HAL
open science

Mixing and fluid dynamics under location uncertainty

Valentin Resseguier

To cite this version:

Valentin Resseguier. Mixing and fluid dynamics under location uncertainty. Fluid mechanics [physics.class-ph]. Université de Rennes, 2017. English. NNT : 2017REN1S004 . tel-01507292v2

HAL Id: tel-01507292

https://theses.hal.science/tel-01507292v2

Submitted on 10 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANNÉE 2016

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Bretagne Loire

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Mathématiques et applications

Ecole doctorale Matisse

présentée par

Valentin Resseguier

préparée à Inria - Centre Inria Rennes | Bretagne Atlantique
et à Ifremer - Centre de Bretagne
(Université de Rennes 1)

Intitulé de la thèse: Mixing and fluid dynamics under location uncertainty

Thèse soutenue à Brest le 10 janvier 2017

devant le jury composé de :

Georg GOTTWALD
Professor, University of Sydney / Rapporteur

Themistoklis P. SAPSIS
Associate professor, MIT / Rapporteur

Bernard DELYON
Professeur, Université de Rennes 1 / Président du jury

Bertrand CHAPRON
Directeur de recherche, Ifremer / Examinateur

Christian FRANZKE
Senior research associate, University of Hamburg / Examinateur

Guillaume LAPEYRE
Chargé de recherche, CNRS / Examinateur

Anne-Marie TREGUIER
Directrice de recherche, CNRS / Examinateur

Etienne MÉMIN
Directeur de recherche, Inria / Directeur de thèse

Remerciements

Je tiens à remercier tout particulièrement mes directeurs de thèse: Etienne Mémin et Bertrand Chapron. Ils m'ont apporté énormément de connaissances scientifiques, toujours dans une atmosphère très conviviale. Avec ses modèles sous incertitude de position, Etienne m'a permis de travailler sur un domaine nouveau où tout ou presque était à découvrir. Il m'a aussi montré que, dans le domaine du traitement de signal, l'informatique ne se bornait pas à l'application de développements théoriques, mais pouvait aussi être un outil de tâtonnement pour guider une recherche scientifique. Durant ma thèse, Etienne s'est beaucoup investi à mes cotés notamment pour l'écriture d'articles. Enfin, il m'a permis de voyager énormément grâce à de multiples conférences. Bertrand m'a permis de retrouver une intuition physique et de la développer. Il a apporté le sens physique et la pédagogie à notre discours pour qu'il puisse trouver un public. Tout ce travail n'aurait pas été possible s'il ne nous avait pas accordés sa confiance en finançant, avec son laboratoire, d'abord la moitié de ma thèse, puis 3 mois supplémentaires à la fin de celle-ci. Il m'a laissé libre de me m'attaquer à de long développements théoriques. Il n'a pas hésité à contacter plusieurs chercheurs réputés et m'a poussé à en contacter d'autres.

Je remercie aussi mes collègues de l'équipe Fluminance et de l'ancien LOS. En particulier, je tiens à remercier Pierre Dérian, un ami avant d'être un collègue, avec qui j'ai travaillé et travaille encore de façon très efficace. Je me dois aussi de remercier Johan Carlier, Pranav Chandramouli et Dominique Heitz pour les sorties de modèles fournies et leurs analyses physiques.

Les sorties de modèles mises à disposition par Jonathan Gula et Jeroen Molemoker ont été, elles aussi, indispensables à cette thèse. Mon travail s'est aussi nourri des discussions que j'ai pu avoir avec ces derniers ainsi qu'avec Georg Gottawld, Darryl Holm, Guillaume Lapeyre, Aurélien Ponte, Guillaume Roulet et Themis Sapsis. Je remercie aussi le jury et les rapporteurs pour avoir accepté ce rôle et pour leur travail qui a permis d'améliorer la thèse et les publications associées.

Je remercie le projet ESA DUEGlobCurrent et les Labex CominLabs, Lebesgue et Mer avec le projet SEACS pour les financements.

Enfin, je remercie Christine, Pierre et Pranav pour la relecture de l'anglais ainsi que toutes les personnes qui m'ont soutenu pendant cette thèse, notamment Bénédicte, Maréva, mes amis et mes parents.

