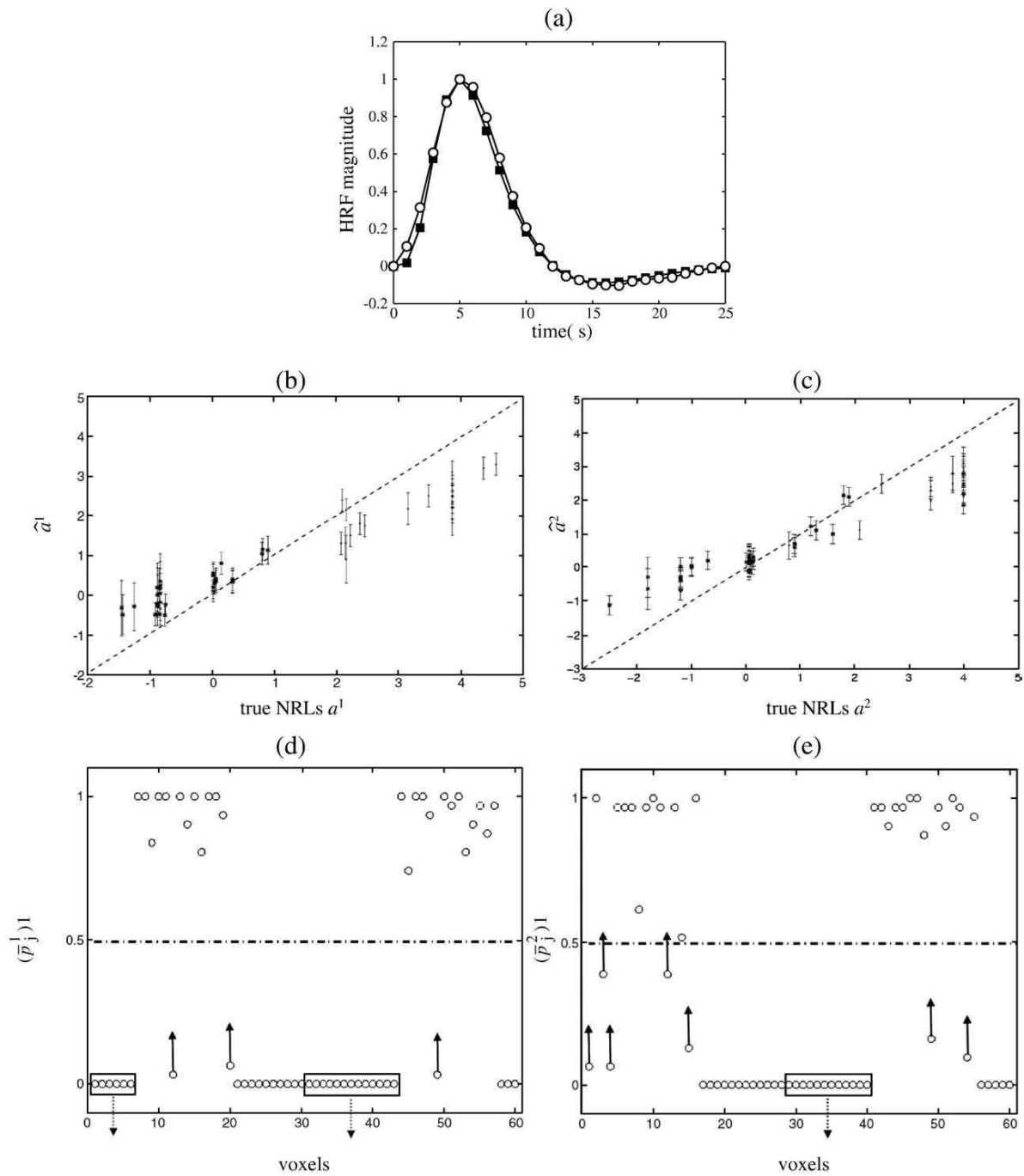
Fig. 7. Simulation results using model $\mathcal{M}_2$. FN voxels are indicated by upward arrows. The truly deactivating voxels that have been mixed in class 0 are surrounded by rectangle.

Experimental paradigm and contrast selection

The reader may refer to (Dehaene-Lambertz et al., 2006) for details about this fMRI experiment. In short, the motivation of this study was to measure the reduction in the neural activity subserving a cognitive representation when this representation is accessed twice (the so-called "repetition suppression" effect), resulting in a detectable adaptation of the measurable signal in fMRI (Grill-Spector and Malach, 2001; Naccache and Dehaene, 2001). The experiment consisted of a single session of $N=216$ scans lasting TR=2.4 seconds

each. Sixty sentences presented in a slow event-related design (SOA=14.4 s) were recorded. Each sentence (S_1) could be repeated two (S_2), three (S_3) or four (S_4) times in a row. The main goal of our subsequent analysis was twofold. First, our primary interest was to exhibit regions which activation to a given sentence either decrease with repetition or keep a constant magnitude across the repetitions. Second, we were interested in inferring the hierarchical temporal organisation from the parcel-based HRF estimates along the superior temporal sulcus (STS).
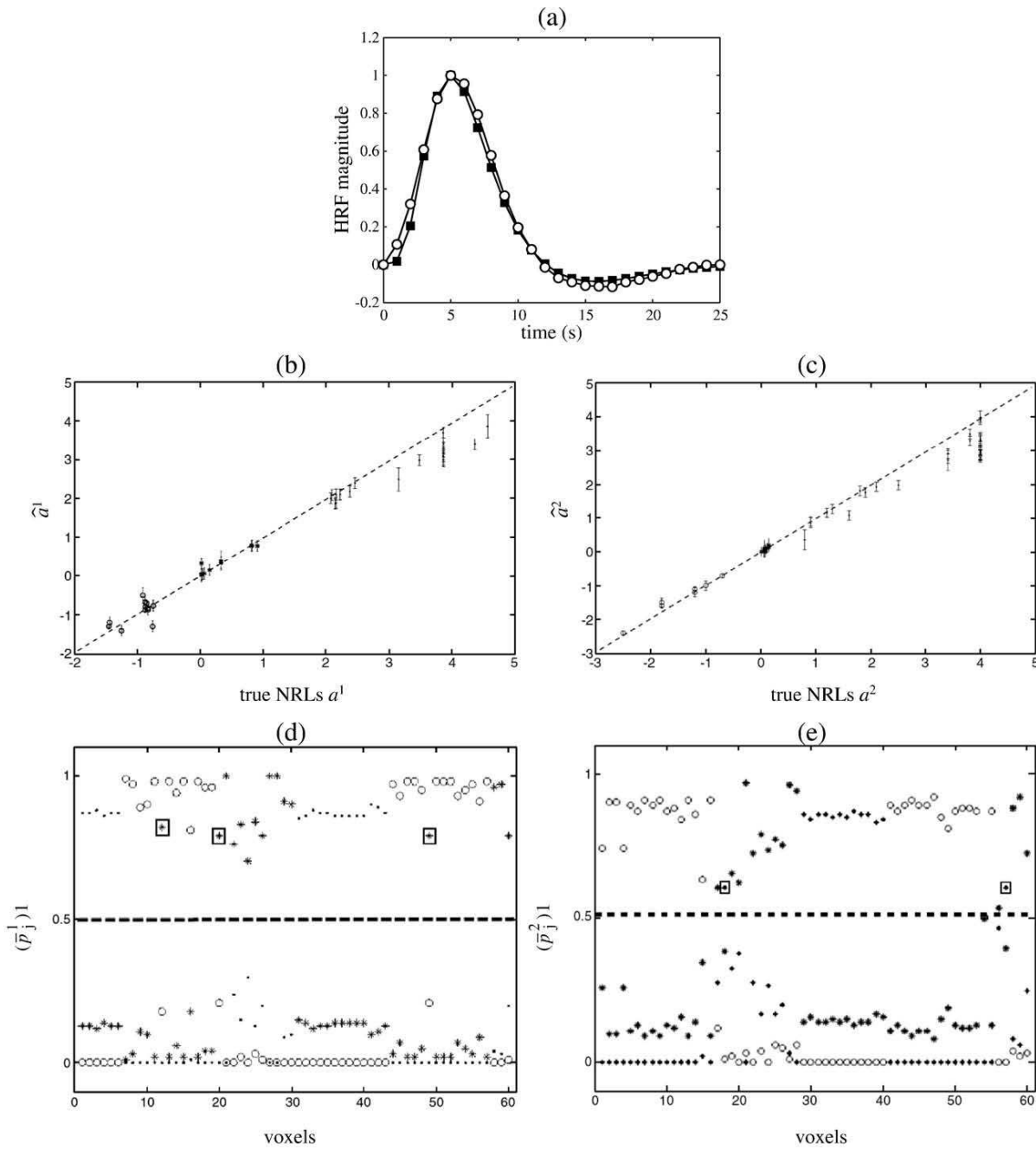
Fig. 8. Simulation results using model $\mathcal{M}_5$. The same legend as in Fig. 2 holds. FP and FN voxels are surrounded by rectangles (3 FNs in (d) and 2 FPs in (e)). $\bigcirc$, · and * symbols represent $(\bar{p}_j^m)_1$, $(\bar{p}_j^m)_{-1}$ and $(\bar{p}_j^m)_0$, respectively.

Table 3
Values of the integrated log-likelihood $\log p(y \mid \mathcal{M}_m)$ computed from a stabilized version of the harmonic mean identity (Raftery et al., 2007) for models $\mathcal{M}_m$ with $m \in \mathbb{N}_5^*$

Model m	$\log p(y \mid \mathcal{M}_m)$	Fig. #	$\log BF_{mn}$, $n = 1:4$			
$\mathcal{M}_1$	-199	Fig. 1	NR	18	316	400
$\mathcal{M}_2$	-217	Fig. 2	-18	NR	298	378
$\mathcal{M}_3$	-515	Fig. 3	-316	-298	NR	378
$\mathcal{M}_4$	-595	Fig. 4	-400	-378	-80	NR
$\mathcal{M}_2$	-700	Fig. 6	Log $BF_{52} = 356$			
$\mathcal{M}_5$	-344	Fig. 7				

Model comparison based on the computation of Bayes factors $\log BF_{mn} = \log p(y \mid \mathcal{M}_m) - \log p(y \mid \mathcal{M}_n)$ for every pair (m, n). **NR** stands for **Not Relevant**.

Since the most significant habituation effect occurs between the first and second sentence repetitions, we modelled the four conditions S_1-S_4 but we only studied the contrast $S_1 - S_2$.

Pre-processings

As explained in Subsection 2.1.1, the grey matter's mask was first computed (see Fig. 9(a)) and then dilated using a 4 mm-radius sphere to account for the width of the cortical ribbon. Fig. 9(b) shows the result of this step. The resulting mask $\mathcal{M}_a$ contains 19719 voxels at the fMRI resolution.

We checked that for nine out of ten subjects the raw fMRI data were motion-free approximately. All T_1-weighted MRI images were normalised onto the MNI template and functional images were transformed accordingly. fMRI volumes were also spatially smoothed using a Gaussian kernel with $FWHM = 6$ mm along each direction. A first level analysis was performed for each subject using SPM2. The GLM modelled the four presentations of a given sentence with two regressors (the canonical HRF and its time derivative). Then, the parcellation was computed from the parameter estimates of this analysis. We chose a relevant F-contrast $c = [1, 0, -1, 0, \cdots ; 0, 1, 0, -1, \cdots, \ldots]$ to study the habituation effect etween the first and second presentations of a given sentence $(S_1 - S_2)$. Fig. 9(c) depicts an axial view of this parcellation for the same slice ($z = -4$ mm).

Our approach strongly relying on a functional homogeneity assumption, we started by comparing the results using increasing parcel numbers (from $K = 100$ to $K = 500$ parcels in $\mathcal{M}_s = \mathcal{M}_a \cup \mathcal{M}_f$.

Fig. 10. (a)-(b) NRL estimates in one slice of subject 1's brain (at $z = -4$ mm) in response to S_1 (a) and S_2 (b). Values correspond to the NRL coefficients only for voxels belonging to $\mathcal{M}_s$. Otherwise they are equal to 0. (c)-(d): detection results in the same slice for S_1 and S_2, respectively. Voxels colour-coded in red are detected as activating. In black are the parcel borders that are superimposed to the different map results to show the parcellisation influence on the estimation of such parameters.

We checked that $K = 200$ is large enough to guarantee a higher and sufficient degree of homogeneity. Here, the smallest and the largest parcels contained 44 and 190 voxels, respectively. Within each parcel, the degree of functional homogeneity was measured by computing a correlation matrix over the parameter estimates of the GLM. Note that this could also be done over the fMRI signals attached to each parcel.

Results

Our method was tested on the nine datasets. Here, we only report results for Subject 1. Although the habituation effect and
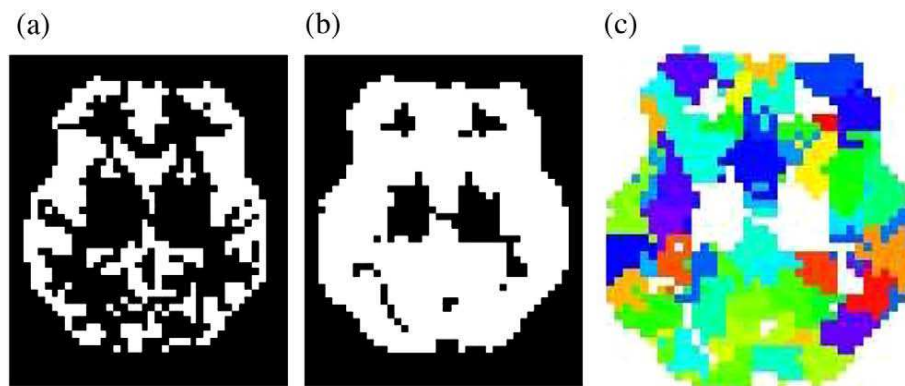

Fig. 9. (a): Slice of Subject 1's anatomical mask ($z = -4$ mm). (b): its dilated version $\mathcal{M}_a$ to match the functional resolution. (c): corresponding parcellation in the same slice. Each colour codes for a different parcel.
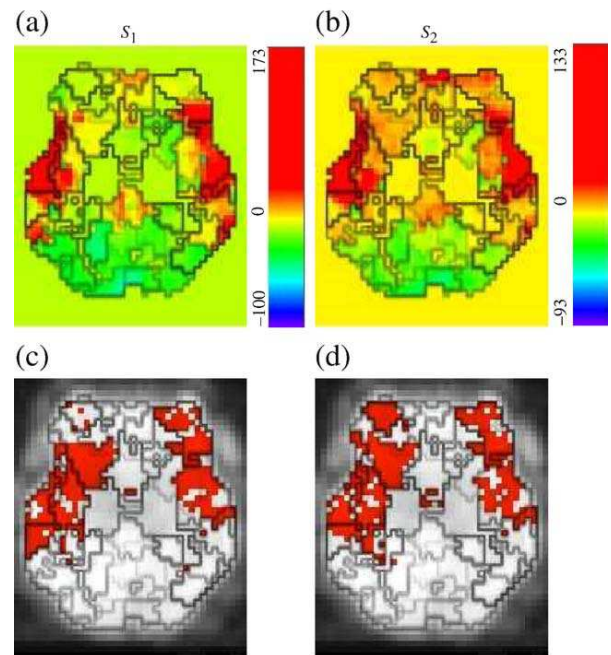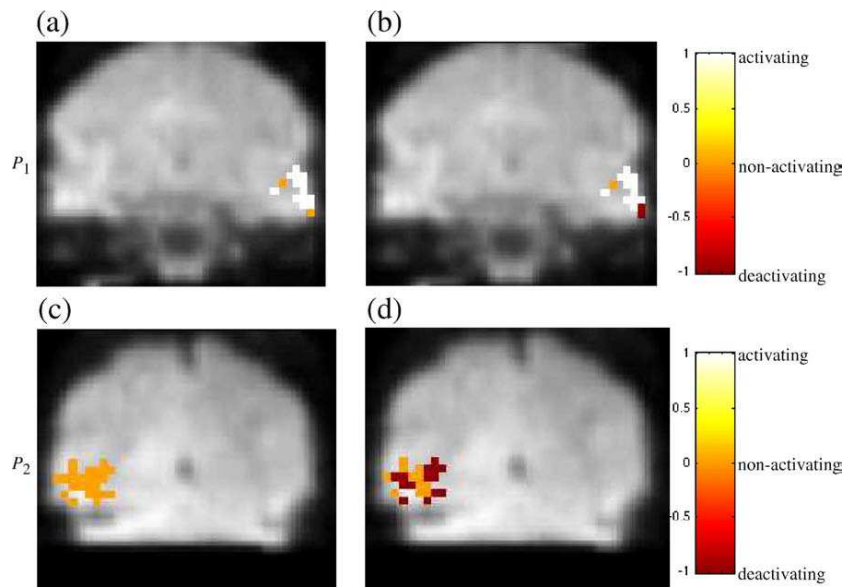
Fig. 11. Comparison of S_1-based classification maps between the *GaGMM* prior (left column, (a)-(c)) and its 3-class extension (right column, (b)-(d)) in two different parcels, $\mathcal{P}_1$ (top row, at $y=-16$ mm) and $\mathcal{P}_2$ (bottom row, at $y=-52$ mm). Square spots in white, orange and brown match with activating, non-activating and deactivating voxels, respectively.

brain dynamics (*i.e.,* the HRF shape) are subject to inter-individual variability in terms of spatial localisation and activation delay, the conclusions drawn for Subject 1 remain quite valid for the others.

In what follows, the proposed joint detection-estimation algorithm was applied to each parcel of Subject 1's brain. Figs. 10(a)-(b) shows the maps of the NRL estimates corresponding to conditions S_1 and S_2, respectively, in a given slice of the brain. In the same slice, Figs. 10(c)-(d) shows the activation probability maps attached to S_1 and S_2 (see (9) for details) our algorithm provides. Activating voxels appear in red colour.

Probing for putative deactivation

This first analysis was devoted to looking at putative deactivations, that is the presence of negative NRLs. We actually

performed tests on all parcels to assess differences between the *GaGMM* prior and its 3-class extension. For illustrative purpose, Fig. 11 depicts the results of such a comparison on two parcels $\mathcal{P}_1$ and $\mathcal{P}_2$, composed of 129 and 135 voxels, respectively. Interestingly, the vast majority of voxels in $\mathcal{P}_1$ elicit a coherent activation in response to the first presentation of a sentence (S_1), while in $\mathcal{P}_2$, most voxels are non-activating. As shown in Fig. 11(a)-(b), the same S_1-based classification map is obtained in $\mathcal{P}_1$ irrespective of the mixture model. The same conclusion holds with respect to S_2 (results not shown). In $\mathcal{P}_2$, Figs. 11(c)-(d) illustrates that a few voxels move from the non-activating state to the deactivating one.