Contents

Remerciements **iii**

Résumé en français **ix**

Notations **xv**

Introduction **xix**

1 Stochastic processes and stochastic calculus **1**

1.1 Gaussian processes and Brownian motions . . . 1
1.1.1 Definition and notations . . . 1
1.1.2 Finite dimensional Brownian motion . . . 1
1.1.3 Infinite dimensional Brownian motion: $\mathbb{I}_d$-cylindrical Wiener process . . . 2
1.2 Basic notions of stochastic calculus . . . 3
1.2.1 Finite variation process . . . 3
1.2.2 Martingale . . . 3
1.2.3 Continuous semimartingale, Itō and diffusion processes . . . 3
1.2.4 Quadratic variation and quadratic cross-variation . . . 3
1.2.5 Integral notations . . . 4
1.3 Important theorems . . . 4
1.3.1 Itō-Wentzell Formula . . . 4
1.3.2 Itō Isometry . . . 4
1.3.3 Girsanov theorem . . . 5
1.3.4 Kolmogorov's backward equation . . . 5
1.3.5 Fokker-Planck equation . . . 5
1.4 Example of classical processes . . . 6
1.4.1 Ornstein-Uhlenbeck process . . . 6
1.4.2 Geometric Brownian motion . . . 6
1.5 Estimation methods for the noise . . . 7
1.5.1 Constant diffusion coefficient . . . 8
1.5.2 The diffusion coefficient is time varying . . . 10
1.6 Stratonovich versus Itō . . . 10
1.6.1 Definitions . . . 11
1.6.2 Links between the two notations . . . 11
1.6.3 Advantages of each setups . . . 12
1.7 Generalization to Markovian processes . . . 12
1.7.1 Analog of the Fokker-Planck equation: one-time density . . . 13
1.7.2 Multi-time density . . . 13
1.7.3 Analog of the mean of Itō formula . . . 13
1.7.4 Multi-point density . . . 14

2 Geophysical flows under location uncertainty, Part I **15**
2.1 Introduction . . . 15
2.2 Transport under location uncertainty . . . 17
2.2.1 A 2-scale random advection-diffusion description . . . 17
2.2.2 Uncertainty formalism . . . 18
2.2.3 Material derivative . . . 19
2.2.4 Scalar advection . . . 19
2.2.5 Transport of extensive properties . . . 22
2.2.6 Summary . . . 23
2.3 Stochastic versions of geophysical flow models . . . 23
2.3.1 Mass conservation . . . 23
2.3.2 Active scalar conservation law . . . 24
2.3.3 Conservation of momentum . . . 24
2.3.4 Atmosphere and Ocean dynamics approximations . . . 25
2.3.5 Summary . . . 27
2.3.6 Guidelines for the derivation of models under location uncertainty . . . 28
2.4 Conclusion . . . 29
Appendices . . . 30
2.A Quadratic variation . . . 30
2.B Link between the material derivative D_t and the operator $\mathbb{D}_t$ . . . 30
2.C The evolution of the variance of a passive tracer . . . 31
2.D Stochastic extension of the Reynolds transport theorem . . . 32
2.E Stochastic Navier-Stokes model . . . 33

3 Geophysical flows under location uncertainty, Part II **35**
3.1 Introduction . . . 35
3.2 Models under location uncertainty . . . 36
3.3 Mesoscales under moderate uncertainty . . . 37
3.3.1 Specific scaling assumptions . . . 37
3.3.2 Stratified Quasi-Geostrophic model under moderate uncertainty . . . 39
3.3.3 Surface Quasi-Geostrophic model under moderate uncertainty . . . 41
3.3.4 Summary . . . 41
3.4 Numerical results . . . 41
3.4.1 Test flow . . . 41
3.4.2 Simulation of the random velocity . . . 42
3.4.3 Resolution gain on a single simulation . . . 45
3.4.4 Ensemble forecasts . . . 45
3.5 Conclusion . . . 51
Appendices . . . 57
3.A Non-dimensional Boussinesq equations . . . 57
3.B QG model under moderate uncertainty . . . 58

4 Geophysical flows under location uncertainty, Part III **65**
4.1 Introduction . . . 65
4.2 Models under location uncertainty . . . 66
4.3 Mesoscale flows under strong uncertainty . . . 67
4.3.1 Specific scaling assumptions . . . 67
4.3.2 Stratified Quasi-Geostrophic model under strong uncertainty . . . 69
4.4 Diagnostic under strong uncertainty . . . 71
4.5 Conclusion . . . 73
Appendices . . . 77
4.A Modified geostrophic balance . . . 77
4.B Non-dimensional Boussinesq equations . . . 77
4.C QG model under strong uncertainty . . . 79