However, the corresponding NRL estimates are of small magnitudes indicating that this new classification may arise by chance. Bayesian model comparison statistically confirms our result since numerical evaluation of Bayes factors gives us $\log BF_{52}=-1.2$ for
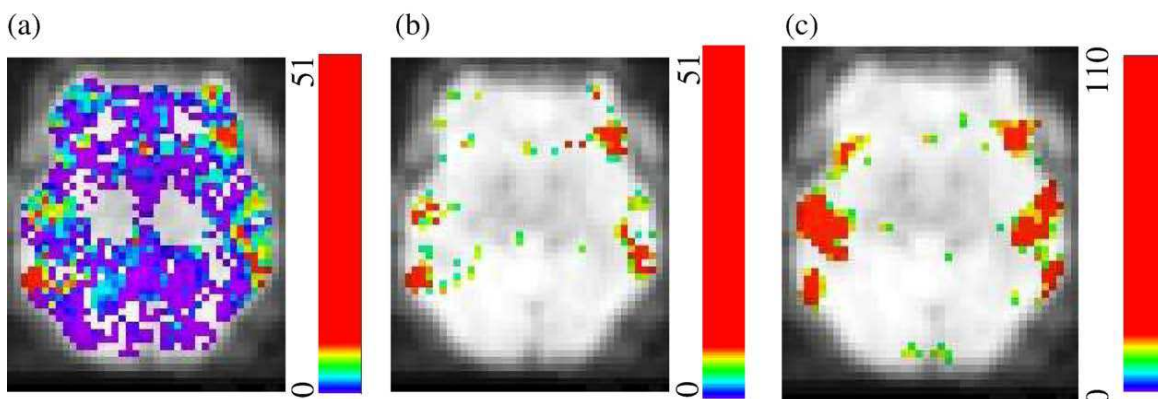

Fig. 12. Statistical maps are superimposed to a functional image and results are given in one slice of the brain (z Statistical maps are superimposed to a functional image and results are given in one slice of the brain ($z=-4$ mm). (a): KL-distance for voxels in $\mathcal{M}_s$ at $z=-4$ mm. (b): The MS most significant voxel KL-distance values. (c): The MS most significant voxel F-values.
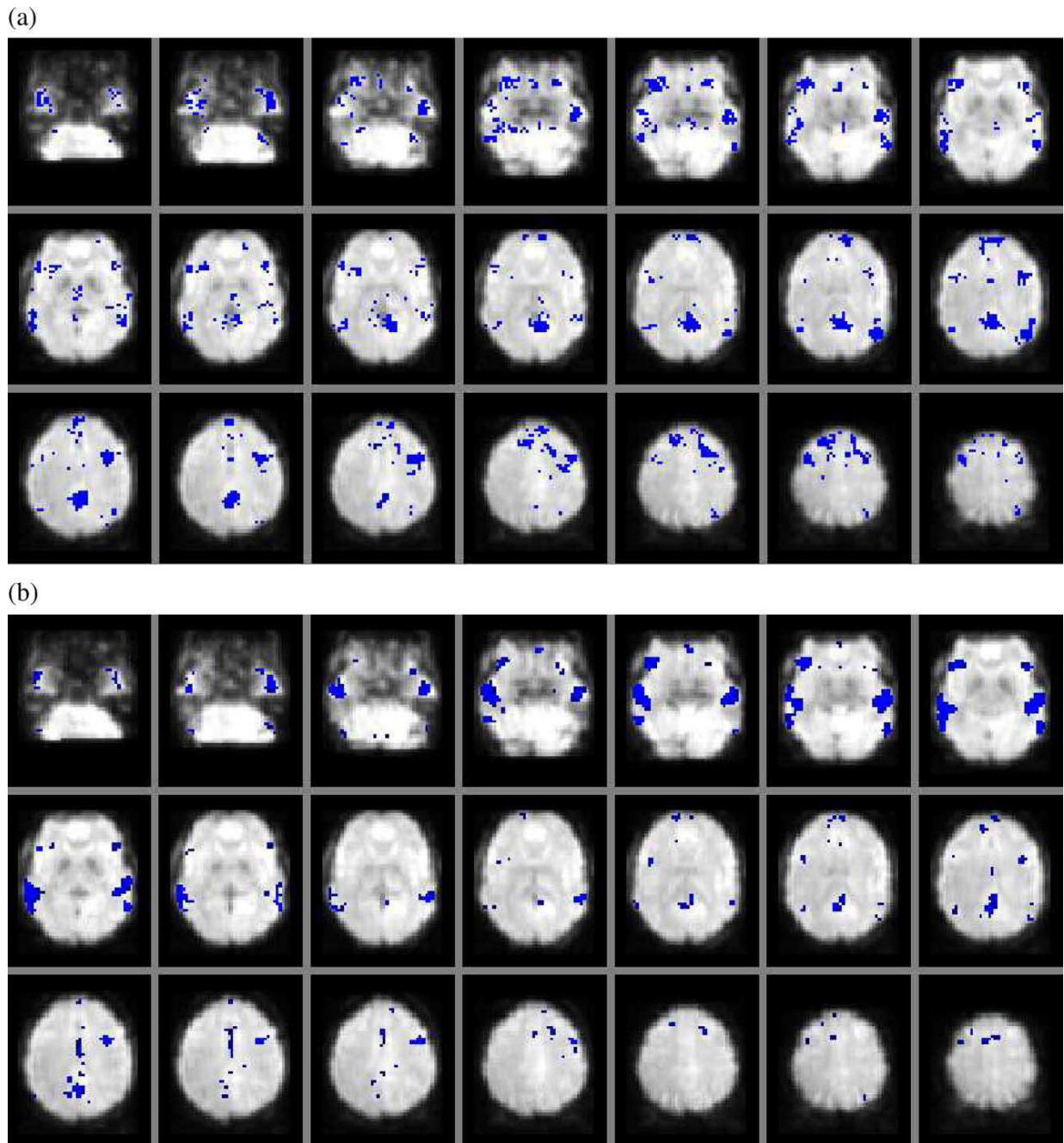
(a)

(b)

Fig. 13. From left to right and from top to bottom, brain slices organised along increasing axial axis (the bottom of the brain appear first in the superior left corner). Study of the contrast S_1. Study of the contrast $S_1 > S_2$. (a): thresholded PPM (see (11)) at $\eta = 0.999$; (b): thresholded SPM t-map at T = 3.09 (corrected for multiple comparisons) obtained using SPM2.

$\mathcal{P}_1$ and log $BF_{52} = -.5$ for $\mathcal{P}_2$. These results show that there is less evidence in this dataset for supporting model $\mathcal{M}_5$. Therefore, the introduction of the third class in the mixture is not necessary to analyse these data, particularly in the brain regions involved in the treatment of phonological stimuli (language comprehension). In the rest of the paper, we restrict ourselves to the *GaGMM* prior.

Fig. 14. (a): Single-slice comparison ($z = 0$ mm, slice 7 in Fig. 13) for the contrast $S_1 > S_2$. **left panels**: thresholded PPM at $\eta = 0.999$; **right panels**: thresholded SPM t-map at T = 3.09 (corrected for multiple comparisons) obtained from the SPM analysis. (b): same comparison in slice $z = 24$ mm (slice 13 in Fig. 13). HRF estimates (in blue) in these regions compared to the canonical hemodynamic response function used in SPM (in red).

(a)
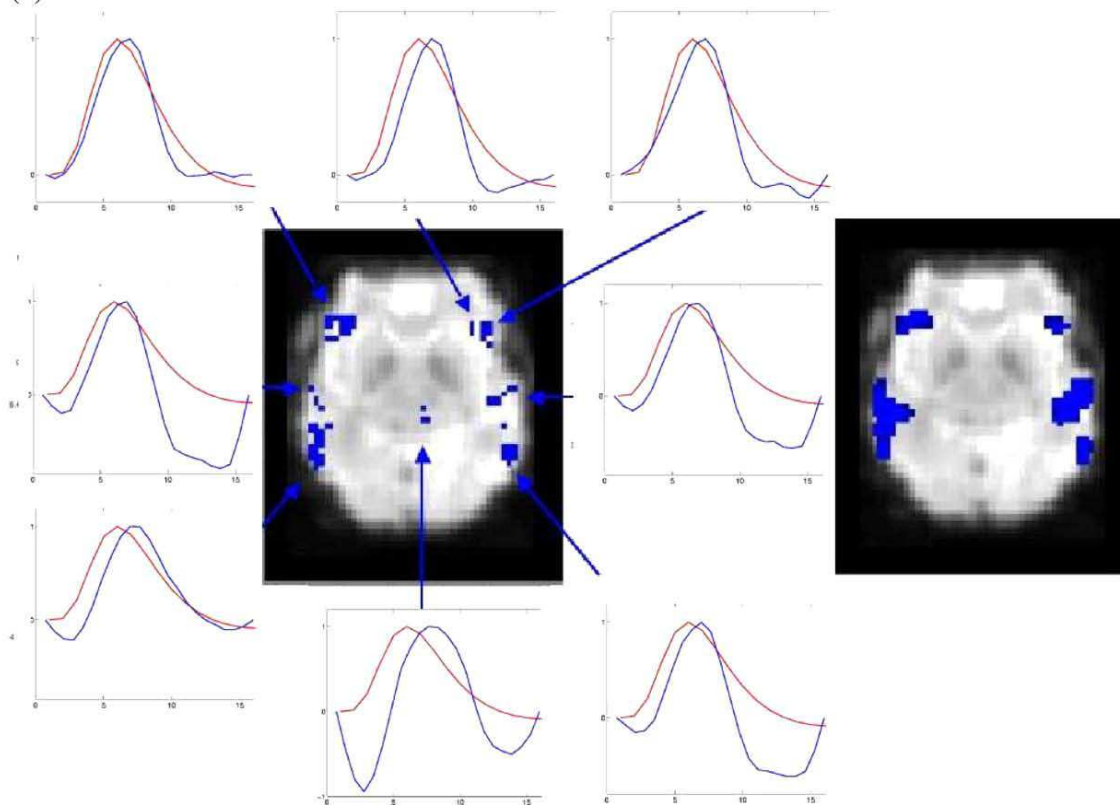
(b)
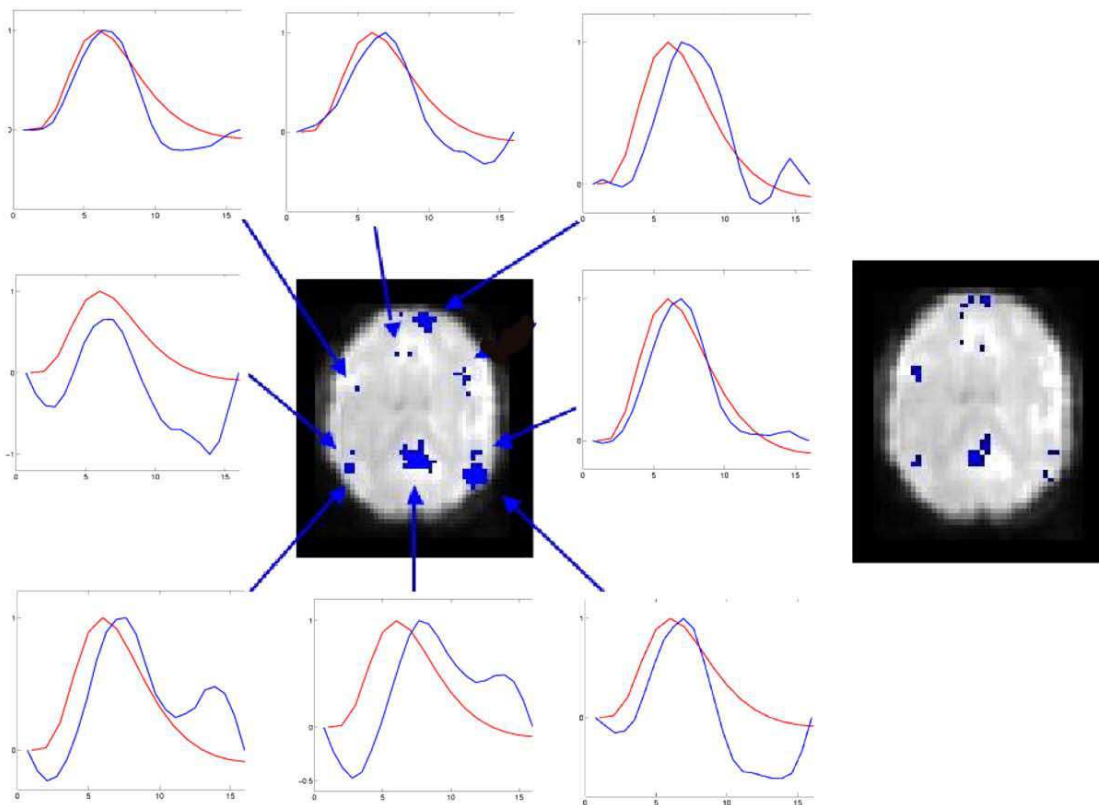
Unsigned comparison between conditions

The result of our KL-distance map is shown in Fig. 12(a) in one slice of the brain ($z=-4$ mm). In Figs. 12(b)-(c), the KL-distance map is then compared to the standard F-test map by extracting the most significant, say MS, voxels from both volumes, according to the corresponding criteria (the largest KL divergence and the highest F value). This thresholding procedure shows common features in the activation patterns but also discrepancies in the temporal lobes that elicit different responses. On this slice, the KL-based criterion provides less activations. This may be due either to our approximation of the KL divergence or to the criterion itself.

Signed comparison between conditions

The mapping of the habituation phenomenon calls for *signed* comparisons since we are looking for voxels where a significant decrease of NRLs between S_1 and S_2 can be observed. Fig. 13 shows such a comparison over the whole brain between the PPM derived using the proposed methodology and the thresholded T-map obtained using SPM2. The activated regions (in blue) for the contrast $S_1 > S_2$ elicit therefore a higher response when a sentence is presented only once.

Statistical differences appear between the proposed PPM and the corresponding SPM. On the one hand, in the majority of slices we observe more activations on the PPM shown in Fig. 13(a). To a certain extent, these differences can be explained by the shape variations of the HRF estimate and its deviation from the canonical shape prescribed in SPM. Moreover, the correction for multiple comparisons used for the *t*-map may dramatically reduce the number of activating voxels. In contrast, no correction has been applied over the PPM.

On the other hand, in the temporal lobes (top row, right slices of Fig. 13(b)), sensitivity of detection seems better on the Student-*t* map: the activated clusters appear larger. Note also the presence of isolated activating voxels in Fig. 13(a). It is likely that this reflects the presence of false positives. To circumvent this issue and decrease the FP rate in the PPM, a spatial correlation model between neighbouring voxels can be introduced; see (Vincent et al., 2007b,a) and the discussion.

In Fig. 14, we represent the estimated HRFs in parcels corresponding to areas where we notice sensitivity differences. It clearly appears that in some parcels our HRF estimates exhibit unexpected timing properties. For instance, in central regions ($x=0$,
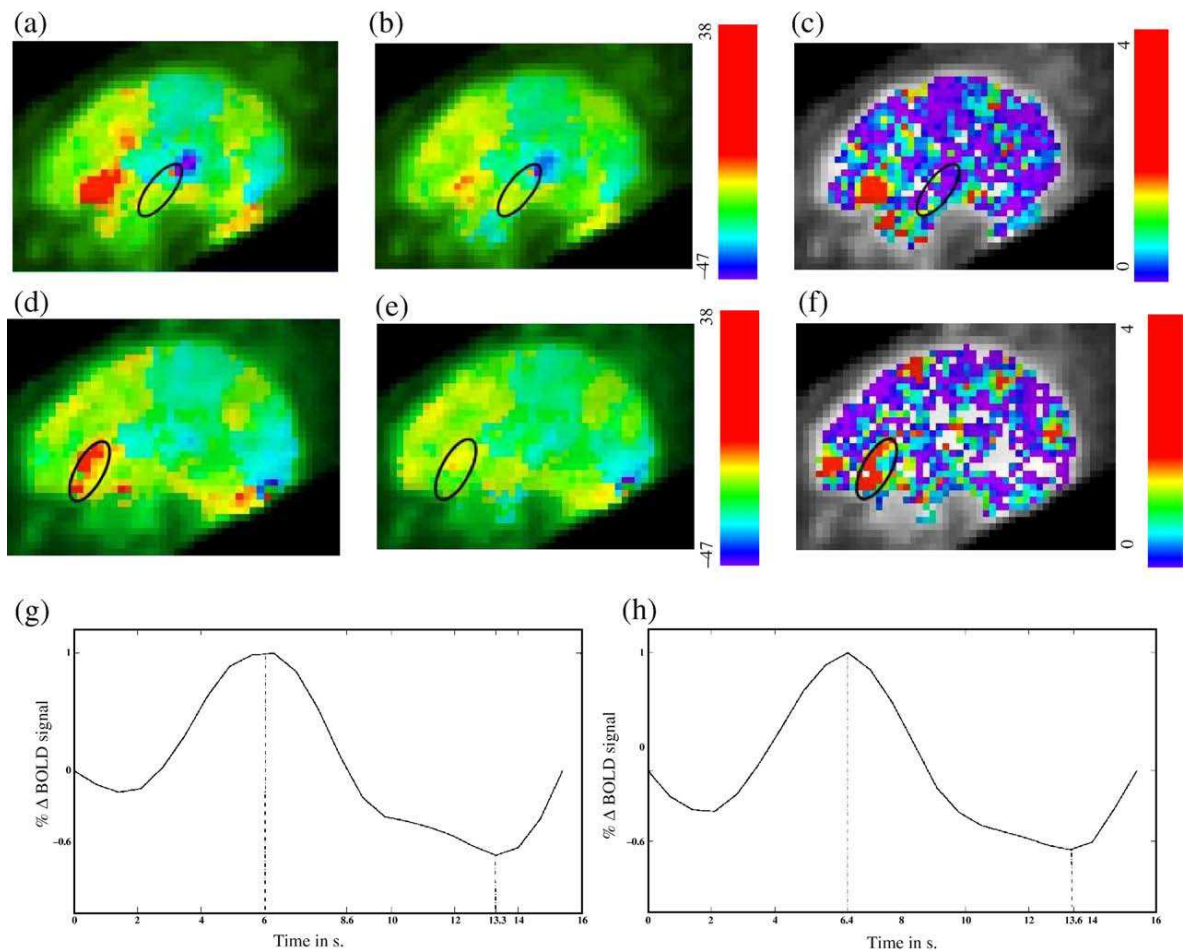

Fig. 15. (a)-(b) and (d)-(e): NRL estimates in the sagittal slices located (a)-(b) and (d)-(e): NRL estimates in the sagittal slices located at $x=-48$ mm and $x=-40$ mm, respectively. (a)-(b) provide the magnitudes in response to S_1 while (d)-(e) give us the NRLs in response to S_2. (c)-(f): KL-distance maps between the corresponding NRLs. Images are superimposed to functional data. The parcels containing the Heschl's gyrus (top row) and Broca's area (bottom row) are surrounded in black and referenced as $\mathcal{P}_{He}$ and $\mathcal{P}_{Br}$ in the text. (g)-(h): HRF estimates in $\mathcal{P}_{He}$ and $\mathcal{P}_{Br}$, respectively.

$y < 0$), we obtained initial dips that were difficult to predict in advance. This requires further analysis (see Discussion). In regions where the HRF estimate is very close to the canonical shape, the PPM and SPM provide similar activation patterns. Finally, some regions were also detected as activating ($S_1 > S_2$) by both methods while there is no evidence in the literature to suppose a priori that they elicit responses to auditory stimuli (see for instance Fig. 14(b) along the interhemispheric axis).

Habituation and temporal organisation

We focus on superior temporal regions ranging from the primary auditory cortex (Heschl's gyrus) to associative areas (middle and posterior STS). Although it has been shown in (Dehaene-Lambertz et al., 2006) that repetitions affect both amplitude and delay of responses, we only model habituation effect on the NRLs by considering the different sentence presentations as different conditions. This procedure is not optimal but remains quite simple.

The first interesting region is Heschl's gyrus located in the primary auditory cortex around voxel with coordinates $(-48, -12, 0)$ mm in the standard Talairach space. This area shows the same response magnitude each time a sentence was presented. Figs. 15(a)-(b) shows the NRL estimates in parcel $\mathcal{P}_{He}$, which is circled in black for S_1 and S_2, respectively. In every voxel of $\mathcal{P}_{He}$, these magnitude parameters are very close to each other making the KL-distance between the marginal posterior distributions of S_1 and S_2 close to zero (see Fig. 15(c)). Hence, the measured difference between S_1 and S_2 is not statistically significant. Summary statistics computed over the NRL estimates in $\mathcal{P}_{He}$ are reported in Table 4 (left col.) and confirm our first analysis quantitatively. The same study was done in $\mathcal{P}_{Br}$ containing Broca's area and centered around voxel $(-40, 24, 0)$ mm in the Talairach space. Figs. 15(d)-(e) clearly indicates a strong decrease in the NRLs between S_1 and S_2. The higher value of the KL divergence reported in Fig. 15(f) confirms the presence of a significant habituation effect in $\mathcal{P}_{Br}$. Our quantitative analysis (see Table 4, right col.) also shows the same trend as outlined by the strong discrepancies between the mean and standard deviations of S_1 and S_2.

In (Dehaene-Lambertz et al., 2006), a sine-wave GLM was designed and fitted to study the speed of habituation. It allowed one to exhibit a temporal organisation of the temporal lobe with fastest responses located in Heschl's gyrus and slowest ones in temporal poles. Here, given the proposed methodology, the temporal organisation is studied more directly by measuring and sorting the timing properties of the parcels $\mathcal{P}_{He}$ and $\mathcal{P}_{Br}$.

In these parcels, we also investigated the temporal aspects of the STS organisation by measuring different features (time-to-peak T_{peak}, time-to-undershoot $T_{undershoot}$) on the estimated HRFs (see Figs. 15 (g)-(h)). We computed these quantities over each parcel ($\mathcal{P}_{He}$ and $\mathcal{P}_{Br}$ and their four closest neighbours) separately before averaging them to get a mean estimate. We found that the responses in $\mathcal{P}_{He}$ ($T_{peak} = 6$ s. and $T_{undershoot} = 13.3$ s.) occur and return to the baseline earlier than

the responses in $\mathcal{P}_{Br}$ ($T_{peak} = 6.4$ s. and $T_{undershoot} = 13.6$ s.). After averaging these quantities over the five parcels, we obtained congruent results ($T_{peak} = 6.1$ s. and $T_{undershoot} = 13.2$ s. around $\mathcal{P}_{He}$ vs. $T_{peak} = 6.5$ s. and $T_{undershoot} = 13.5$ s. around $\mathcal{P}_{Br}$) meaning that the region embedding Hesch's gyrus elicits brain activations faster than the region including Broca's area. This confirms more directly what has been already derived in (Dehaene-Lambertz et al., 2006) although no statistical test is provided to assess the significance of this result. The next question concerns of course the putative reasons of these earlier responses in Hesch's gyrus. There is a large evidence in these datasets for supporting faster neurodynamics as the main origin of these results instead of faster hemodynamics. This can be checked for instance using complementary analysis like a phase analysis conducted in (Dehaene-Lambertz et al., 2006).

Discussion

In this paper, we have proposed an original method to perform a parcel-based joint detection-estimation of brain activity from fMRI data. It has been shown on simulated datasets that a gamma-Gaussian mixture as prior pdf on the NRLs outperforms a Gaussian mixture in terms of sensitivity/specificity trade-off. It has also been reported that the noise model has an influence over this compromise, particularly at lower SNR: a first-order AR model provides lower false positive and negative rates in comparison with a white noise model.

Our method extends previous works (Makni et al., 2005, 2006b,a) to deal with anatomically informed whole brain analysis. As already done in (Smith et al., 2003; Nieto-Castanon et al., 2003; Flandin et al., 2002), analysis was constrained to the mask of the grey matter obtained from a segmentation of the T1-weighted MRI. Our approach also relies on functional homogeneity assumptions at a regional scale that can be assessed either from the fMRI time series themselves or from the GLM parameter estimates. To meet these conditions, we resort to an automatic parcellation technique developed in (Thirion et al., 2006) but alternative clustering strategies may be thought of. Our approach therefore depends on this prior decomposition making the global within-subject analysis a two-steps procedure. The quality of the parcels will have an impact on the model fitting and a slight modification of the parcellation may generate different results especially in case of identifiability problems. By varying the number of parcels, we have checked that our results remain quite stable for different parcellations. The solution to this problem actually lies in the coupling of the parcellation procedure with our detection-estimation approach. This a very appealing direction of research but remains beyond the scope of the present work. At the expense of an increased computational complexity, the two steps could be merged in a combined approach through a hierarchical Bayesian model: one might be interested in improving the parcellation from the results of the detection-estimation stage using an iterative strategy: neighbouring parcels would be grouped if their underlying hemodynamics share similar features. The algorithm should take place in the context of reversible jumps MCMC to properly handle fusion/segregation moves between parcels Green (1995); Richardson and Green (1997). Besides, the parcellation identification issue could also be attacked using triplet Markov fields (Benboudjema and Pieczynski, 2007), which seem suitable for modelling nonstationarities in image segmentation.

A strong feature of our approach is the possibility to derive parcel-based HRF time courses throughout the brain. It allows us to assess the spatial variability of the HRF shape and to check that this shape greatly fluctuates across parcels. Since the parcellation

Table 4
Summary statistics in parcels containing Heschl's gyrus and Broca's area

Statistics	$\mathcal{P}_{He}$, Heschl's gyrus		$\mathcal{P}_{Br}$, Broca's area	
	S_1 ($m=1$)	S_2 ($m=2$)	S_1 ($m=1$)	S_2 ($m=2$)
$\max_j (\hat{a}_j^m)^{PM}$	39.45	31.29	46.57	22.33
$\min_j (\hat{a}_j^m)^{PM}$	0.35	-0.88	-4.21	-0.32
Mean $(\hat{a}_j^m)^{PM}$	7.96	3.76	19.81	9.91
Median $(\hat{a}_j^m)^{PM}$	5.09	1.44	19.95	10.49
Std $(\hat{a}_j^m)^{PM}$	7.62	6.72	8.95	4.92

procedure is derived at the group level, one is able to compare subject specific HRFs in a given parcel. Doing so, we have noticed that the between-subject variability in the HRF shape seems to be larger than the within-subject spatial variability, as already suggested in (Handwerker et al., 2004).

Our results also suggest that the modelling of spatial HRF fluctuations is important to segregate brain regions involved in the experimental paradigm (Heschl's gyrus, Broca's area, …). The adaptation effect was particularly evident between the first and second sentences. The pattern of adaptation was different across regions with a set of regions demonstrating the same response each time a sentence was presented (i.e., Heschl's gyrus) and regions showing a more or less strong decrease between the first and second presentation (eg, Broca's area, superior STS).

Nonetheless, there exist fMRI experiments for which the proposed approach may fail because of the inhomogeneity of the HRF shape both in space and across conditions at a regional scale. This may occur when neurodynamic and hemodynamic fluctuations intermix. Indeed, recent studies of the fine structure of the fusiform face area (FFA) (Grill-Spector et al., 2006) have shown that the FFA is actually highly heterogeneous. It appears that the FFA is composed in reality of several small scale subregions that respond strongly, not only to faces, but also to cars and sculptures. The subregions discovered in (Grill-Spector et al., 2006) were associated with very distinct HRFs. For such studies, the best way to capture these HRF fluctuations is to perform a voxel-based non-parametric analysis on the basis of which statistical comparisons can be done across conditions, as demonstrated in (Ciuciu et al., 2003; Marrelec et al., 2004).

We have also distinguished some differences between our PPMs and the SPMs derived from a classical GLM-based analysis. Our results confirm the interest of a simultaneous procedure for detecting and estimating brain activity. The proposed procedure actually improves the sensitivity of detection in some regions where the temporal characteristics (time-to-peak, time-to-under-shoot, …) of our HRF estimate deviate from those of the canonical shape. Unfortunately, this effect is not systematic on the datasets since in other regions a loss of sensitivity was observed. The reasons underlying this unexpected decrease have to be identified. To elucidate this issue, future work will be devoted to the comparison of a degraded version of our joint detection-estimation procedure with the actual one. The degraded version corresponds to an inference scheme where the HRF is maintained fixed and is not sampled at every iteration of the Gibbs sampler. We will quantify statistically this sensitivity difference on simulated data where we know exactly the "ground" truth. For doing so, we could generate fMRI data over the whole brain using the fMRI simulator developed by (Drobnjak et al., 2006).

In the present paper, we get rid of the scale ambiguity problem due to the bilinearity of (1) w.r.t. the pair $(\boldsymbol{h}, \mathtt{a})$ by imposing a unitary norm constraint over $\boldsymbol{h}$. This could have a dramatic impact on the convergence of our sampling scheme and then on the recovered HRF shape by distorting the target distribution of the MCMC scheme. Alternative strategies may first consist in cancelling this normalisation step. However, the resulting sampling scheme is too slow to converge in a reasonable amount of time (Veit and Idier, 2007). An efficient alternative has been proposed in (Veit and Idier, 2007). It has been applied to the joint detection-estimation of brain activity in (Ciuciu et al., 2007). It consists in adding to the MCMC procedure a sampling step of a positive scalar parameter s coding for the HRF scale. It can be shown that its sampling is fast, follows a *generalized*

inverse Gaussian distribution in case of Gaussian mixtures and guarantees the theoretical convergence of the generated Markov chain to the posterior distribution. Deriving the target distribution of this scale parameter for inhomogeneous mixtures is beyond the scope of this paper.

To conclude about the real impact of our within-subject analysis, inference should take place at the group level. In other words, we should compare the results of two random effect analyses (RFX) based on the same group statistics (*e.g.,* mean effect) and statistical test (Student *t*-test). The first RFX analysis would correspond to the gold standard, in which the input data are given by the normalised effects of a standard individual SPM analysis. The second analysis would take the results of our algorithm for each subject as inputs. Of course, this is only feasible in case of multi-subjects parcellation, what is currently obtained using the procedure described in (Thirion et al., 2006, 2007).

From a methodological point of view, we have shown that our joint detection-estimation technique is able to identify deactivations in the brain. This is owing to the introduction of a third class in the prior mixture model associated to the NRLs. Nonetheless, we did not exhibit real deactivations on the analysed datasets. In the future, we should therefore validate the 3-class extension on specific datasets. A good candidate could be a dataset acquired during an event-related auditory paradigm in which silence events are presented randomly to compare activations to a baseline derived from such events. As already shown in (Ciuciu et al., 2003), silence events may generate deactivations in the temporal lobe if they are presented to the subject when the gradients of the scanner are switched off. This will be the subject of further work.

Smoothing the data spatially provides a reliable manner for recovering clusters of activation instead of isolated spots, at the expense of a loss of resolution. To avoid this preprocessing, the proposed method could be extended by introducing spatial correlation in the prior model. This could be done either on the NRLs (a) or on the underlying states (labels q). We argue in favour of the second solution for simplicity reasons. As already derived for Gaussian mixtures in (Vincent et al., 2007b,a), it is quite simple to sample from an Ising (2-class model) or Potts (3-class model) Markov random field (MRF) that enforce neighbouring voxels to be classified in the same state (*e.g.,* activating). This approach actually seems more reasonable in terms of computational load than considering edge-preserving MRF based on non-quadratic potentials (Green, 1990; Geman and McClure, 1987). Also, for computational reasons this extension has been developed in a supervised framework meaning that the hyper-parameter encoding spatial regularity of the hidden MRF is set by hand. Future work will be focused on an spatially adaptive extension in which this parameter is estimated as well, as already done in (Woolrich et al., 2005; Woolrich and Behrens, 2006).

Another extension that could be introduced at little expense concerns the analysis on the cortical surface, as proposed in (Andrade et al., 2001). This will probably improve the sensitivity of detection by constraining the analysis to the cortical surface. Such study needs first a segmentation of the anatomical MRI, then an extraction of the grey-white matter interface (*e.g.,* as a mesh), and finally requires an interpolation of the fMRI signal on the nodes of the mesh (see for instance (Grova et al., 2006) for a suitable approach).

Finally, the model presented here assumes that the NRLs are constant in time. Hence, to account for putative habituation effects, it requires to model repetitions of the same stimulus as different

experimental conditions, what may be not optimal in terms of sensitivity of detection. To account for trial-varying NRLs due to adaptation or learning effects arising either as a direct consequence of the paradigm or as a alteration of subject's arousal, the proposed model can be generalised in a way that makes the number of unknown parameters not too large. As proposed in (Ciuciu et al., 2006), habituation can be modelled at the voxel level by a pair of parameters: the NRL to the first trial of the stimulus and a mean habituation speed across the consecutive trials that follows a hyperbolic parametric model depending on the inter-stimulus intervals.

Hopefully, all these additional points will induce improvements in the detection-estimation results and will help to a better comprehension of brain functions.

Acknowledgments

Appendix A. Densities

We give the definitions of the densities used throughout this paper. We also provide numerical recipes for efficient simulations according to complex distributions when necessary.

A.1. Multivariate normal density

The multivariate Normal density for d-dimensional variable $\boldsymbol{x}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is given by

$$\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = |2\pi\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right). \quad \text{(A.1)}$$

The general formula of Kullback-Leibler (KL) divergence between a test density, say $q(\boldsymbol{x})$ and a reference density $p(\boldsymbol{x})$ is

$$KL(q||p) = \int q(\boldsymbol{x}) \log \frac{q(\boldsymbol{x})}{p(\boldsymbol{x})} d\boldsymbol{x}. \quad \text{(A.2)}$$

For multivariate normal densities $q(\boldsymbol{x}) = \mathcal{N}\left(\boldsymbol{x}|\boldsymbol{\mu}_q, \boldsymbol{\Sigma}_q\right)$ and $p(\boldsymbol{x}) = N\left(\boldsymbol{x}|\boldsymbol{\mu}_q, \boldsymbol{\Sigma}_p\right)$, Eq. (A.2) takes the following form:

$$KL(q||p) = \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_p|}{|\boldsymbol{\Sigma}_q|} + \frac{1}{2}tr\left(\boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_q\right)$$
$$+ \frac{1}{2}\left(\boldsymbol{\mu}_q - \boldsymbol{\mu}_p\right)^t\boldsymbol{\Sigma}_p^{-1}\left(\boldsymbol{\mu}_q - \boldsymbol{\mu}_p\right) - \frac{d}{2}.$$

For univariate normal densities $q(x) = \mathcal{N}\left(x|\mu_q, \sigma_q\right)$ and $p(x) = \mathcal{N}\left(x|\mu_p, \sigma_p^2\right)$, the KL (A.2) distance becomes:

$$KL(q||p) == \frac{\left(\sigma_p^2 - \sigma_q^2\right)^2 + (\mu_p - \mu_q)^2\left(\sigma_p^2 + \sigma_q^2\right)}{4\sigma_p^2\sigma_p^2}. \quad \text{(A.3)}$$

A.2. Positive normal density

The truncated normal distribution for a scalar variable $x \in \mathbb{R}_+$ with *parameters* (m, υ) expresses as follows:

$$\begin{cases} \mathcal{N}^+(x|m,\upsilon) = C^{-1}\exp\left(-(x-m)^2/2\upsilon\right)\mathbb{I}_{\mathbb{R}_+}(x), \\ C = \sqrt{\pi\upsilon}\left[1 + \text{erf}\left(m/\sqrt{2\upsilon}\right)\right]/\sqrt{2}, \end{cases} \quad \text{(A.4)}$$

where erf is the error function (Abramowitz and Stegun, 1970, p. 297): $\text{erf}(z) = 2/\sqrt{\pi}\int_0^z e^{-t^2}dt$. Parameter m defines the mode of the density if $m > 0$. Note that the knowledge of C is not required for simulating a realisation of the density. Indeed, its mean μ and variance σ^2 are given by

$$\mu = m + \sqrt{\frac{2\upsilon}{\pi}}\frac{\exp(-m^2/2\upsilon)}{1 + \text{erf}(m/\sqrt{2\upsilon})}, \quad \text{(A.5)}$$

$$\sigma^2 = \upsilon + \frac{m^2}{4} - \left[\mu - \frac{m}{2}\right]^2. \quad \text{(A.6)}$$

Hence, the standard inversion technique of the cumulative distribution function (Devroye, 1986; Gelfand et al., 1992) may be used. First, it consists in simualting an uniform variate $u \sim \mathcal{U}([0, 1])$ and then in calculating:

$$x = m + \sqrt{2\upsilon}\ \text{erf}^{-1}\left(u + \text{erf}\left(m/\sqrt{2\upsilon}\right)(u-1)\right),$$

The erf function is approximated numerically in practice. In cases where the approximation error becomes important (*i.e.,* when $|m|$ is too large), this simulation method is inefficient. Instead, we use efficient alternatives which are based on accept-reject algorithms (Robert, 1995), the most powerful relies on multiple instrumental distributions (Mazet et al., 2005)[9].

A.3. Gamma density

The gamma density for variable $x \in \mathbb{R}_+$ with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, is defined by

$$\mathcal{G}(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\exp(-\beta x)\mathbb{I}_{\mathbb{R}_+}(x) \quad \text{(A.7)}$$

where $\Gamma(x)$ is the gamma function defined as

$$\Gamma(x) = (x-1)\Gamma(x-1) = \int_0^\infty t^{x-1}e^{-t}dt. \quad \text{(A.8)}$$

The mean and variance are respectively given by E $[x] = \alpha/\beta$ and var $[x] = \alpha/\beta^2$, while the mode of the distribution is given by $M_0 = (\alpha-1)/\beta$. Particular cases of the gamma distribution are the *Erlang distribution* $\mathcal{G}(x|\alpha, 1)$, the *exponential* distribution $\mathcal{E}(x|\beta) = \mathcal{G}(x|1, \beta)$ and the chi-squared distribution $\mathcal{G}(x|v/2, 1/2)$ denoted by χ_v^2. The Erlang distribution may be used in pratice as a preliminary step for simulating a gamma variate $x \sim \mathcal{G}(\alpha, \beta)$ since we get a right sample if $x = u/\beta$ and $u \sim \mathcal{G}(\alpha, 1)$.

Note also that the inverse gamma distribution, denoted by $\mathcal{IG}(\alpha, \beta)$ throughout the paper is the distribution of x^{-1} when $x \sim \mathcal{G}(\alpha, \beta)$.