5 Bifurcations and location uncertainty **81**
5.1 Introduction . . . 81
5.2 Transport under location uncertainty . . . 82
5.2.1 Informal description . . . 82
5.2.2 Uncertainty formalism . . . 83
5.2.3 Scalar advection . . . 84
5.2.4 SQG model under moderate uncertainty . . . 84
5.3 Chaotic test flow and resolution issues . . . 85
5.4 Stochastic analysis . . . 88
5.5 Conclusion . . . 95
Appendices . . . 97
5.A Stochastic material derivative . . . 97
5.B Simulation of the random velocity . . . 98

6 Reduced flow models **99**
6.1 Introduction . . . 99
6.2 The proposed stochastic model . . . 101
6.2.1 Stochastic representation of the Reynolds-transport theorem . . . 102
6.2.2 Stochastic Navier-Stokes model . . . 103
6.3 Classical model reduction using POD . . . 104
6.4 Stochastic POD . . . 105
6.4.1 Model . . . 105
6.4.2 Choice of the time step . . . 106
6.4.3 Estimation of the uncertainty variance tensor . . . 106
6.5 Numerical results . . . 109
6.5.1 Characteristics of the data . . . 109
6.5.2 Reconstruction of *chronos* . . . 110
6.6 Conclusion . . . 114
Appendices . . . 115
6.A Basic notions of stochastic calculus . . . 115
6.B Equivalence of formulation for the stochastic transport and Navier-Stokes theorem 117
6.B.1 Stochastic Reynolds transport model . . . 117
6.B.2 Stochastic Navier-Stokes model . . . 117
6.C Dissipative effect of sub-grid tensor τ . . . 117

7 Diffusion modes for POD models and flow analysis **119**
7.1 Introduction . . . 119
7.2 Dynamics stochastic modelling . . . 120
7.2.1 Stochastic conservation equations . . . 121
7.2.2 Navier-Stokes equations associated with a stochastic representation of the small-scales . . . 122
7.3 Reduced order models . . . 123
7.3.1 POD model reduction . . . 123
7.3.2 Reduced order modelling associated with the stochastic representation . . . 124
7.3.3 Time scale characterisation . . . 124
7.3.4 Estimation of the small-scale variance tensor . . . 125
7.4 Flow configuration and numerical simulations . . . 126
7.5 Diffusion modes results . . . 127
7.5.1 Estimation and decomposition of the turbulent velocity components . . . 127
7.5.2 Small-scale energy density, stationarity and anisotropy . . . 129
7.5.3 Chronos reconstruction . . . 137
7.6 Conclusion . . . 140

8 Transport of fluid velocity under location uncertainty 143
8.1 Transport under location uncertainty of the velocity . . . 143
8.1.1 Structure of the model . . . 143
8.1.2 Kinetic energy budget . . . 144
8.1.3 Incompressible Navier-Stokes under location uncertainty with and without noise . . . 145
8.2 Galerkin projection . . . 145
8.2.1 A finite-dimensional system with correlated additive and multiplicative noise 145
8.2.2 Energy leak . . . 146
8.2.3 Proper Orthogonal Decomposition . . . 147
8.3 Estimations of subgrid terms . . . 148
8.3.1 Estimation formulas . . . 148
8.3.2 Modal time step . . . 149
8.4 Ensemble forecasting . . . 149
8.5 Conclusion . . . 156
Appendices . . . 156
8.A Navier-Stokes model under location uncertainty . . . 156
8.B Energy dissipation . . . 157
8.C Estimation formulas . . . 159

9 Smooth flow, tracer gradients and tracer spectra 161
9.1 Introduction . . . 161
9.2 Time evolution for the tracer gradient norm . . . 163
9.2.1 Exact flow properties . . . 163
9.2.2 Decorrelation approximations . . . 164
9.2.3 Link with the mixing criterion of Mezić et al. (2010) . . . 165
9.2.4 Time dependance . . . 166
9.2.5 Numerical illustrations . . . 170
9.3 Tracer spectral tail . . . 179
9.3.1 Moment conservation . . . 179
9.3.2 Smooth scalar approximation . . . 180
9.3.3 Tail change . . . 181
9.3.4 Self-similar approximation . . . 181
9.4 Estimation of eddy diffusivity . . . 184
Appendices . . . 188
9.A Omnidirectional spectra of tracers for smooth flow . . . 188