A.4. Beta density

The Beta density for variable $x \in [0, 1]$ with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, is defined by

$$\mathcal{B}e(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}\mathbb{I}_{[0,1]}(x) \quad \text{(A.9)}$$

---

[9] This code available on-line at http://www.iris.cran.uhp-nancy.fr/francais/si/Personnes/Perso Mazet/rpnorm-fr.htm.

where $B\,(\alpha,\,\beta)$ is the Beta function:

$$B(x) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

The mean and variance are respectively given by

$$E[x] = \frac{\alpha}{\alpha+\beta}, \quad \text{and} \quad \text{var}[x] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

The mode of the Beta distribution evolves w.r.t. the domains of $(\alpha,\,\beta)$:

$$M_0 = \begin{cases} \dfrac{\alpha-1}{\alpha+\beta-2} & \text{if } \alpha>1 \text{ and } \beta>1 \\ 0 \text{ and } 1 & \text{if } \alpha<1 \text{ and } \beta<1 \\ 0 & \text{if } \begin{cases} \alpha<1 \text{ and } \beta\geq1 \\ \alpha=1 \text{ and } \beta>1 \end{cases} \\ 1 & \text{if } \begin{cases} \alpha\geq1 \text{ and } \beta<1 \\ \alpha>1 \text{ and } \beta=1 \end{cases} \\ \text{does not exist if } \alpha=\beta=1 \end{cases}$$

Importantly, for simulating Beta-distributed random variate $x\sim\mathcal{B}e(\cdot|\alpha,\beta)$, we proceeds as follows. First, generate two exponential variables $e_1\sim\mathcal{E}(\cdot|\alpha)$ and $e_2\sim\mathcal{E}(\cdot|\beta)$. Second, compute x as $e_1/(e_1+e_2)$.

A.5. gamma-Gaussian density

The non-standard gamma-Gaussian density for variable $x\in\mathbb{R}_+$ with shape $\alpha>0$, mean $\mu>0$ and variance parameters $\upsilon>0$, is defined by

$$\mathcal{GN}(x|\alpha,\mu,\upsilon) = K_\alpha^{-1}x^{\alpha-1}\exp\left(-(x-\mu)^2/2\upsilon\right)\mathbb{I}_{\mathbb{R}_+}(x) \qquad (A.10)$$

where K_α is the normalising constant:

$$\begin{aligned} K_\alpha &= \int_0^{+\infty} x^{\alpha-1}\exp\left(-(x-\mu)^2/2\upsilon\right)dx \\ &= \exp\left(-\mu^2/4\upsilon\right)\upsilon^{\frac{\alpha}{2}}\Gamma(\alpha)\mathcal{D}_\alpha\left(-\mu/\sqrt{\upsilon}\right) \end{aligned} \qquad (A.11)$$

The last equation follows from (Gradshteyn and Ryzhik, 1994, p. 337, Eq 3.462-1) that relies upon the gamma function (see (A.8)) and the Parabolic Cylinder functions D_ν (Gradshteyn and Ryzhik, 1994, p. 885, Eq 7.711-1 and p. 1065)[10]. Importantly, the first two centered moments of a gamma-Gaussian random variable $x\sim\mathcal{GN}(x|\alpha,\mu,\upsilon)$ can be computed analytically:

$$E[x|\alpha,\mu,\upsilon] = \frac{1}{K_\alpha}\int_0^{+\infty} x^\alpha\exp\left(-(x-\mu)^2/2\upsilon\right)dx = \frac{K_{\alpha+1}}{K_\alpha} \qquad (A.12)$$

$$\text{var}[x|\alpha,\mu,\upsilon] = E\left[x^2|\alpha,\mu,\upsilon\right] - E[x|\alpha,\mu,\upsilon]^2 = \frac{K_{\alpha+2}}{K_\alpha} - \frac{K_{\alpha+1}^2}{K_\alpha^2}. \qquad (A.13)$$

Simulating gamma-Gaussian random variables $x\sim\mathcal{GN}(\cdot|\alpha,\mu,\upsilon)$ is not straightforward and thus more effortful as compared to

sampling from standard laws. Following (Moussaoui et al., 2006), to solve for this problem, we resort to a Metropolis-Hastings algorithm which needs the specification of an instrumental distribution q (Hastings, 1970; Robert, 2001). To avoid high rejection rate, this instrumental pdf has to be chosen to fit the target distribution $f = \mathcal{GN}(\cdot|\alpha,\mu,\upsilon)$ at best. In this regard, expression (A.10) is useful to characterise f in terms of mode, mean or variance from which the instrumental distribution q may be adjusted. Calculating the first-order derivative of (A.10) w.r.t. x and equating to zero, the mode of f is obtained as the solution of the following second order equation:

$$x^2 - \mu x - \upsilon(\alpha-1) = 0, \quad \text{subject to} \quad x\geq0.$$

Let us denote $\Delta=\mu^2+4\upsilon(\alpha-1)$. The mode of f, which is non-negative by definition, expresses as follows:

$$v = \begin{cases} 0 & \text{if } \Delta<0 \\ \max\left((\mu+\sqrt{\Delta})/2, 0\right) & \text{otherwise.} \end{cases} \qquad (A.14)$$

Hence, the instrumental density q is taken as a truncated normal distribution $\mathcal{N}^+(x|v,\upsilon)$, which is easier to sample from as detailed in Appendix A.2.

A.6. Multinomial density

The density of a multinomial discrete distribution for variable $x=\{x_1,\ldots,x_N\}$ with parameters $\boldsymbol{\pi}=\{\pi_1,\ldots,\pi_N\}$ is defined by

$$\mathcal{MN}_M(x|\boldsymbol{\pi}) = \frac{M!}{\prod_{i=1}^N x_i!}\prod_{i=1}^N \pi_i^{x_i}\mathbb{I}_{\sum_i x_i=M}$$

where $x_i\geq0$, $\pi_i>0$ and $\Sigma_{i=1}^N \pi_i=1$.

A.7. Dirichlet density

Let $\boldsymbol{\delta}=\{\delta_1,\ldots,\delta_N\}$ be some positive parameters. The probabilistic density function of the N-state Dirichlet distribution for variable $\boldsymbol{\pi}=\{\pi_1,\ldots,\pi_N\}$ satisfying $\pi_i\geq0$ with parameters $\boldsymbol{\delta}$, is defined by

$$\mathcal{D}_N(\boldsymbol{\pi}|\boldsymbol{\delta}) = \frac{\Gamma\left(\sum_{i=1}^N \delta_i\right)}{\prod_{i=1}^N \Gamma(\delta_i)}\prod_{i=1}^N \pi_i^{\delta_i-1}\mathbb{I}_{\sum_i \pi_i=1},$$

where $\Gamma(x)$ is, as before, the gamma function (A.8). Parameters δ_i are *prior observation counts* for events governed by π_i. The Dirichlet distribution is the conjugate prior of the parameters of a multinomial distribution. One special case is the *symmetric* Dirichlet distribution where $\delta_i=\delta_0 \ \forall i$. In this case, the density becomes

$$\mathcal{D}_N(\boldsymbol{\pi}|\delta_0) = \frac{\Gamma(N\delta_0)}{\Gamma(\delta_0)^N}\prod_{i=1}^N \pi_i^{\delta_0-1}.$$

The real vector $(X_1/S,\ldots,X_N/S)$ follows a Dirichlet distribution denoted as $\mathcal{D}_N(\cdot|\boldsymbol{\delta})$ if $X_i\sim\mathcal{G}(\delta_i,\beta)$ are independent, and $S=\Sigma_i X_i$. This holds true for any β, so in practice we choose $\beta=1$. This result is very useful in practice for simulating realisations of a Dirichlet process.

Appendix B. Computational details for the MCMC procedure

In this section, we derive the full conditional distributions of the quantities $(\boldsymbol{h}, (\mathtt{a}, \mathbb{q}), \ell$ and $\Theta)$ to be sampled. When the sampling

---

[10] See also http://mahieddine.ichir.free.fr for implementation of Parabolic Cylinder functions.

procedure w.r.t. a given parameter cannot be implemented as a Gibbs sampling step, we provide the reader with the derivations of some relevant instrumental distribution needed in the corresponding Metropolis Hastings move.

B.1. The HRF h and its scale σ_h^2

Let us denote $S_j = \Sigma_m a_j^m X^m$. h is $\mathcal{N}(\mu_h, \Sigma_h)$-distributed with:

$$\Sigma_h^{-1} = \sigma_h^{-2} R^{-1} + \sum_{j=1}^{J} S_j^t \Lambda_j S_j \text{ and } \mu_h = \Sigma_h \sum_{j=1}^{J} \sigma_{\varepsilon_j}^{-2} S_j^t \Lambda_j \left(y_j - P\ell_j \right).$$

$$(B.1)$$

Variance σ_h^2 is simulated according to $p(\sigma_h^2|h) = \mathcal{IG}(D/2, h^t R^{-1} h/2)$.

B.2. The nuisance variables ℓ and their scale σ_ℓ^2

Vectors ℓ_j being independent ($j = 1 : J$), they can be sampled in parallel according to $N\left(\mu_{\ell_j}, \Sigma_{\ell_j}\right)$ where

$$\Sigma_{\ell_j}^{-1} = \sigma_\ell^{-2} I_Q + \sigma_{\varepsilon_j}^{-2} P^t \Lambda_j P \text{ and } \mu_{\ell_j} = \sigma_{\varepsilon_j}^{-2} \Sigma_{\ell_j} P^t \Lambda_j \left(y_j - S_j h \right).$$

$$(B.2)$$

Variance σ_ℓ^2 is simulated according to $\mathcal{IG}((QJ+1)/2, \Sigma_j \|\ell_j\|^2/2)$.

B.3. The voxelwise mixtures (q, a)

Although we do not introduce any spatial correlation between the NRLs, the latter are sampled one-at-a time since the distribution $p(a_j^m \mid rest)^{11}$ depends on $a_j^{\backslash m} = [a_j^1, , \ldots, a_j^{m-1}, a_j^{m+1}, \ldots, a_j^M]$ due to the linearity of model (1) with respect to a_j. The sampling of the NRLs is therefore implemented through two nested loops, the inner corresponding to the stimulus types (*e.g.,* index m) and the outer to voxels (*e.g.,* index j). Since the full conditional posterior $p(a_j^m \mid rest)$ is a mixture, its sampling can be achieved in two steps. The first one consists in drawing a realisation of class q_j^m while the second one proceeds conditionally on class q_j^m. To carry out the first step, we need to identify the posterior mixture in voxel j and for condition m:

$$p\left(a_j^m | rest\right) \propto \exp\left(-\frac{1}{2\sigma_{\varepsilon_j}^2} \|e_{j,m} - a_j^m g_m\|_{\Lambda_j}^2\right) \sum_{i=-1}^{1} \lambda_{i,m} f_i\left(a_j^m | \theta_{i,m}\right)$$

$$(B.3)$$

where $g_m = X^m h$, and $e_{j,m} = y_j - P\ell_j - \Sigma_{n \neq m} a_j^n g_n$.

After straightforward calculations including the normalisation of (B.3), we get the following expression:

$$p\left(a_j^m | rest\right) = \sum_{i=-1}^{1} \lambda_{i,j}^m f_i\left(a_j^m | \theta_{i,m}^m\right), \quad \text{with :}$$

$$(B.4)$$

$$f_0\left(a_j^m | \theta_{0,j}^m\right) = \mathcal{N}\left(a_j^m | \mu_{0,j}^m, \upsilon_{0,j}^m\right)$$

$$(B.5)$$

$$f_1\left(a_j^m | \theta_{1,j}^m\right) = \mathcal{GN}\left(a_j^m | \alpha_{1,m}, \mu_{1,j}^m, \upsilon_{1,j}^m\right)$$

$$(B.6)$$

---

[11] *rest* stands for the "remaining variables".

$$f_{-1}\left(a_j^m | \theta_{-1,j}^m\right) = -\mathcal{GN}\left(-a_j^m | \alpha_{-1,m}, -\mu_{-1,j}^m, \upsilon_{-1,j}^m\right)$$

$$(B.7)$$

The mixing probabilities are given by

$$\lambda_{i,j}^m = \left(1 + \sum_{i' \neq i} \widetilde{\lambda}_{i',j}^m / \widetilde{\lambda}_{i,j}^m\right)^{-1}, \quad \forall i = -1 : 1$$

$$\widetilde{\lambda}_{0,j}^m = \lambda_{0,m} \left(\upsilon_{0,j}^m/\upsilon_{0,m}\right)^{1/2} \exp\left(\left(\mu_{0,j}^m\right)^2 / 2\upsilon_{0,j}^m\right),$$

$$\widetilde{\lambda}_{i,j}^m = \lambda_{i,m} \frac{\beta_{i,m}^{\alpha_{i,m}}}{\Gamma(\alpha_{i,m})} K_{i,j}^m \exp\left(\left(\mu_{i,j}^m\right)^2 / 2\upsilon_{i,j}^m\right), \quad \text{for} \quad i \in \{-1, 1\},$$

and depend on the normalising constants $K_{i,j}^m$, for $i = \pm 1$, of the corresponding gamma-Gaussian densities; see (A.11) for its closed-form expression.

The parameters of the conditional posterior densities $f_i(a_j^m \mid \theta_{i,j}^m)$ are given by:

$$\begin{cases} \upsilon_{1,j}^m = \upsilon_{-1,j}^m = \sigma_{\varepsilon_j}^2 \left(g_m^t \Lambda_j g_m\right)^{-1}, \quad \upsilon_{0,m}^m = \left(\upsilon_{0,m}^{-1} + \left(\upsilon_{1,j}^m\right)^{-1}\right)^{-1} \\ \mu_{i,j}^m = \upsilon_{1,j}^m \left(\sigma_{\varepsilon_j}^{-2} g_m^t \Lambda_j e_{m,j} - i\beta_{i,m}\right), \quad \forall i = -1 : 1. \end{cases}$$

$$(B.8)$$

Sampling the class q_j^m first amounts to generating $u_j^m \sim \mathcal{U}([0, 1])$ and then to applying the following rules:

$$q_j^m \begin{cases} -1 \text{ if } u_j^m \leqslant \lambda_{-1,j}^m, \\ 0 \text{ if } \lambda_{-1,j}^m < u_j^m \leqslant \lambda_{-1,j}^m + \lambda_{0,j}^m, \\ 1 \text{ otherwise.} \end{cases}$$

Once q_j^m is correctly set, it remains to sample from the conditional distribution $f_i(\cdot \mid \theta_{i,j}^m)$ as suggested by (B.4). If $q_j^m = 0$, this operation is straightforward because $f_0(\cdot \mid \mu_{0,j}^m, \upsilon_{0,j}^m)$ is Gaussian (*cf.* (B.5)). However, if $q_j^m = \pm 1$, this operation is computationally more expensive since are gamma-Gaussian; see Appendix A.5 for details. Strictly speaking, the sampling of $f_{-1}(\cdot \mid \theta_{1,j}^m)$ consists first in simulating a realisation of $(-a_j^m)$ using a well-suited positive normal density and then negating that realisation.

Interestingly, when $\alpha_{i,m} = 1$, which corresponds to taking an exponential prior for the NRL distribution, the conditional posterior density $f_i(\cdot \mid \theta_{i,j}^m)$ is exactly a truncated normal distribution with parameters equal to those of p_i. In that case, the Metropolis-Hastings is not necessary since all proposals are accepted (the acceptation rate equals to 1). The sampling of the truncated normal density can be achieved efficiently as detailed in Appendix A.

B.4. Mixture probabilities

Since we have $\Pr(q_j^m = i \mid \lambda_m) = \lambda_{i,m}$, for $i = -1 : 1$ and because the prior is conjugate, *i.e.,* a symmetric Dirichlet distribution $D(\lambda_m \mid \delta)$ with $\delta > 0$, the full conditional posterior distribution is also Dirichlet and reads:

$$p\left(\lambda_m | q_j^m = i, \delta\right) \propto Pr\left(q_j^m = i | \lambda_m\right) p(\lambda_m | \delta) \sim D(\delta'), \text{ with } \delta_i'$$
$$= \delta + 1, \text{ and } \delta_l' = \delta, \quad \forall l \neq i.$$

The spatial correlation being not modelled in (6), we may write $\Pr(q^m = \kappa \mid \lambda_m) = \Pi_j \Pr(q_j^m = \kappa_j \mid \lambda_m)$, with $\kappa_j = -1 : 1$ and the joint posterior distribution $p(\lambda_m \mid \kappa, \delta)$ is given by

$$p(\lambda_m | j, \delta) \propto \Pi_j \Pr\left(q_j^m = \kappa_j | \lambda_m\right) p(\lambda_m | \delta) \sim D(\delta')$$
$$\text{with} \quad \delta_i' = \delta + \underbrace{\text{Card}\left[C_{i,m} = \left\{j \in 1 : J | q_j^m = i\right\}\right]}_{= J_{i,m}}. \quad (\text{B.9})$$

B.5. Mixture hyper-parameters

Variance $v_{0,m}$ is very easy to sample because $p(v_{0,m} \mid z^m) = IG((J_{0,m} - 1)/2, \ v_{0,m}/2)$, where $v_{0,m} = \Sigma_{j \in C_0, m}(a_j^m)^2$.

For the two other classes, we proceed to the sampling of the scale and shape gamma distribution parameters $\alpha_{\pm 1, m}$ and $\beta_{\pm 1, m}$, respectively. Following (Moussaoui et al., 2006), we use a Metropolis-Hastings step for $\alpha_{\pm 1, m}$ parameters with a gamma instrumental density defined below. Simulating parameters $\beta_{\pm 1, m}$ is easier since they follow a gamma distribution.

The posterior density of each hyper-parameter $\alpha_{i,m}$ takes the form

$$p(\alpha_{i,m} | rest) \propto \prod_{j \in C_{i,m}} \frac{\beta_{i,m}^{\alpha_{i,m}}}{\Gamma(\alpha_{i,m})} \left(a_j^m\right)^{\alpha_{i,m}-1} p(\alpha_{i,m} | s_i) \propto g(\alpha_{i,m})^{J_{i,m}} \mathbb{I}_{\mathbb{R}_+}(\alpha_{i,m}),$$
$$(\text{B.10})$$

where

$$g(\alpha_{i,m}) = \exp(\tau_{i,m} \alpha_{i,m}) / \Gamma(\alpha_{i,m}),$$
$$\tau_{i,m} = \ln \beta_{i,m} + \sum_{j \in C_{i,m}} \left(a_j^m - s_i\right) / J_{i,m}.$$

This posterior distribution does not belong to a known family, so its simulation requires a MH jump. Akin to (Moussaoui et al., 2006), to obtain a good instrumental law $q(\alpha_{i,m})$, we propose to approximate function $g(\alpha_{i,m})$ using a gamma density $\mathcal{G}(t_{i,m}, u_{i,m})$. More precisely, parameters $(u_{i,m}, t_{i,m})$ of this density are determined in order for its mode and inflexion points match those of function $g(\alpha_{i,m})$. After some simple manipulations, we obtain:

$$t_{i,m} = 1 + \alpha_{mode}^2 / (\alpha_{mode} - \alpha_{infl})^2, \ u_{i,m} = \alpha_{mode} / (\alpha_{mode} - \alpha_{inf1})^2,$$
$$(\text{B.11})$$

where α_{mode} and α_{infl} are the mode and the superior inflexion point $(\alpha_{infl} > \alpha_{mode})$ of $g(\alpha_{i,m})$. Calculating the first and second derivatives of $g(\alpha_{i,m})$ yields these two non-linear equations that implicitly define α_{mode} and α_{infl}:

$$\psi(\alpha_{mode}) = \tau_{i,m} \quad \text{and} \quad \psi^{(1)}(\alpha_{inf1}) = \left(\psi(\alpha_{inf1}) - \tau_{i,m}\right)^2, \quad (\text{B.12})$$

where ψ is the digamma function defined by $\psi(x) = \frac{d}{dx} \log \Gamma(x)$ and $\psi^{(1)}$ is its first derivative (trigamma function). Details about these functions are provided in (Abramowitz and Stegun, 1970, p. 253). The resolution of the two Eq.s (B.12) is done using a root finding numerical method (cf. Press et al., 1992, Ch. 9)). Finally, the posterior density (B.10) is simulated using a Metropolis-Hastings algorithm with a instrumental density $q(\alpha_{i,m})$ chosen as a gamma distribution $\mathcal{G}(t'_{i,m}, u'_{i,m})$ whose parameters are given by

$$t'_{i,m} = J_{i,m}(t_{i,m} - 1) + 1, \ u'_{i,m} = J_{i,m} u_{i,m}. \quad (\text{B.13})$$

The sampling of shape parameters $\beta_{i,m}$ is done according to the full conditional posterior distribution

$$p(\beta_{i,m} | rest) \propto \beta_{i,m}^{(J_{i,m}\alpha_{i,m} + b_i)} \exp\left(-\beta_{i,m}\left(\sum_{j \in C_{i,m}} a_j^m + c_i\right)\right) \quad (\text{B.14})$$
$$\sim \mathcal{G}\left(J_{i,m}\alpha_{i,m} + b_i + 1, \ \sum_{j \in C_{i,m}} a_j^m + c_i\right).$$

B.6. Noise variances

Sampling the noise variances σ_ε^2 can be performed in parallel. Drawing a noise variance is straightforward because $p\left(\sigma_{\varepsilon_j}^2 | rest\right) = \mathcal{IG}\left((N+1)/2, \| \widetilde{y}_j \|_{A_j}^2 / 2\right)$.

B.7. AR parameters

For each voxel V_j, we have:

$$p_j(p_j) = p(p_j | rest) \propto \sqrt{1 - p_j^2} \ \exp\left(-a_j(p_j - m_j)^2\right) \mathbb{1}_{(-1,1)}(p_j),$$
$$(\text{B.15})$$

where $a_j = A_j / 2\sigma_{\varepsilon_j}^2$ and $m_j = B_j / A_j$, with $A_j = \sum_{n=2}^{N-1} \widetilde{y}_{j,n}^2$ and $B_j = \sum_{n=1}^{N-1} \widetilde{y}_{j,n} \widetilde{y}_{j,n+1}$.

The density p_j is log-concave, unfortunately it does not seem to belong to a referenced family of pdf, from which an efficient sampling technique would be available. Here, we propose to resort to a Metropolis-Hastings independence algorithm that uses a beta pdf $g_j \sim Be(\zeta_j, \kappa_j)$ defined over $(-1, 1)$[12] as the instrumental distribution:

$$g_i(\rho) \propto (1 + \rho)^{\zeta_j - 1} (1 - \rho)^{\kappa_j - 1}, \quad \forall |\rho| < 1. \quad (\text{B.16})$$

The parameters ζ_j and κ_j have to be tuned in an appropriate way, so that g_j approximates p_j as closely as possible. Here, ζ_j and κ_j are chosen in such a way that $\log g_j$ and $\log p_j$ have the same curvature around a common maximizer over $(-1, 1)$. Let us first remark that the maximizer r_j of $\log p_j$ is uniquely defined by

$$2a_j(r_j - m_j)\left(1 - r_j^2\right) + r_j = 0, \quad |r_j| < 1.$$

Moreover, r_j takes an explicit expression, as the root of a polynomial of degree three. Then, (ζ_j, κ_j) can be found by solving

$$\begin{cases} (\log g_j)'(r_j) = 0 \\ (\log g_j)''(r_j) = (\log p_j)''(r_j) \end{cases}$$

which is a linear system. After some straightforward calculations, the solution can be expressed as follows:

$$\zeta_j = a_j(1 + r_j)^2 (1 + m_j - 2r_j) + 3/2$$
$$\kappa_j = a_j(1 - r_j)^2 (1 - m_j + 2r_j) + 3/2$$

It can be practically checked that $g_j(\rho)$ and $p_j(\rho)$ take very similar values for all $\rho \in (-1, 1)$. Therefore, the proposal ρ'_j (sampled from g_j) has a high acceptation probability min $\{1, p_j(\rho'_j)g_j(\rho_j)/p_j(\rho_j)g_j(\rho'_j)\}$. In practice, the worse acceptation ratio that we observed was about 0.92.

---

[12] If $x \in (0, 1)$ and $x \sim Be(\zeta, \kappa)$ then $\rho = 2x - 1$ is said $Be(\zeta, \kappa)$-distributed over $(-1, 1)$.

Appendix C. Bayesian model comparison

We have introduced Bayesian model comparison through the computation of Bayes factors:

$$BF_{mn} \triangleq \frac{p(y|\mathcal{M}_m)}{p(y|\mathcal{M}_n)} = \frac{\int p(y|\boldsymbol{\theta}_m, \mathcal{M}_m) p(\boldsymbol{\theta}_m|\mathcal{M}_m) d\boldsymbol{\theta}_m}{\int p(y|\boldsymbol{\theta}_n, \mathcal{M}_n) p(\boldsymbol{\theta}_n|\mathcal{M}_n) d\boldsymbol{\theta}_n},$$

with $(m, n) \in \mathbb{N}_4^*$.

These quantities are thus computed as the ratios of integrated likelihoods or *model evidence* of the different models. Of course, they are approximated from the MCMC outputs using the methodology proposed in (Kass and Raftery, 1995; Chib, 1995; Chib and Jeliazkov, 2001) and further developed in (Raftery et al., 2007). The latter relies on the *harmonic mean identity*:

$$\frac{1}{p(y|\mathcal{M}_*)} = E\left[\frac{1}{p(y|\boldsymbol{\theta}, \mathcal{M}_*)} | y, \mathcal{M}_*\right]$$

This suggests that the model evidence can be approximated by the harmonic mean of the likelihoods $p(y | \boldsymbol{\theta}^{(l)}, \mathcal{M}_*)$ based on L draws $\boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(L)}$ from the posterior distribution $p(\boldsymbol{\theta} | y, \mathcal{M}_*)$:

$$p(\widetilde{y|\mathcal{M}_*}) = \left[\frac{1}{L}\sum_{l=L_0}^{L_1} \frac{1}{p(y|\boldsymbol{\theta}^{(l)}, \mathcal{M}_*)}\right]^{-1} \qquad (C.1)$$

with $L = L_1 - L_0 + 1$. These sample might come out of a standard MCMC implementation. Although $p(\widetilde{y | \mathcal{M}_*})$ is consistent as the sample size L increases, its precision is not guaranteed: it may have an infinite variance. Therefore, we have implemented a *stabilized version* of $p(\widetilde{y | \mathcal{M}_*})$ which is presented in detail in (Raftery et al., 2007). In short, it consists in replacing $p(y | \boldsymbol{\theta}^{(l)}, \mathcal{M}_*)$ by $p(y | f(\boldsymbol{\theta}^{(l)}), \mathcal{M}_*)$ in (C.1) such that f is a measurable function of $\boldsymbol{\theta}$ and a *dimension reduction* transformation. Since the voxels are assumed independent in space, one may proceed separately for each voxel: $p(y | f(\boldsymbol{\theta}^{(l)}), \mathcal{M}_*) = \Pi_j p(y_j | f(\boldsymbol{\theta}_j^{(l)}), \mathcal{M}_*)$. More precisely, this means that f can be derived by integrating out analytically the NRLs $\mathbf{a}_j$ and the noise variance $\sigma_{\varepsilon_j}^2$ for each voxel V_j. This seems sufficient to ensure that $\mathrm{var}\left[(p(y | f(\boldsymbol{\theta}), \mathcal{M}_*))^{-1} | y\right] < \infty$. Hence, in practice we consider the following estimator:

$$p(\widetilde{y|\mathcal{M}_*}) = \left[\frac{1}{L}\sum_{l=L_o}^{L_1} \frac{1}{\Pi_j p(y_j | f(\boldsymbol{\theta}_j^{(l)}), \mathcal{M}_*)}\right]^{-1}.$$

Doing so, we have computed the log-evidence $\log p(y | \mathcal{M}_m)$ of models $\mathcal{M}_1$-$\mathcal{M}_5$ once they have been fitted against the first set of artificial data. Then, the logarithms of Bayes factors have been derived:

$$\log BF_{mn} = \log p(y|\mathcal{M}_m) - \log p(y|\mathcal{M}_n), \quad \forall (m,n) \in \mathbb{N}_4^*.$$

The same procedure has been applied to models $\mathcal{M}_2 - \mathcal{M}_5$ with artificial data eliciting deactivations.

References

Abramowitz, M., Stegun, I.A., 1970. Handbook of mathematical functions. Dover publications, New York, ny.

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The variability of humain bold hemodynamic responses. NeuroImage 7, 574.

Andrade, A., Kherif, F., Mangin, J.-F., Worsley, K., Paradis, A.-L., Simon, O., Dehaene, S., Poline, J.-B., 2001. Detection of fMRI activation using cortical surface mapping. Hum. Brain Mapp. 12, 79–93.

Bagshaw, A.P., Hawco, C., Bénar, C.-G., Kobayashi, E., Aghakhani, Y., Dubeau, F., Pike, G.B., Gotman, J., 2005. Analysis of the EEG-fMRI response to prolonged bursts of interictal epileptiform activity. Neuro-Image 24, 1099–1112.

Bénar, C.-G., Grova, C., Kobayashi, E., Bagshaw, A.P., Aghakhani, Y., Dubeau, F., Gotman, J., 2006. EEG-fMRI of epileptic spikes: Concordance with EEG source localization and intracranial EEG. NeuroImage 30, 1161–1170.

Benboudjema, D., Pieczynski, W., August 2007. Unsupervised statistical segmentation of nonstationary images using triplet Markov fields. IEEE Trans. Pattern Anal. Mach. Intell. 29 (8), 1367–1378.

Buxton, R., Frank, L., 1997. A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. J. Cereb. Blood Flow Metab. 17 (1), 64–72.

Chib, S., 1995. Marginal likelihood from the Gibbs output. J. Am. Stat. Assoc. 90, 1313–1321.

Chib, S., Jeliazkov, I., 2001. Marginal likelihood from the Metropolis-Hastings output. J. Am. Stat. Assoc. 96 (453), 270–281.

Ciuciu, P., Poline, J.-B., Marrelec, G., Idier, J., Pallier, C., Benali, H., October 2003. Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment. IEEE Trans. Med. Imag. 22 (10), 1235–1251.

Ciuciu, P., Idier, J., Roche, A., Pallier, C., April 2004. Outlier detection for robust region-based estimation of the hemodynamic response function in event-related fMRI. 2nd Proc. IEEE ISBI. IEEE, Arlington, VA, pp. 392–395.

Ciuciu, P., Idier, J., Makni, S., June 11-15 2006. Modeling non-linear and non-stationary effects of the BOLD response using mixture models in fMRI. Proc. 12th HBM CD-Rom. Elsevier, Florence, Italy.

Ciuciu, P., Idier, J., Veit, T., Vincent, T., September 2007. Application du rééchantillonnage stochastique de l'échelle en détection-estimation de l'activité cérébrale par IRMf. Actes du 21e colloque GRETSI. GRETSI, Troyes, France, pp. 373–376.

Cohen, M.S., 1997. Parametric analysis of MRI data using linear systems methods. NeuroImage 6, 93–103.

Dehaene-Lambertz, G., Dehaene, S., Anton, J.-L., Campagne, A., Ciuciu, P., Dehaene, G.P., Denghien, I., Jobert, A., Le Bihan, D., Sigman, M., Pallier, C., Poline, J.-B., 2006. Functional segregation of cortical language areas by sentence repetition. Hum. Brain Mapp. 27, 360–371.

Devroye, L., 1986. Non-Uniform Random Variate Generation. Springer Verlag, New York, USA. Available from http://jeff.cs.mcgill.ca/~luc/rnbookindex.html.

Drobnjak, I., Gavaghan, D., Suli, E., Pitt-Francis, J., M., J., 2006. Development of a functional magnetic resonance imaging simulator for modeling realistic rigid-body motion artifacts. Magn. Reson. Med. 56, 364–380.

Everitt, B.S., Bullmore, E.T., 1999. Mixture model mapping of brain activation in functional magnetic resonance images. Hum. Brain Mapp. 7, 1–14.

Flandin, G., Penny, W.D., February 2007. Bayesian fMRI data analysis with sparse spatial basis function priors. NeuroImage 34 (3), 1108–1125.

Flandin, G., Kherif, F., Pennec, X., Malandain, G., Ayache, N., Poline, J.-B., September 2002. Improved detection sensitivity of functional MRI data using a brain parcellation technique. Proc. 5th MICCAI. LNCS 2488 (Part I). Springer Verlag, Tokyo, Japan, pp. 467–474.

Friston, K., 1994. Statistical parametric mapping. In: Thatcher, R., Hallet, M., Zeffiro, T., John, E., Huerta, M. (Eds.), Functional Neuroimaging: Technical Foundations, pp. 79–93.

Friston, K., 1998. Imaging neuroscience: Principles or maps? Proc. Natl. Acad. Sci. U. S. A. 95, 796–802.

Gelfand, A., Smith, A., Lee, T.-M., 1992. Bayesian analysis of constrained parameter and truncated problems using Gibbs sampling. J. Am. Stat. Assoc. 87 (518), 523–532.

Geman, S., McClure, D., 1987. Statistical methods for tomographic image reconstruction. Proceedings of the 46th Session of the ICI. Bulletin of the ICI, vol. 52, pp. 5–21.

Genovese, C., 2000. A Bayesian time-course model for functional magnetic resonance imaging data (with discussion). J. Am. Stat. Assoc. 95, 691–719.

Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage 9, 416–429.

Gössl, C., Auer, D.P., Fahrmeir, L., June 2001. Bayesian spatio-temporal modeling of the hemodynamic response function in BOLD fMRI. Biometrika 57, 554–562.

Goutte, C., Nielsen, F.A., Hansen, L.K., December 2000. Modeling the haemodynamic response in fMRI using smooth filters. IEEE Trans. Med. Imag. 19 (12), 1188–1201.

Gradshteyn, I., Ryzhik, I., 1994. Table of Integrals, Series, and Products, 5th ed. Academic Press, 1250 Sixth Avenue, San Diego CA.

Green, P.J., March 1990. Bayesian reconstructions from emission tomography data using a modified EM algorithm. IEEE Trans. Med. Imag. 9 (1), 84–93.

Green, P.J., 1995. Reversible jump MCMC computation and Bayesian model determination. Biometrika 82, 711–732.

Grill-Spector, K., Malach, April 2001. fmri-adaptation: a tool for studying the functional properties of human cortical neurons. Acta Psychol. (Amst) 107 (1-3), 293–321.

Grill-Spector, K., Sayres, R., Ress, D., Sep 2006. High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. Nat. Neurosci. 9 (9), 1177–1185.

Grova, C., Makni, S., Flandin, G., Ciuciu, P., Gotman, J., Poline, J.-B., 2006. Anatomically informed interpolation of fMRI data on the cortical surface. NeuroImage 31, 1475–1486.

Handwerker, D.A., Ollinger, J.M., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. NeuroImage 21, 1639–1651.

Hastings, W.K., January 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97.

Henson, R., Price, C., Rugg, M., Turner, R., Friston, K., 2002. Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. NeuroImage 15 (1), 83–97.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Amer. Statist. Assoc. 90, 773–795.

Kay, S.M., 1988. Modern Spectral Estimation. Prentice-Hall, Englewood Cliffs, nj.

Kruggel, F., Von Crammon, D.Y., 1999. Modeling the hemodynamic response in single-trial functional MRI experiments. Magn. Reson. Med. 42, 787–797.

Lange, N., 1997. Empirical and substantive models, the Bayesian paradigm, and meta-analysis in functional brain imaging. Hum. Brain Mapp. 5, 259–263.

Liu, J., 2001. Monte Carlo strategies in scientific computing. Springer series in Statistics. Springer-Verlag, New-York.

Makni, S., Ciuciu, P., Idier, J., Poline, J.-B., September 2005. Joint detection-estimation of brain activity in functional MRI: a multichannel deconvolution solution. IEEE Trans. Signal Process. 53 (9), 3488–3502.

Makni, S., Ciuciu, P., Idier, J., Poline, J.-B., May 2006a. Bayesian joint detection-estimation of brain activity using MCMC with a Gamma-Gaussian mixture prior model. Proc. 31th Proc. IEEE ICASSP, vol. V. IEEE, Toulouse, France, pp. 1093–1096.

Makni, S., Ciuciu, P., Idier, J., Poline, J.-B., April 2006b. Joint detection-estimation of brain activity in fMRI using an autoregressive noise model. 3th Proc. IEEE ISBI. IEEE, Arlington, VA, pp. 1048–1051.

Mangin, J.-F., Frouin, V., Bloch, I., Régis, J., López-Krahe, J., 1995. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. J. Math. Imaging Vis. 5, 297–318.

Marrelec, G., Benali, H., Ciuciu, P., Pélégrini-Issac, M., Poline, J.-B., May 2003. Robust Bayesian estimation of the hemodynamic response function in event-related BOLD MRI using basic physiological information. Hum. Brain Mapp. 19 (1), 1–17.

Marrelec, G., Ciuciu, P., Pélégrini-Issac, M., Benali, H., August 2004. Estimation of the hemodynamic response function in event-related

functional MRI: Bayesian networks as a framework for efficient Bayesian modeling and inference. IEEE Trans. Med. Imag. 23 (8), 959–967.

Mazet, V., Brie, D., Idier, J., 2005. Simulation of positive normal variables using several proposal distributions. IIEEE workshop on statistical signal processing. Bordeaux, France.

Miezin, F.M., Maccotta, L., Ollinger, J.M., Petersen, S.E., Buckner, R.L., 2000. Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. NeuroImage 11, 735–759.

Moussaoui, S., Brie, D., Mohammad-Djafari, A., Carteret, C., November 2006. Separation of non-negative mixture of non-negative sources using a Bayesian approach and MCMC sampling. IEEE Trans. Signal Process. 54 (11), 4133–4145.

Naccache, L., Dehaene, S., 2001. Unconscious semantic priming extends to novel unseen stimuli. Cognition 80, 215–229.

Neumann, J., Lohmann, G., 2003. Bayesian second-level analysis of functional magnetic resonance images. NeuroImage 20 (2), 1346–1355.

Nielsen, F.A., Hansen, L.K., Toft, P., Goutte, C., Lange, N., Stroher, S.C., Morch, N., Svarer, C., Savoy, R., Rosen, B., Rostrup, E., Born, P., 1997. Comparison of two convolution models for fMRI time series. NeuroImage 5, S473.

Nieto-Castanon, A., Ghosh, S., Tourville, J., Guenther, F., 2003. Region of interest based analysis of functional imaging data. NeuroImage 19 (4), 1303–1316.

Ogawa, S., Lee, T., Kay, A., Tank, D., 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc. Natl. Acad. Sci. USA 87 (24), 9868–9872.