10 Additional results 189
10.1 Models under location uncertainty with Stratonovich notations . . . 189
10.1.1 Flow . . . 189
10.1.2 Transport . . . 189
10.1.3 Modified drift . . . 190
10.2 Comparison with Holm (2015) . . . 190
10.3 Review of the derivation of the models under location uncertainty . . . 192
10.3.1 Mass conservation . . . 192
10.3.2 Active scalar conservation law . . . 192
10.3.3 Navier-Stokes model under location uncertainty . . . 193
10.3.4 Boussinesq equations . . . 194
10.3.5 Non-dimentionalized Boussinesq equations . . . 195
10.3.6 Quasi-Geostrophic model under Moderate Uncertainty (QG_{MU}) . . . 197
10.3.7 Surface Quasi-Geostrophic model under Moderate Uncertainty (SQG_{MU}) . 201
10.3.8 QG and SQG models under Strong Uncertainty (SQG_{SU}) . . . 201

Conclusion 205

Résumé en français

Cette thèse concerne le développement, l'extension et l'application d'une formulation stochastique des équations de la mécanique des fluides introduite par Mémin (2014). La vitesse petite échelle, non-résolue, est ici modélisée au moyen d'un champ aléatoire décorrélé en temps. Cela change l'expression de la dérivée particulaire et donc les équations de la mécanique des fluides. Les modèles qui en découlent sont dénommés modèles sous incertitude de position. La thèse s'articulent autour de l'étude successive de modèles réduits, de versions stochastiques du transport et de l'advection à temps long d'un champ de traceur par une vitesse mal résolue.

Le chapitre 1 est une introduction qui rassemble différents résultats théoriques. Le chapitre 6 est un article publié. Les chapitres 2, 3, 4, 5 et 7 correspondent à des manuscripts qui ont été soumis. Les chapitres 8 et 9 sont des manuscripts en préparation. Le chapitre 10 présente quelques résultats complémentaires.

Résumé du chapitre 1: Processus stochastiques et calcul stochastique

Dans le premier chapitre, nous rappelons plusieurs notions relatives aux processus aléatoires et au calcul stochastique. Après avoir présenté les processus gaussiens et les mouvements browniens en dimensions finie et infinie, nous donnons brièvement les définitions de base et les principales propriétés du calcul stochastique: les processus à variations finies, les martingales, les semimartingales et les variations quadratiques. Puis, plusieurs théorèmes sont détaillés: formule d'Itō-Wentzell, théorème de Girsanov ainsi que les équations de Kolmogorov et Fokker-Planck. Ces résultats fournissent des outils très puissants pour gérer des fonctions aléatoires. En particulier, la formule d'Itō-Wentzell exprime la dérivée temporelle de la composée de deux fonctions aléatoires. La modélisation sous incertitude de position développée dans les chapitres 2, 3, 4, 5, 6, 7 et 8 s'appuie fortement sur ce théorème. Nous présentons également les versions multidimensionnelles de deux processus célèbres: le processus d'Ornstein-Uhlenbeck (OU) et le mouvement brownien géométrique. Dans ce cadre de dimension finie, nous donnons un aperçu de la structure algébrique des lois d'évolution impliquées dans la modélisation sous incertitude de position. Après cela, nous examinons quelques méthodes pour estimer les coefficients de diffusion. Ce point sera utile pour les chapitres 6, 7 et 8. Avant de conclure, les avantages et les inconvénients des notations de Stratonovich et d'Itō sont discutés. Enfin, nous ouvrons la discussion avec des outils plus généraux, issue de la physique statistique. Avec ces outils, les opérateurs différentiels impliqués dans les théorèmes précédemment cités ne sont plus limités au second ordre.