Ou, W., Golland, P., July 2005. From spatial regularization to anatomical priors in fMRI analysis. IPMI. Glenwood Springs, Colorado.

Penny, W., Friston, K., Apr 2003. Mixtures of general linear models for functional neuroimaging. IEEE Trans. Med. Imag. 22 (4), 504–514.

Penny, W.D., Kiebel, S., Friston, K.J., 2003. Variational Bayesian inference for fMRI time series. Neuroimage 19 (3), 727–741.

Press, W., Teukolsky, S., Vetterling, W., Flannery, B., 1992. Numerical Recipes in C- The Art of Scientific Computing, vol. 15.7. Cambridge University Press, Ch, pp. 699–706.

Raftery, A.E., Newton, M.A., Satagopan, J.M., Krivitsky, P.N., 2007. Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In: Bernardo, J., Bayarri, M., Berger, O., David, A., Heckermann, D., Smith, A., West, M. (Eds.), Bayesian statistics 8. Oxford University Press, pp. 1–45.

Rajapakse, J.C., Kruggel, F., Maisog, J.M., Von Cramon, D., 1998. Modeling hemodynamic response for analysis of functional MRI time-series. Hum. Brain Mapp. 6, 283–300.

Richardson, S., Green, P.J., 1997. On Bayesian analysis of mixtures with an unknown number of components (with discussion). J. R. Stat. Soc., B 59 (4), 731–792.

Robert, C., 1995. Simulation of truncated normal variables. Stat. Comput. 5, 121–125.

Robert, C.P., 2001. The Bayesian Choice, Springer Texts in StatisticsSecond Edition. Springer Verlag, New York, ny.

Roberts, S.J., Penny, W.D., September 2002. Variational Bayes for generalized autoregressive models. IEEE Trans. Signal Process. 50 (9), 2245–2257.

Smith, M., Pütz, B., Auer, D., Fahrmeir, L., 2003. Assessing brain activity through spatial Bayesian variable selection. NeuroImage 20, 802–815.

Svensen, M., Kruggel, F., von Crammon, D., January 2000. Probabilistic modeling of single-trial fMRI data. IEEE Trans. Med. Imag. 19, 19–35.

Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J.-B., August 2006. Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets. Hum. Brain Mapp. 27 (8), 678–693.

Thirion, B., Pinel, P., Tucholka, A., Roche, A., Ciuciu, P., Mangin, J.-F., Poline, J.-B., September 2007. Structural analysis of fMRI data revisited: Improving the sensitivity and reliability of fMRI group studies. IEEE Trans. Med. Imag. 26 (9), 1256–1269.

Thyreau, B., Thirion, B., Flandin, G., Poline, J.-B., May 2006. Anatomo-functional description of the brain: a probabilistic approach. Proc. 31th Proc. IEEE ICASSP, vol. V. IEEE, Toulouse, France, pp. 1109–1112.

Vaever Hartvig, N., Jensen, J., 2000. Spatial mixture modeling of fMRI data. Hum. Brain Mapp. 11 (4), 233–248.

Veit, T., Idier, J., September 2007. Rééchantillonnage de l'échelle dans les algorithmes MCMC pour les problémes inverses bilinéaires. Actes du 21e colloque GRETSI. GRETSI, Troyes, France, pp. 1233–1236.

Vincent, T., Ciuciu, P., Idier, J., August 2007a. Application and validation of spatial mixture modelling for the joint detection-estimation of brain activity in fMRI. Proc. of the 29th IEEE EMBS Annual international conference. Lyon, France, pp. 5218–5222.

Vincent, T., Ciuciu, P., Idier, J., April 2007b. Spatial mixture modelling for the joint detection-estimation of brain activity in fMRI. 32th Proc. IEEE ICASSP, vol. I. IEEE, Honolulu, Hawaii, pp. 325–328.

Woolrich, M., Behrens, T., October 2006. Variational Bayes inference of spatial mixture models for segmentation. IEEE Trans. Med. Imag. 25 (10), 1380–1391.

Woolrich, M., Ripley, B., Brady, M., Smith, S., December 2001. Temporal autocorrelation in univariate linear modelling of fMRI data. NeuroImage 14 (6), 1370–1386.

Woolrich, M., Jenkinson, M., Brady, J., Smith, S., February 2004a. Fully Bayesian spatio-temporal modelling of fMRI data. IEEE Trans. Med. Imag. 23 (2), 213–231.

Woolrich, M., Jenkinson, M., Brady, J.M., Smith, S., 2004b. Constrained linear basis set for HRF modelling using variational Bayes. NeuroImage 21 (4), 1748–1761.

Woolrich, M., Behrens, T., Beckmann, C., Smith, S., January 2005. Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data. IEEE Trans. Med. Imag. 24 (1), 1–11.

Worsley, K., Liao, C., Aston, J., Petre, V., Duncan, G., Morales, F., Evans, A., January 2002. A general statistical analysis for fMRI data. NeuroImage 15 (1), 1–15.

Chapitre IX

IMAGERIE FONCTIONNELLE ULTRA-RAPIDE PAR SÉQUENCE *Echo Volumar Imaging*

[a.12] C. Rabrait, P. Ciuciu, A. Ribés, C. Poupon, P. Leroux, V. Lebon, G. Dehaene-Lambertz, D. Le Bihan et F. Lethimonnier, « High temporal resolution functional MRI using parallel echo volume imaging », accepté pour publication dans *Journal of Magnetic Resonance Imaging*, mars 2008. Volume 27, Numéro 4, pages 744-753.

Original Research

High Temporal Resolution Functional MRI Using Parallel Echo Volumar Imaging

Cécile Rabrait, MS,[1]* Philippe Ciuciu, PhD,[1] Alejandro Ribés, PhD,[1] Cyril Poupon, PhD,[1] Patrick Le Roux, PhD,[2] Ghislaine Dehaine-Lambertz, MD, PhD,[1,3] D. Le Bihan, MD, PhD,[1] and F. Lethimonnier, PhD[1]

Purpose: To combine parallel imaging with 3D single-shot acquisition (echo volumar imaging, EVI) in order to acquire high temporal resolution volumar functional MRI (fMRI) data.

Materials and Methods: An improved EVI sequence was associated with parallel acquisition and field of view reduction in order to acquire a large brain volume in 200 msec. Temporal stability and functional sensitivity were increased through optimization of all imaging parameters and Tikhonov regularization of parallel reconstruction. Two human volunteers were scanned with parallel EVI in a 1.5T whole-body MR system, while submitted to a slow event-related auditory paradigm.

Results: Thanks to parallel acquisition, the EVI volumes display a low level of geometric distortions and signal losses. After removal of low-frequency drifts and physiological artifacts, activations were detected in the temporal lobes of both volunteers and voxelwise hemodynamic response functions (HRF) could be computed. On these HRF different habituation behaviors in response to sentence repetition could be identified.

Conclusion: This work demonstrates the feasibility of high temporal resolution 3D fMRI with parallel EVI. Combined with advanced estimation tools, this acquisition method should prove useful to measure neural activity timing differences or study the nonlinearities and nonstationarities of the BOLD response.

Key Words: echo volumar imaging; parallel MRI; SENSE; fMRI; high temporal resolution
J. Magn. Reson. Imaging 2008;27:744–753.
© 2008 Wiley-Liss, Inc.

UNTIL NOW, functional MRI (fMRI) data were mostly acquired using echo planar imaging (EPI), allowing the

744

acquisition of multislice brain volumes with a spatial resolution of about 3 mm and a temporal resolution of 1–2 seconds. EPI is thus well suited to the detection of cerebral activations and the mapping of activated areas. Nevertheless, with the development of event-related fMRI (1) and the growing interest in the temporal features of the hemodynamic response function (HRF) (2,3), higher scanning rates are called for by neuroscientists.

High temporal resolution fMRI is feasible with echo volumar imaging (EVI), a 3D extension of EPI, in which a 3D Fourier space is encoded in a single-shot acquisition. Some applications of EVI have been reported, especially in fMRI (4–6). As compared with EPI, EVI presents several advantages for fMRI. First, EVI acquisition offers very short TR, on the order of 200 msec. Second, 3D single-shot acquisition makes slice-timing correction unnecessary and reduces the risk of intravolume motion of the subject. Finally, true 3D acquisition is known to reduce vascular inflow effects (7) that often confuse the interpretation of fMRI results (8).

Nevertheless, EVI has seldom been used in fMRI, due to its heavy demand on gradient hardware. Actually, because of hardware limitations the echo train duration (ETD) is high in EVI and the voxel bandwidth along the direction of partition (third encoding direction) is very low. Thus, if high spatial resolution and wide brain coverage are required in an EVI acquisition, susceptibility-induced distortions and signal losses due to T2* relaxation dramatically alter image quality. To overcome these difficulties compromises between acquisition time and spatial parameters must be found. Previously reported EVI studies have preferentially used the frequency encoding direction (less expensive in acquisition time) and acquired highly anisotropic volumes. However, this geometry is not suitable to every fMRI experiment and, because of the small number of samples along the partition direction, the Fourier transform is not a reliable estimator of the object's real spin density.

In this context, our purpose was to develop an EVI method allowing single-shot acquisition of a nearly isotropic volume covering most of the human brain. It was achieved, for the first time, by combining EVI with parallel acquisition and SENSE reconstruction, along both
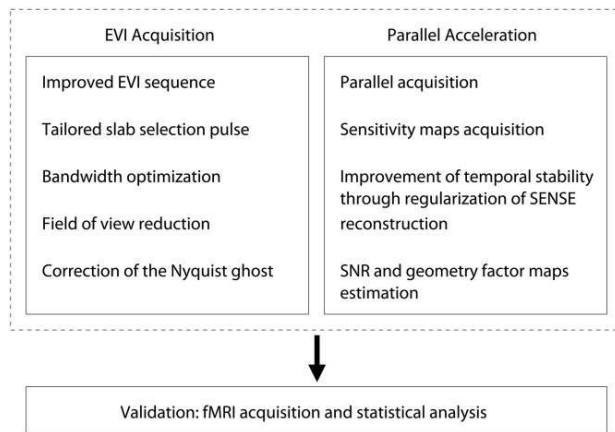
Figure 1. Schematic diagram illustrating the different steps involved in parallel acquisition of EVI volumes.

the phase and partition directions (9). An acceleration factor of 4 was reached, making it possible to acquire a $120 \times 120 \times 144$ mm^3 brain volume in 200 msec. To achieve 3D single-shot acquisition of a large brain volume and keep the level of susceptibility-induced distortions low, the spatial resolution was limited to 6 mm. This spatial scale is commonly used in neuroscience group studies, which involve brain normalization.

In this article the adaptation of 2D-SENSE reconstruction to localized EVI volumes is described. The optimization of all the imaging parameters to maximize the SNR is discussed and the effect of regularization on functional sensitivity is investigated. An auditory, slow event-related functional paradigm is used to demonstrate the feasibility of detecting brain activation and estimating high temporal resolution hemodynamic response functions with parallel EVI.

MATERIALS AND METHODS

To clarify the presentation of the different steps required to achieve parallel EVI, we introduce a schematic diagram describing the whole process in Fig. 1.

EVI Acquisition

In order to prevent the formation of an N/2 ghosting artifact along the partition direction, a modified EVI sequence with rewind gradients was implemented, as previously (6). The area of the rewind gradient, applied along the phase direction after the encoding of each section, exactly compensates for the accumulated area of the phase blipped gradients. Thus, contrary to conventional EVI (10), alternate planes in Fourier space are acquired under phase encoding gradients with the same polarity, as illustrated in Fig. 2.

Slab selective excitation was performed by a spectro-spatial radiofrequency (RF) pulse (11), so that subcutaneous fat was not excited. This pulse was composed of eight spatially selective pulses (duration 1.5 msec), with variable magnitudes and separated by 2.27-ms wait periods, as illustrated in Fig. 2. A slab-selection gradient was applied during the RF pulses and magnetiza-

tion was rephased by another gradient during wait periods. The complete pulse lasts 18.2 msec and its bandwidth is equal to 10 kHz. The spatial selectivity of the pulse, defined as the ratio of the pass bandwidth to the transition bandwidth, was high (around 15) and the suppression of the fat signal was efficient. The spatial selectivity of the slab selection pulse was sufficient to define the thickness of the acquired volume along the partition direction.

In order to limit susceptibility-induced distortions and signal losses, the ETD must be minimized. The ETD was estimated as a function of all relevant imaging parameters. It appears that, in parallel EVI with trapezoidal gradient pulses, ramp-sampling acquisition does not allow one to significantly reduce the ETD, mainly because of the rewind gradient and of the small matrices acquired. Hence, it has not been introduced. Moreover, since the frequency encoding gradient amplitude depends on the acquisition bandwidth and because of the limited slew rate of the gradient hardware, there is an optimal readout bandwidth minimizing the ETD, which depends on Nx, Ny, field of view (FOV) along the readout direction, and gradient rise time. The choice of the optimal bandwidth significantly reduces the ETD. For example, for a $20 \times 20 \times 24$ matrix, with a $120 \times 120 \times 144$ mm^3 FOV and an acceleration factor of 2 along two directions (R = 2×2), the optimal acquisition bandwidth is 100.0 kHz, corresponding to an ETD of 58 msec. Bandwidth of 50 kHz and 200 kHz correspond respectively to ETD of 68 and 74 msec.

To further reduce the ETD the FOV was reduced as proposed previously (5,6). Outer volume suppression (OVS) was applied before excitation in order to avoid wraparound artifacts. Two quadratic phase pulses, designed with the Shinnar–Le Roux algorithm, each of which being followed by spoiler gradient pulses, were applied in conjunction with a slab-selective gradient along the phase direction (12). The quadratic phase pulses (duration: 4 msec, bandwidth: 10 kHz) offer a high spatial selectivity (around 20) and a low peak B_1. To ensure a good suppression of the external signal, OVS pulses were repeated twice, before each excitation.
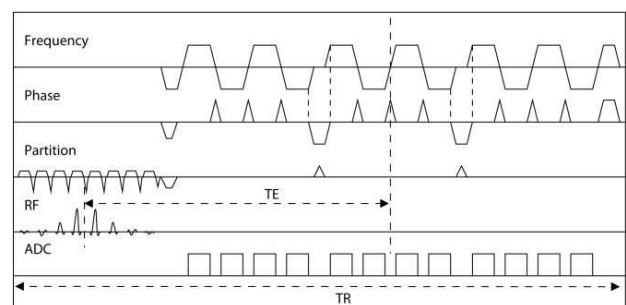
Figure 2. Timing diagram of the EVI sequence: a spectro-spatial slab excitation pulse is followed by a modified EVI echo train. In this example only four phase encoding lines and three sections are acquired. For sake of simplicity, the outer volume suppression module is not included on this diagram. This module takes place before excitation and is presented in Ref. (12).

After 3D reconstruction, EVI volumes are affected by a Nyquist N/2 ghosting artifact along the phase direction, as in EPI. It is due to small timing differences between odd and even echoes. The well-known calibration scan method (13) was used to correct for these differences. A calibration scan was acquired at the beginning of the functional time series, with neither phase nor partition gradients. Calibration data were first Fourier transformed with respect to the frequency encoding direction. Then, Ahn's coefficients (constant and linear) were computed for every phase encoding step (14) and used to correct all the volumes of the temporal series. Contrary to (13), Ahn's coefficients have not been modeled since discontinuities appeared between k-space sections. Moreover, if the ETD is kept short, Ahn's correction is sufficient to reduce ghosting artifact efficiently in EVI volumes at 1.5T.

Parallel Acceleration

Parallel Acquisition

Minimization of the ETD by choosing the optimal acquisition bandwidth and restricting the FOV does not permit reaching short ETD while covering a large part of the brain. As shown by the previous example ($20 \times 20 \times 24$ matrix, $120 \times 120 \times 144$ mm^3 FOV), without parallel acceleration the minimized ETD would be in the 230–240 msec range. The only way to dramatically reduce the ETD is to use parallel imaging in order to undersample k-space acquisition along both the phase and partition directions. In 3D imaging the choice of 2D-SENSE parallel acquisition and reconstruction is advantageous since geometry factors are lower than in 1D-SENSE for a given acceleration factor; the SNR of the reconstructed volumes is thus higher (9). All 2D-SENSE acquisitions were performed in either sagittal or coronal orientations. Indeed, the nearly cylindrical coil we used (8-channel head coil) prohibits SENSE reconstruction along its main axis, since its sensitivity is nearly uniform along this direction. Parallel acquisition also directed the optimization of the imaging parameters. Indeed, the FOV has to be relatively large along both undersampling directions, since parallel reconstruction is easier when the difference of sensitivities between signal components superimposed in a given voxel is more important. Moreover, the additional degrees of freedom offered by parallel imaging (15) allowed reducing acoustic noise through minor parameter tuning. This optimization resulted in a decrease from about 115 dB to an acceptable level of 104 dB.

Parallel Reconstruction: Acquisition of the Sensitivity Maps

Parallel reconstruction of EVI undersampled volumes was performed with an in-house-developed multidimensional SENSE algorithm (16). To perform SENSE reconstruction it was necessary to acquire sensitivity maps for each RF channel. In EVI this step was crucial. It was indeed impossible to use an EVI volume as a sensitivity map, since the high level of geometric distortions and signal losses would make the map useless. EPI images could not be used as sensitivity maps either,

since the distortions are not the same as in parallel EVI. Therefore, conventional 2D gradient echo images were acquired as sensitivity maps. One critical issue was to make sure that distortions remain low in EVI, so that EVI volumes and sensitivity maps could be superimposed. Empirical optimization of the ETD was thus performed by experimenting parallel reconstruction of EVI volumes for a large range of ETD (data not shown). This study led to restricting the ETD below 60 msec, although EPI acquisitions at 1.5T frequently use longer ETD. Full FOV sensitivity maps have been preferred since SNR is higher in full FOV images compared to reduced FOV images. Sensitivity maps were automatically cropped to match the EVI volume after the building of a binary mask of the brain. This mask was used to remove voxels outside of the brain from parallel reconstruction. This "voxel exclusion" operation improved the SNR of the reconstructed volumes (9). Given the low spatial resolution of parallel EVI volumes, the impact of Gibb's ringing artifacts should be considered, as in parallel spectroscopic imaging (17,18). In parallel EVI, as spatial resolution is the same in sensitivity maps and parallel data, Gibb's artifact should be similar. Nevertheless, Gibb's side lobes could be important in parallel data while being excluded from the masked sensitivity maps. To avoid this problem the binary brain mask was dilated by a few voxels as proposed previously (17). Alternatively, a cosine filter could be applied in k-space, which apodizes the reconstructed data while degrading the spatial resolution. A more efficient solution, which improves SENSE reconstruction by taking into account the real point spread function, has also been proposed (18).

As explained previously (19), noise properties should be obtained from artifact-free data, preferably from noise-only samples. In the present study noise properties were estimated from an empty area in the sensitivity maps to reduce the total acquisition duration and because 2D gradient echo is not very sensitive to image artifacts. Gibb's artifact, which is the more likely to occur, should have a small effect since noise properties are estimated outside of the dilated mask.