Résumé du chapitre 2: Ecoulements géophysiques sous incertitude de position, Partie I: Transport aléatoire et modèles généraux

La première partie développe le cadre général de la modélisation sous incertitude de position introduite par Mémin (2014). Ce principe repose sur l'hypothèse d'une vitesse, vivant à petite échelle, aléatoire et non corrélée dans le temps. Cette hypothèse provient de l'observation que les pas de temps utilisés dans les simulations et les mesures des écoulements géophysiques, sont beaucoup plus grands que les plus petits temps hydrodynamiques. Cette hypothèse modifie l'expression de la dérivée particulaire introduisant une correction d'advection, un bruit multiplicatif et une diffusion inhomogène et anisotrope. Pour des vitesses non divergentes, cette dérivée particulaire stochastique a la remarquable propriété de conserver l'énergie de chaque réalisation. Les effets équilibrés

du bruit et de la diffusion permettent à notre modèle de mieux prendre en compte les échanges énergétiques entre les différentes composantes du traceur, en comparaison à une dérivée particulaire déterministe. En particulier, une partie de l'énergie de la moyenne (d'ensemble) du traceur est transférée de façon continu à la composante aléatoire du traceur. Cet échange d'énergie aléatoire constitue une propriété bien adpatée à la quantification d'incertitude. Enfin, la loi d'évolution du Jacobien et les similitudes avec le modèle de Kraichnan sont présentées.

Dans la seconde partie, les conservations générales de la mécanique des fluides (masse, énergie interne, salinité et quantité de mouvement) sont exprimées dans le cadre de l'incertitude de position. Nous pouvons en déduire un modèle de Boussinesq stochastique. En présence d'un gradient de background stationnaire (*e.g.* stratification), les forçages stochastiques de quantités transportées de façon aléatoire sont des combinaisons de bruits additifs et multiplicatifs corrélés. Le bruit additif transfère l'énergie du background à l'anomalie et augmente sa variance. Les effets du bruit additif sur le transport de la flottabilité sont étudiés à l'aide d'un modèle jouet d'oscillations de flottabilité. En supposant que la vitesse à grande échelle est lisse dans le temps, nous pouvons séparer l'équation de quantité de mouvement en deux équations: une pour les termes de grande échelle et une pour les termes de petite échelle. Ce chapitre résume également les principales étapes de la dérivation de modèles sous incertitude de position ainsi que les paramétrages possibles.

Résumé du chapitre 3: Ecoulements géophysiques sous incertitude de position, Partie II: Quasi-géostrophie et dispersion efficace d'ensembles

Dans la suite du chapitre 2, des modèles simplifiés de mésoéchelle sont dérivés. En plus des hypothèses de forte stratification et de forte rotation, les ordres de grandeur des termes de tenseurs sous-maille doivent être évalués. L'ordre de grandeur considéré dans ce chapitre conduit à un modèle quasi-géostrophique (QG) appelé sous incertitude modérée. La Vorticité Potentielle (PV) est conservée à trois termes sources près. Le premier est dû à des inhomogénéités de la correction d'advection et de la diffusion turbulente de la vitesse; Le second exprime les interactions entre la fréquence de Coriolis et les tourbillons sous-maille inhomogènes; Le dernier est le résultat de l'action conjointe de la vitesse résolue et de la vitesse non résolue. Comme ce dernier terme n'est pas corrélé dans le temps, il augmente l'enstrophie potentiel. En supposant une PV nul à l'intérieur du fluide, on obtient la relation classique de surface quasi-géostrophique (SQG) entre la vitesse à grande échelle et la flottabilité. En tant que tel, le transport sous incertitude de la flottabilité, à la surface, a été appelée modèle SQG sous incertitude modérée (SQG_{MU}).

Pour illustrer le potentiel de la modélisation sous incertitude de position, des simulations numériques du modèle SQG_{MU} sont effectuées avec une vitesse non-résolue homogène. Deux simulations du modèle SQG déterministe sont également décrites: une à la même résolution grossière, pour fournir un élément de comparaison, et une à une résolution plus élevée, choisie comme référence. La simulation utilisant le modèle SQG_{MU} résout mieux les structures à petite échelle que le modèle déterministe à la même résolution. Plus précisément, le modèle SQG_{MU} résout les instabilités de filaments et la création de petits tourbillons appelés "colliers de perles" qui s'en suit. Ensuite, l'équilibre entre le bruit et la diffusion prescrit par notre cadre est confirmé visuellement par des champs de flottabilité dans les espaces spatiale et spectrale. Après cela, un ensemble de simulations a été réalisé avec le modèle SQG_{MU}. L'ensemble évalue avec précision les amplitudes et les positions de ses propres erreurs à chaque pas de temps dans les espaces spatiale et spectrale. Au contraire, un ensemble généré par le modèle déterministe avec des conditions initiales aléatoires sous-estime les erreurs d'un ordre de grandeur. De plus, une analyse des moments d'ordre supérieur de l'ensemble SQG_{MU} identifie correctement les "colliers de perles" comme des événements extrêmes froids ou chauds.