Parallel Reconstruction: Regularization

Typically, regularization techniques are used in parallel reconstruction to improve the visual quality of static images. Nevertheless, since parallel EVI aims at detecting cerebral activations, temporal signal stability is a predominant criterion over image quality. In the present work, as pioneered previously (20,21), regularization of parallel reconstruction was used to improve signal stability and activation detection. Tikhonov weighting (22) was applied, with a regularization condition minimizing the magnitude of the MR signal in reconstructed volumes, as before (23). The relative importance of the regularization term was modulated by a regularization parameter, λ^2, empirically set.

To investigate the impact of regularization, SNR maps, geometry factor maps, and statistical scores were computed for increasing values of λ^2. This values range from 10^{-4} to 0.2. At $\lambda^2 = 10^{-4}$, reconstructed volumes are highly similar to volumes reconstructed

without regularization, while avoiding computer errors due to small singular values. At $\lambda^2 = 0.2$ statistical scores have reached a plateau, while reconstruction artifacts become important. The number of intermediate settings was kept small because of the strong efforts required to perform the statistical analysis for each setting.

To assess signal stability the SNR was defined in each voxel as the ratio of the average signal over its standard deviation in the voxel time course. This definition, often used in cerebral functional imaging (24,25), is reliable when used with parallel imaging. We note that due to the inhomogeneous spatial distribution of the noise other SNR definitions suffer from serious bias (26). SNR measurements were performed after removal of low-frequency drifts and filtering out of physiological artifacts, as described in the next section. The quality of parallel reconstruction was also assessed through the computation of spatial maps of the geometry factor, *g*, (16). These maps quantify the difficulty to separate the different signal components superimposed in a given voxel. The *g* factor depends on the geometry of the acquired volume and the coil array.

Data Acquisition and Analysis

Two right-handed healthy male adults were scanned using a Signa 1.5T Excite II MR system (General Electric Healthcare, Milwaukee, WI), employing a gradient coil with a maximum amplitude of 40 mT.m^{-1} and a 266-μs minimum rise time. An 8-channel high resolution brain array by MRI Devices was used for parallel acquisition. Acquisitions on human volunteers were approved by a local ethics committee and the subjects gave written informed consent.

For the acquisition of parallel EVI time series the imaging parameters were set as follows: $20 \times 20 \times 24$ matrix, TE/TR = 40/200 msec, 3D-FOV = $120 \times 120 \times 144$ mm^3, BW = 100 kHz, flip angle = 35°. The volumes were acquired in sagittal orientation, with the frequency encoding gradient applied in the superior–inferior direction, and the partition encoding gradient applied in the mediolateral direction. A total parallel reduction factor of 4 was applied, corresponding to a reduction factor of 2 along both phase and partition directions. The ETD was equal to 58 msec, while typical ETD in EPI ranged from 30–80 msec on this scanner. The volume of interest had been optimized for auditory paradigms and included the two temporal lobes of an "average" subject. Moreover, the chosen geometry minimizes vascular artifacts due to inflow of venous blood, because of 3D acquisition and low excitation flip angle (7), and since no "fresh" blood can enter the excited volume along the partition direction.

For SENSE parallel reconstruction, 2D gradient-echo sensitivity maps were acquired in sagittal orientation. Acquisition parameters were set as follows: 40×40 matrix, 24 slices, slice thickness = 6 mm, TE/TR = 10/500 msec, FOV = 240×240 mm^2, BW = 62.5 kHz, flip angle = 30°. A high-resolution T1-weighted 3D volume was acquired for anatomical localization ($256 \times 256 \times 128$ matrix, voxel size = $0.9 \times 0.9 \times 1.2$ mm^3).

The fMRI paradigm used to validate the detection of cerebral activation consisted in a slow-event presentation of auditory sentences, one short sentence (2200 msec on average) every 14.4 seconds. Each sentence was repeated four times in a row, in order to induce habituation effects as in Ref. (27). Each run comprised 12 different sentences for a total duration of 11 minutes, 41 seconds. The first sentence was preceded by a 9.6-second rest period, corresponding to 48 dummy scans. One run was acquired in each participant.

Postprocessing and statistical analysis of the EVI temporal series were conducted using the SPM2 software (www.fil.ion.ucl.ac.uk). Illustrations were created using the Anatomist software (http://brainvisa.info). First, EVI volumes were rotated to match the axial orientation of anatomical data. Second, reconstructed time series were realigned to an EVI volume devoid of artifacts, using a six-parameter linear transformation, and resliced. Third, a numerical band-rejection filter was applied to reduce cardiac (0.8–1.2 Hz) and respiratory artifacts (0.15–0.3 Hz), as in Ref. (6). Physiological artifacts are easily filtered out in EVI thanks to the high sampling frequency. Finally, the volumes were spatially smoothed, using a 10-mm isotropic Gaussian filter. This operation is generally applied to EPI data to assure the accuracy of the statistical inference made from the data in SPM and increase the sensitivity of detection in group studies. It also increases the SNR and thus the sensitivity of detection of brain activations. Rigid coregistration to anatomical data was performed using the Brainvisa software (http://brainvisa.info). Since the EVI volume is relatively large, computation of the coregistration parameters was straightforward and accurate. The removing of low-frequency drifts was performed through modeling of the trend by a basis of discrete cosine functions (28) with a 32-second cutoff period.

A generalized linear model was generated by entering four distinct variables corresponding to the four repetition positions of a sentence. The variables, convolved with the standard SPM HRF and its temporal derivative were included in the model. The significance of the activations in response to the first presentation of each sentence was assessed using Student's *t*-test corrected for multiple comparisons (family-wise error) and thresholded at 10^{-4}.

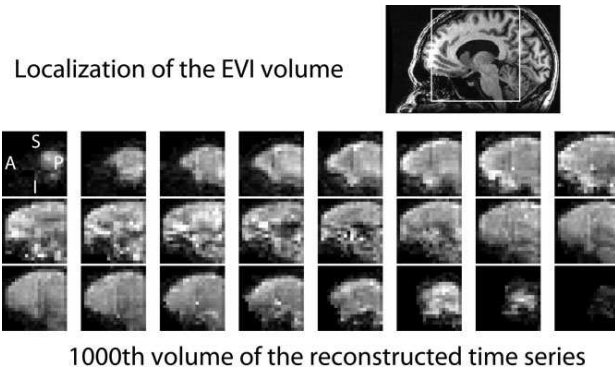## RESULTS

Image Quality and SNR

As illustrated in Fig. 3, parallel imaging improves EVI image quality, as demonstrated for single-shot acquisition methods (19,29). A spherical water phantom was scanned and EVI volumes were acquired, either with no parallel acceleration or with a reduction factor of 4. In the latter case, undersampling was applied along both phase (Ry = 2) and partition (Rz = 2) directions. Parallel acquisition results in a decrease of the ETD by a factor of 4, therefore reducing susceptibility-induced distortions and signal dropout. For comparison, reference images (sum of squares of the sensitivity maps) are also displayed, as well as the folded EVI volume (sum of
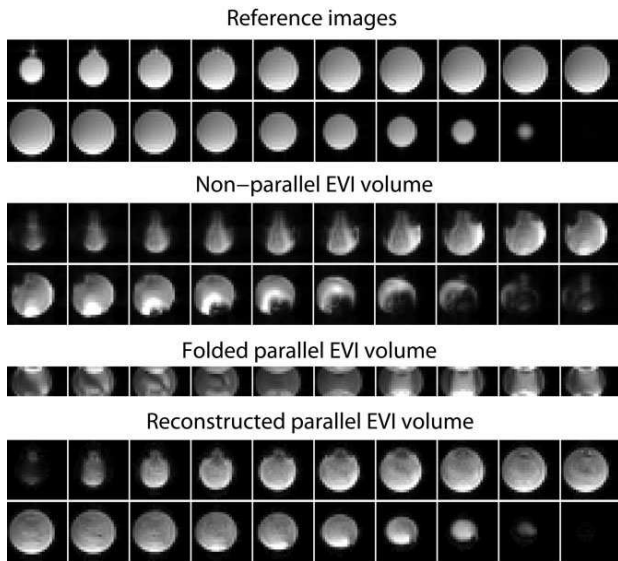
Figure 3. Comparison between nonparallel and parallel EVI volumes. The parallel EVI volume has been acquired with a total parallel acceleration factor of 4. 2D gradient-echo images acquired with the same FOV are displayed as reference. Acquisition parameters: nonparallel EVI volume: Nx = Ny = 24, Nz = 20, Lx = Ly = 160 mm, Lz = 120 mm, TE/TR = 116/252 msec, ETD = 216 msec, BW = 125 kHz, flip angle = 35°; parallel EVI volume: Nx = 24, Ny = 12, Nz = 10, Lx = Ly = 160 mm, Lz = 120 mm, TE/TR = 40/200 msec, ETD = 54 msec, BW = 125 kHz, flip angle = 35°; 2D gradient-echo images: Nx = Ny = 24, 20 slices, Lx = Ly = 160 mm, slice thickness = 6 mm, TE/TR = 10/500 msec, BW = 62.5 kHz, flip angle = 30°.

Figure 4. Parallel EVI volume acquired on a human subject and its localization over one slice of the subject's anatomical data. The 1000th volume of the time series is presented, since it is representative of the volumes studied for activation detection. The sections are presented from right to left of the subject's head. Acquisition parameters for the parallel EVI time series: Nx = 20, Ny = 10, Nz = 12, Lx = Ly = 120 mm, Lz = 144 mm, TE/TR = 40/200 msec, ETD = 58 msec, BW = 100 kHz, flip angle = 35°, regularization parameter $(\lambda^2) = 10^{-2}$.

squares of the folded volumes reconstructed from the undersampled data). The latter comprises one phase-encoding line over two and one section over two. The reconstructed parallel EVI volume is relatively homogenous and appears highly similar to the references images, although the TE (40 msec) is longer than the TE used in the acquisition of reference images (10 msec). This similarity can be partially explained by the application of a mask computed from sensitivity data. Nevertheless, the folded EVI volume, which has not been masked, also displays reduced distortions and better signal homogeneity compared to the nonparallel EVI volume.

The 1000th volume extracted from a parallel EVI temporal series acquired on a subject is shown in Fig. 4. This volume was acquired after the magnetization reached a steady state, and is thus representative of the volumes studied for activation detection. In these volumes the signal is relatively homogenous. The only area of low signal corresponds to the lateral ventricles, filled with cerebral spinal fluid (CSF), which presents a lower steady-state signal than gray and white matter at short TR. The homogeneity of the signal along the partition direction validates the good quality of the slab selection pulse. The slight signal diminution in the middle of the volume can be explained by the so-called "SENSE ghost" (30) which comes from residual signal in the external areas. The small amplitude of this effect as-

sesses the high level of signal suppression achieved. In 3D sagittal acquisitions, distortions could occur along the phase (A/P) and partition (R/L) directions. The low level of distortions is illustrated in Fig. 5, showing the superimposition of a parallel EVI volume over anatomical images, displayed as a set of (R/L, A/P) planes. Accurate coregistration between parallel EVI data and anatomical data is observed, as required for correct localization of cerebral activations.

SNR maps were computed after filtering out cardiac and respiratory artifacts. Before filtering, signal values
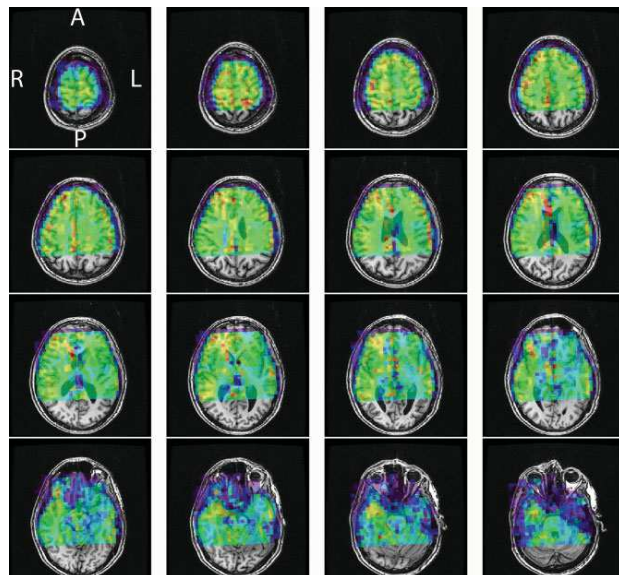
Figure 5. Superimposition of the EVI volume presented in Fig. 4 over anatomical data. This superimposition is presented as a set of (R/L, A/P) planes, since in 3D sagittal acquisitions susceptibility-induced distortions could occur along the (R/L) and (A/P) directions.
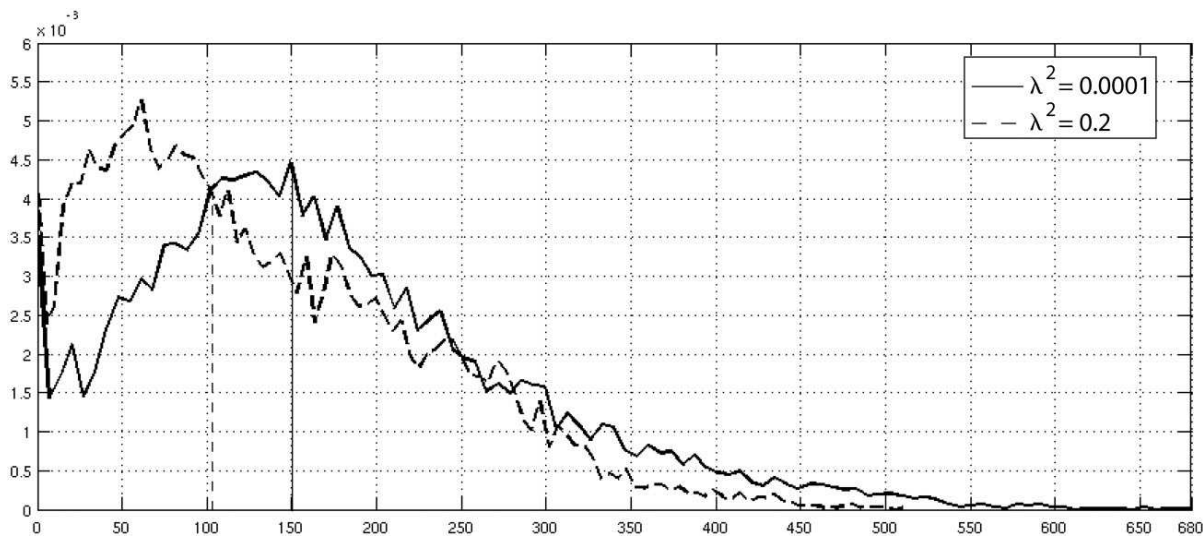
Figure 6. Distributions of temporal SNR values in the voxels of the volume of interest, for subject 1, for two different values of the regularization parameter, $\lambda^2 = 10^{-4}$ (dotted line) and $\lambda^2 = 10^{-2}$ (full line).

were not normally distributed in numerous voxels. Nevertheless, the application of the band-rejection filter made such distributions nearly Gaussian so that SNR could be calculated using the chosen definition. Figure 6 presents the distributions of SNR values in the whole volume, for the smallest and highest regularization parameters (λ^2) applied. These gamma-shaped histograms illustrate the global SNR improvement brought by regularization. Median SNR values, computed from the whole brain, are reported in Table 1 as a function of λ^2. In agreement with (20,21), regularization increases the temporal stability of parallel EVI data. Moreover, SNR values are very similar for the two subjects and relatively high in spatially smoothed time series.

Geometry factor maps have been computed using the following expression, derived from the geometry factor definition (31), for each voxel ρ of the reconstructed volume:

$$g_\rho^{\lambda^2} = \sqrt{[(S^H S + \lambda^2 \cdot Id)^{-1}]_{\rho\rho} \cdot [(S^H S)]_{\rho\rho}}$$

where S is the matrix of the complex coil sensitivities measured in the voxels superimposed in one voxel of the undersampled volume (16). The average, median, and maximum values of the geometry factor experimentally measured are highly similar for the two subjects and decrease to moderate values as λ^2 increases, as

Table 1
Median SNR as a Function of the Regularization Factor.

λ^2	Subject 1		Subject 2	
	before smoothing	after smoothing	before smoothing	after smoothing
0,0001	45	104	51	115
0,01	51	113	56	122
0,04	60	129	63	135
0,1	69	142	69	146
0,2	77	151	75	154

illustrated in Table 2. These values indicate that parallel reconstruction is feasible in the chosen geometry and justifies the SNR improvement observed when regularization is applied. The impact of regularization on the geometry factor is discussed in more detail in Ref. (31).

Activation Detection

In the two subjects, activation in response to the first presentation of the sentences was found along the superior temporal gyri and sulci, and for all λ^2 settings. In all cases, cluster level P-values were inferior to 10^{-4} after family-wise error correction for multiple comparisons. A representative activation map is displayed in Fig. 7 (top), superimposed over anatomical data, for subject 1. In Fig. 7 the presented results were computed from parallel EVI time series reconstructed with $\lambda^2 = 0.2$. The highest SNR results are presented after having checked that no ill-localized activations appear at this λ^2 value. In response to the fourth presentation, the extent of the activating area was reduced (data not shown), as observed previously (27).

In parallel EPI fMRI acquisitions, regularization of parallel reconstruction generally increases the functional contrast-to-noise ratio (CNR), through a noise reduction larger than the contrast reduction (20). This increase in CNR mechanically results in an increased sensitivity to the blood oxygenation level-dependent (BOLD) contrast (20,21). This effect is also observed in parallel EVI. For different values of λ^2, median t-scores were computed from the 200 most activated voxels in the statistical maps corresponding to the first presentation of the sentences. As illustrated in Fig. 8a, median t-scores increase when λ^2 is increased, as foretold by the SNR increase. Nevertheless, the increase slows down at high λ^2, as observed in Ref. (21). Besides, if λ^2 is further increased, the regularization condition becomes predominant over the accuracy of the reconstruction, and the reconstructed EVI volumes display

Table 2
Average, Median and Maximum Geometry Factor as a Function of the Regularization Factor

| λ^2	Subject 1			Subject 2		
	average g	median g	max g	average g	median g	max g
0,0001	2,53	2,81	4,65	2,47	2,64	5,1
0,01	2,3	2,55	3,84	2,26	2,44	3,94
0,04	1,92	2,09	2,79	1,9	2,06	2,79
0,1	1,57	1,68	2,03	1,56	1,67	2,01
0,2	1,31	1,38	1,56	1,31	1,37	1,56

important artifacts, as illustrated in Fig. 8b. Therefore, the λ^2 value corresponding to the optimal trade-off must be found.