Résumé du chapitre 4: Ecoulements géophysiques sous incertitude de position, Partie III: SQG et dynamique des fronts sous des conditions de forte turbulence

Ce chapitre explore un autre adimentionnement du modèle de Boussinesq aléatoire dérivée dans le chapitre 2. Comparée au chapitre 3, la dynamique sous-maille a une influence plus forte. Cependant, les autre termes présentent les même ordres de grandeurs que dans le chapitre 3. L'hypothèse

de forte rotation conduit à un équilibre géostrophique modifié impliquant une divergence de la vitesse horizontale. Puis, la forte rotation donne directement une PV nul à l'intérieur du fluide et une relation SQG modifiée. Ce modèle a été désigné par le nom de modèle SQG sous forte incertitude (SQG_{SU}).

L'équilibre géostrophique modifié fournit une relation simple entre la vorticité et la divergence de la vitesse horizontale. Ce diagnostic est testé avec succès sur des sorties de modèle réaliste à très haute résolution (Gula et al., 2015). Ce test nous permet également d'estimer un coefficient de diffusion horizontale réaliste et de vérifier *a posteriori* notre hypothèse de forte incertitude. La composante de vitesse divergent (ageostrophique) exprime la frontolyse sur les côtés chauds des fronts et la frontogénèse sur les côtés froids des fronts à la surface de l'océan.

Résumé du chapitre 5: Bifurcations et incertitude de position dans les écoulements fluides géophysiques

Ce chapitre est une suite naturelle du chapitre 3. La modélisation sous incertitude de position est encore une fois étudiée à l'aide de l'exemple du modèle SQG_{MU}, dans le cas homogène et pour un écoulement ayant une géométrie similaire. Néanmoins, l'accent est mis ici sur une bifurcation apparaissant après un temps de prédicatiblité.

Après avoir brièvement rappelé les principaux aspects de la modélisation sous incertitude de position et le modèle SQG_{MU}, une bifurcation due à une rupture de symétrie est étudiée avec le modèle SQG déterministe à haute résolution. Une légère modification des conditions initiales conduit à un changement majeur après un mois d'advection: deux scénarios sont identifiés. Avec la même condition initiale, en raison de l'inévitable tenseur sous-maille, le modèle SQG déterministe à deux résolutions différentes donne deux scénarios différents. A une résolution grossière, une approche stochastique est donc nécessaire. Deux ensembles sont simulés: un avec la loi d'évolution classique et des conditions initiales aléatoires et un avec le modèle SQG_{MU}. Pour analyser visuellement les ensembles, une décomposition de Karhunen-Loève de l'ensemble est réalisée à chaque pas de temps. Dans cette étude de cas, le premier mode de la décomposition est suffisant pour une description qualitative. Sa densité de probabilité est estimée pour les deux ensembles. Seul l'ensemble associé au modèle SQG_{MU} est convergé statistiquement et décrit correctement les deux scénarios. En outre, cet ensemble est plus proche de la référence. Pour la méthode avec des conditions initiales aléatoires, la nécessité d'un plus grand nombre de particules est liée à son comportement sous-dispersif bien connu. Ceci est dû à la fois à l'échantillonnage initial et à la mauvaise résolution de loi d'évolution. En revanche, la modélisation sous incertitude de position possède un forçage aléatoire inhomogène même dans le cas d'une vitesse non-résolue homogène; Cette propriété rend l'élargissement de l'ensemble plus efficace et nous permet d'utiliser moins de réalisations pour représenter l'ensemble.

Résumé du chapitre 6: Modèle réduit d'écoulement issue d'une representation stochastique de Navier-Stokes

Ce chapitre utilise la modélisation sous incertitude de position pour dériver un modèle d'ordre réduit (ROM) déterministe pour les écoulements fluides.