The high acquisition efficiency of parallel EVI acquisition is illustrated in Fig. 7 (bottom), showing HRFs of individual voxels during the four repetitions of a sentence. HRFs were obtained through selective averaging of the experiment time course (32). No temporal smoothing was applied. The BOLD contrast represents around 1% of the baseline signal and a decrease in amplitude and a speeding up of the BOLD response are typically observed with habituation as before (27).
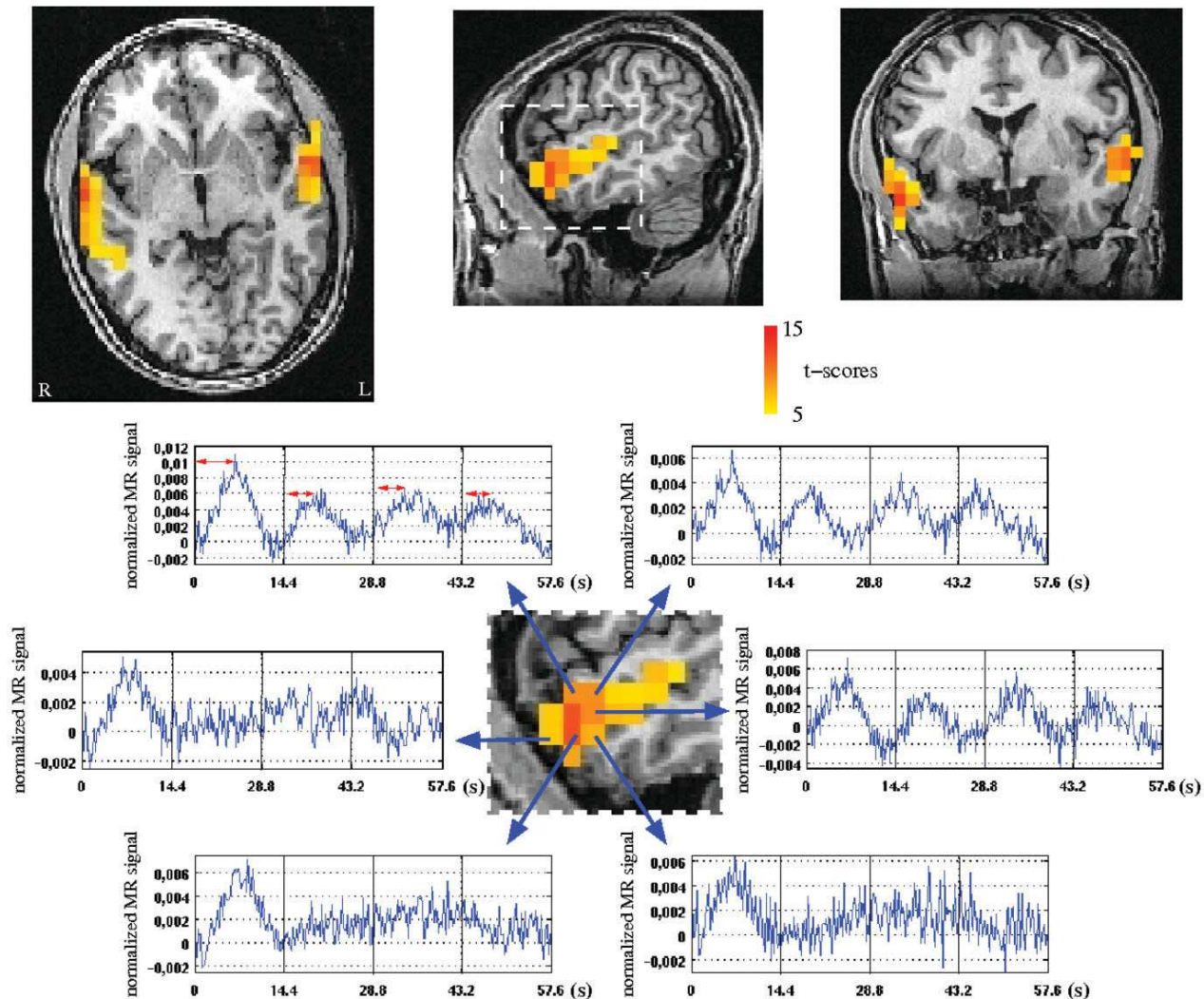
Figure 7. Top: representative activation maps illustrating the areas activated by the first repetition of the sentences, for subject 1 (thresholded at $P < 10^{-4}$, with Family-Wise Error FWE correction for multiple comparisons). Bottom: voxel-wise hemodynamic response for the four presentations of a sentence, obtained through selective averaging of the voxels time courses, after suppression of low-frequency drifts and filtering out of physiological artifacts. Vertical lines materialize the beginning of each presentation of a sentence.
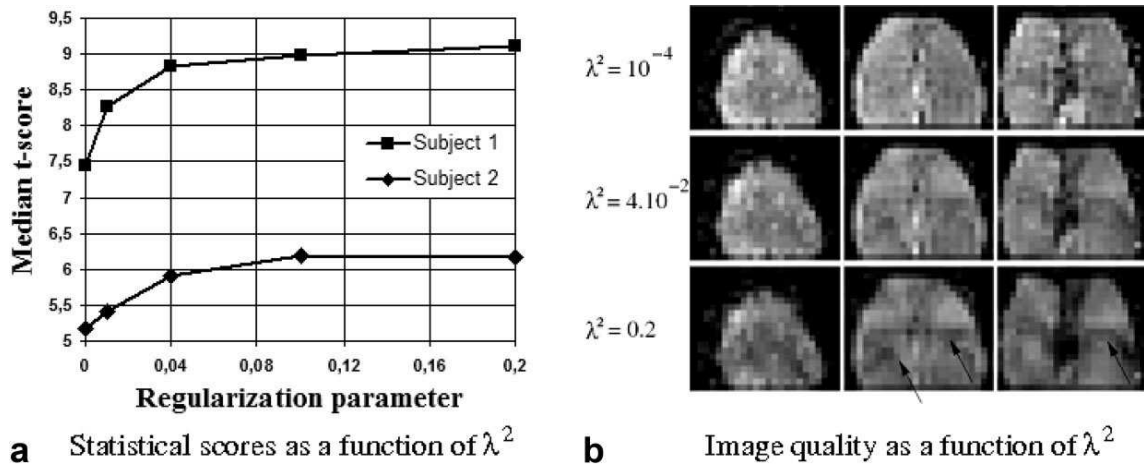
Figure 8. a: Median *t*-scores computed from the 200 most activated voxels in response to the first presentation of a sentence, as a function of the regularization parameter, for the two subjects. **b:** Three slices from a parallel EVI brain volume reconstructed with increasing values of the regularization parameter.

Moreover, different habituation behaviors can be distinguished among the voxels eliciting activation in response to the first sentence presentation.

DISCUSSION

To our knowledge, this study presents the first images and fMRI data acquired using parallel EVI. Pioneer studies have demonstrated the high temporal resolution allowed by EVI acquisitions and detection of cerebral activations at relatively high spatial resolution ($3 \times 3 \times 1.5$ mm^3 (6)). Nevertheless, those studies were limited to anisotropic volumes ($192 \times 96 \times 12$ mm^3 (6)) due to the important ETD required by EVI. In the present work, parallel imaging allows the reduction of ETD by a factor of 4, which permits covering a larger, nearly isotropic volume of the brain ($120 \times 120 \times 144$ mm^3), with a low level of susceptibility-induced distortions and signal losses. Although parallel EVI is currently limited in spatial resolution ($6 \times 6 \times 6$ mm^3), it offers a new compromise between FOV and spatial and temporal resolution. Moreover, faster gradient coils and optimized receiver coils with a higher number of acquisition channels will allow improving spatial resolution.

Several methodological difficulties were overcome to perform parallel EVI acquisition of localized volumes. First, a good adequacy between parallel data and sensitivity maps is required for SENSE reconstruction, and is not easily obtained in 3D single-shot imaging. This was achieved through minimization of the ETD in order to reduce distortions and signal losses. Therefore, 2D gradient echo images could be used as sensitivity maps. Second, combining outer volume suppression and parallel imaging requires a very high quality of external signal suppression. Otherwise, a wraparound artifact appears in the reconstructed volume. This efficient signal suppression is allowed by Shinnar–Le Roux designed quadratic phase RF pulses, repeated twice, and combined with spoiler gradients. Third, all imaging parameters were optimized before fMRI acquisitions in order to find the optimal trade-off between short ETD,

acceptable spatial resolution and brain coverage, and adequate SNR. For instance, we studied the impact of the regularization parameter, λ^2, on functional sensitivity and demonstrated that a trade-off must be found between high SNR and reconstruction accuracy. So far, λ^2 has been empirically set for the whole volume. To avoid this supervised operation and improve the efficiency of the regularization, automatic tuning of λ^2 for each voxel in the folded volume (and each subject) should be considered and an optimal criterion based on functional CNR and spatial reconstruction accuracy should be defined. The tuning could be performed either using the L-curve technique, as before (31), or more efficiently, in the maximum likelihood framework, provided that the penalization remains quadratic (Gaussian likelihood).

After this demonstration of feasibility, several methodological improvements could be made, especially regarding the acquisition strategy and the reconstruction algorithm. Actually, parallel EVI would benefit from a reconstruction algorithm in the *k*-space domain, like GRAPPA (33), since it would suppress the SENSE-ghost artifact (34). Besides, TGRAPPA (35) takes advantage of the repetitive scanning by temporal shifting the acquired *k*-space lines in an interleaved manner, which suppress the need for supplementary data. Applying TGRAPPA to parallel EVI would not only reduce the total acquisition time, but also improve image quality by providing sensitivity data with the same geometric distortions as the parallel data and reducing the sensitivity to subject's movement.

In fMRI acquisitions the main requirement is a high temporal SNR (24). In EPI acquisitions, some brain areas always present a low SNR due to distortions and signal losses. Parallel imaging allows reducing the ETD and study these areas with a better SNR (36). Nevertheless, parallel imaging introduces another source of spatial variability of the SNR, because of the spatially inhomogeneous noise in reconstructed images. Thus, spatial maps of the temporal SNR must be considered while interpreting parallel fMRI data and the same dif-

ficulty occurs with parallel EVI. Information about the spatial variations of the SNR is also provided by geometry factor maps. Nevertheless, geometry factors are not reliable estimators of the SNR loss associated with parallel acquisition since they are computed from the sensitivity maps only. For instance, when outer volume suppression is applied, as in the present study, the potential noise magnification due to the "SENSE ghost" is not taken into account. Besides, especially in single-shot acquisitions, the impact of the decrease in TE and ETD allowed by parallel imaging is not considered, and the SNR loss is generally smaller than predicted by the geometry factor (19). Thus, reliable quantitative assessment of the temporal stability is only found in SNR maps.

Increasing the scanning rate and offering 3D single-shot acquisition provides numerous advantages for fMRI studies and could improve the understanding of brain dynamics (37). Indeed, the high acquisition efficiency of parallel EVI allows the extraction of high temporal resolution single-voxel HRF using selective averaging. After this minimal postprocessing, spatial information about the dynamics of the brain response could be extracted and mapped. For example, the large number of samples acquired during the first seconds of the HRF could lead to precise estimation of its initial slope, allowing comparisons of response velocity between different experimental conditions, or voxels. As illustrated in this study, habituation effects, and more generally nonstationary effects, could benefit from parallel EVI, which combines a very high scanning rate and a large coverage of the brain. Moreover, trial-by-trial event-related responses can often be distinguished in activated voxels time courses, which could be interesting for performance-related fMRI studies. High scanning rates are also required in "rest" studies of functional connectivity (38), to avoid aliasing of physiological artifacts over the low frequencies of interest. Finally, the likely reduction of vascular artifacts, which should be investigated in the future, could potentially enhance the specificity of fMRI analysis.

The analysis of parallel EVI data requires further adaptation of activation detection and HRF estimation tools. For instance, the efficiency of the movement correction algorithms should be evaluated in parallel EVI since the mask could hinder the estimation of movement parameters. Moreover, SNR could be further improved by a more efficient correction of the physiological artifacts. Besides, at such a high scanning rate, serial correlation of the noise becomes important and the first-order autoregression model used in SPM2 may not correct efficiently for it (39). Finally, these data will also benefit from more sophisticated estimation tools, incorporating physiological and regional a priori information (2) or performing joint detection and estimation of cerebral responses (40).

In conclusion, parallel EVI was implemented and successfully applied to the detection of brain activations. Combining FOV reduction and 2D parallel acceleration with a reduction factor of 4, 3D single-shot acquisition of a nearly isotropic brain volume could be performed at 1.5T, using a very high scanning rate. Moreover, distortions and signal losses were kept low

since the ETD was similar to the ones used in EPI acquisitions. Tikhonov regularization of parallel reconstruction was applied to improve activation detection and its impact on the sensitivity to the BOLD contrast was studied. Finally, detection of brain activations was demonstrated in a slow event-related auditory paradigm and high temporal resolution, single voxel, hemodynamic response functions were recovered. Combined with advanced estimation tools, parallel EVI should prove useful to measure neural activity timing differences or study the nonlinearities and nonstationarities of the BOLD response.

REFERENCES

1. Rosen B, Buckner R, Dale A. Event-related functional MRI: past, present, and future. Proc Natl Acad Sci U S A 1998;95:773–780.
2. Ciuciu P, Poline JB, Marrelec G, Idier J, Pallier C, Benali H. Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment. IEEE Trans Med Imaging 2003;22:1235–1251.
3. Wager T, Vazquez A, Hernandez L, Noll D. Accounting for nonlinear BOLD effects in fMRI: parameter estimates and a model for prediction in rapid event-related studies. Neuroimage 2005;25:206–218.
4. Mansfield P, Coxon R, Hykin J. Echo-volumar imaging (EVI) of the brain at 3,0 T: first normal volunteer and functional imaging results. J Comput Assist Tomogr 1995;19:847–852.
5. Yang Y, Mattay V, Weinberger D, Frank J, Duyn J. Localized echo-volume imaging methods for functional MRI. J Magn Reson Imaging 1997;7:371–375.
6. van der Zwaag W, Francis S, Bowtell R. Improved echo volumar imaging (EVI) for functional MRI. Magn Reson Med 2006;56:1320–1327.
7. Frahm J, Merboldt KD, Hanicke W, Kleinschmidt A, Boecker H. Brain or vein — oxygenation or flow? On signal physiology in functional MRI of human brain activation. NMR Biomed 1994;7:45–53.
8. Lu H, Golay X, van Zijl P. Intervoxel heterogeneity of event-related functional magnetic resonance imaging responses as a function of T1 weighting. Neuroimage 2002;17:943–955.
9. Weiger M, Pruessmann M, Boesiger P. 2D SENSE for faster 3D MRI. MAGMA 2002;14:10–19.
10. Mansfield P. Multi-planar image formation using NMR spin echoes. J Phys C Solid State Phys 1977;10:L55–L58.
11. Meyer C, Pauly J, Macovski A. Simultaneous spatial and spectral selective excitation. Magn Reson Med 1990;15:287–304.
12. Le Roux P, Gilles R, McKinnon G, Carlier P. Optimized outer volume suppression for single-shot fast spin-echo cardiac imaging. J Magn Reson Imaging 1998;8:1022–1032.
13. Bruder H, Fischer H, Reinfelder H, Schmitt F. Image reconstruction for echo planar imaging with nonequidistant k-space sampling. Magn Reson Med 1992;23:311–323.
14. Ahn C, Cho Z. A new phase correction method in NMR imaging based on autocorrelation and histogram analysis. IEEE Trans Med Imaging 1987;6:32–36.
15. de Zwart J, van Gelderen P, Kellman P, Duyn J. Reduction of gradient acoustic noise in MRI using SENSE-EPI. Neuroimage 2002;16:1151–1155.
16. Pruessmann K, Weiger M, Scheidegger M, Boesiger P. SENSE: sensitivity encoding for fast MRI. Magn Reson Med 1999;42:952–962.
17. Dydak U, Weiger M, Pruessmann KP, Meier D, Boesiger P. Sensitivity-encoded spectroscopic imaging. Magn Reson Med 2001;46:713–722.
18. Zhao X, Prost RW, Li Z, Li S-J. Reduction of artifacts by optimization of the sensitivity map in sensitivity encoded spectroscopic imaging. Magn Reson Med 2005;53:30–34.

19. de Zwart J, van Gelderen P, Kellman P, Duyn J. Application of sensitivity-encoded echoplanar imaging for blood oxygen level-dependent functional brain imaging. Magn Reson Med 2002;48: 1011–1020.

20. Lin F, Huang TY, Chen NK, et al. Functional MRI using regularized parallel imaging acquisition. Magn Reson Med 2005;54:343–353.

21. Ribés A, Poupon C, Rabrait C, Le Bihan D, Lethimonnier F. Tikhonov regularization optimisation for PreLearn: effects on the detection of activation in functional MRI. In: Proc 15th Annual Meeting ISMRM, Berlin; 2007 (abstract 1757).

22. Tikhonov AN, Arsenin VA. Solution of ill-posed problems. New York: Winston-Wiley; 1977:258.

23. King K. SENSE image quality improvement using matrix regularization. In: Proc 9th Annual Meeting ISMRM, Glasgow; 2001 (abstract 1771).

24. Parrish T, Gitelman D, Labar K, Mesulam MM. Impact of signal-to-noise on functional MRI. Magn Reson Med 2000;44:925–932.

25. Goerke U, Möller H, Norris D, Schwarzbauer C. A comparison of signal instability in 2D and 3D EPI resting-state fMRI. NMR Biomed 2005;18:534–542.

26. Dietrich O, Reeder S, Reseir SM, Schoenberg S. Influence of parallel imaging and other reconstruction techniques on the measurements of signal to noise ratios. In: Proc 13th Annual Meeting ISMRM, Miami; 2005 (abstract 158).

27. Dehaene-Lambertz G, Dehaene S, Anton JL, et al. Functional segregation of cortical language areas by sentence repetition. Hum Brain Mapp 2006;27:360–371.

28. Holmes A, Josephs O, Buchel C, Friston K. Statistical modelling of low-frequency confounds in fMRI. In: Proc 3rd Int Conf Funct Mapp Hum Brain, Copenhagen; 1997:S480.

29. Pruessmann KP, Weiger M, Boesiger P. Sensitivity encoded cardiac MRI. J Cardiovasc Magn Reson 2001;3:1–9.

30. Goldfarb J. The SENSE ghost: field-of-view restrictions for SENSE imaging. J Magn Reson Imaging 2004;20:1046–1051.

31. Lin F, Kwong K, Belliveau J, Wald L. Parallel imaging reconstruction using automatic regularization. Magn Reson Med 2004;51: 559–567.

32. Dale A, Buckner R. Selective averaging of rapidly presented individual trials using fMRI. Hum Brain Mapp 1997;5:329–340.

33. Griswold MA, Jakob PM, Heidemann RM, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med 2002;47:1202–1210.

34. Griswold MA, Kannengiesser S, Heidemann RM, Wang J, Jakob PM. Field-of-view limitations in parallel imaging. Magn Reson Med 2004;52:1118–1126.

35. Breuer FA, Kellman P, Griswold MA, Jakob PM. Dynamic autocalibrated parallel imaging using temporal GRAPPA (TGRAPPA). Magn Reson Med 2005;53:981–985.

36. Bellgowan PSF, Bandettini PA, van Gelderen P, Martin A, Bodurka J. Improved BOLD detection in the medial temporal region using parallel imaging and voxel volume reduction. Neuroimage 2006;29: 1244–1251.

37. Buckner R. The hemodynamic inverse problem: making inferences about neural activity from measured MRI signals. Proc Natl Acad Sci U S A 2003;100:2177–2179.

38. Cordes D, Haughton V, Arfanakis K, et al. Frequencies contribution to functional connectivity in the cerebral cortex in "resting state" data. Am J Neuroradiol 2001;22:1326–1333.

39. Constable R, Spencer D. Repetition time in echo planar functional MRI. Magn Reson Med 2001;46:748–755.

40. Makni S, Ciuciu P, Idier J, Poline JB. Joint detection-estimation of brain activity in functional MRI: a multichannel deconvolution solution. IEEE Trans Signal Process 2005;53:3488–3502.