Les ROM expriment des solutions d'EDP (Equations aux Dérivées Partielles) ou d'EDPS (Equations aux Dérivées Partielles Stochastiques) sur des bases réduites. Ces bases sont issues de données et optimisent généralement un critère spécifique. Cette thèse porte sur la Proper Orthogonal Decomposition (POD). La base associée est un ensemble de fonctions de l'espace appelées modes spatiaux qui encodent un maximum de variabilité spatio-temporelle de la solution. Les coefficients de la vitesse dans cette base sont appelés modes temporels. Après une projection de Galerkin de la loi d'évolution sur la base, des termes supplémentaires sont souvent nécessaires pour modéliser l'influence des modes non résolus. Dans cette thèse, nous proposons deux ROM POD basés sur deux représentations stochastiques différentes de Navier-Stokes pour deux applications différentes: un modèle déterministe pour les simulations de type LES dans les chapitres 6 et 7 et un modèle stochastique pour la quantification d'incertitude (UQ) dans le chapitre 8. Les deux s'appuient sur la modélisation sous incertitude de position mais avec des hypothèses différentes.

Dans la première approche, la vitesse à grande échelle et donc les modes temporels résolus sont supposés être différentiables par rapport au temps. Comme dans Mémin (2014) et dans le chapitre 2, ceci conduit dans le contexte de la modélisation sous incertitude de position à une représentation grande échelle de Navier-Stokes sans bruit. Une correction d'advection et une diffusion inhomogène et anisotrope apparaissent à la fois dans l'EDP et dans le ROM associé. Elles sont paramétrées par la matrice de covariation quadratique à un point et à un temps de l'écoulement. Cette matrice inhomogène spatialement est appelée tenseur de variance. Le chapitre 6 propose deux méthodes pour l'estimer en fonction de la vitesse résiduelle. Lorsque le tenseur de variance est supposé stationnaire, une moyenne temporelle donne facilement le résultat. Pour le cas non stationnaire, le calcul stochastique permet encore d'estimer le tenseur de variance en le décomposant sur une base temporelle (Genon-Catalot et al., 1992). En utilisant les modes temporels de la POD comme base temporelle, le ROM reste autonome. De plus, nous estimons un pas de temps optimal sur les modes résolus pour accélérer la simulation du ROM et améliorer l'estimation du tenseur de variance. Enfin, le ROM reconstruit avec succès les modes temporels de deux écoulements de sillage à Reynolds 300 et 3900. Le tenseur sous-maille stabilise le système réduit, même à Reynolds élevé, et l'advection efficace corrige le décalage de fréquence.

Résumé du chapitre 7: Modélisation stochastique et modes de diffusion pour modèles POD et analyse des petites échelles d'un écoulement

Ce chapitre fait suite au précédent. Nous rappelons le principe de la modélisation sous incertitude de position, la représentation de Navier-Stokes associée pour le cas d'une vitesse grande échelle lisse, le ROM POD qui en résulte et les deux types d'estimation du tenseur de variance. La décomposition du tenseur de variance sur une base temporelle définit ce que nous appelons les modes de diffusion. Ce chapitre propose une estimation de temps caractéristiques modales au lieu d'un seul temps comme dans le chapitre précédent. Pour chaque mode temporel résolu, un temps caractéristique est associé. Comme l'estimateur du tenseur de variance dépend de ce temps caractéristique, l'équation d'évolution de chaque mode résolu repose sur un tenseur de variance différent et donc un tenseur de sous-maille différent. Cette méthode améliore fortement la reconstruction des modes temporels. Sans aucun ajustement, notre reconstruction devient meilleure, ou au moins aussi bonne, que le ROM POD, utilisant une viscosité turbulente (Rempfer and Fasel, 1994), qui est ajustée de manière optimale sur les séries temporelles des modes résolus. En outre, le chapitre montre comment analyser l'influence de la vitesse résiduelle à travers les modes de diffusion. Ainsi, ce chapitre propose de nouveaux outils d'analyse de données, complémentaires à la décomposition POD (ou aux fonctions empiriques orthogonales (EOF)) habituelle. Une telle analyse est impossible avec l'hypothèse de viscosité turbulente traditionnelle. Ces modes de diffusion dévoilent les effets de la composante de vitesse non résolue sur la composante résolue. Nous identifions les régions et les directions principale de la diffusion turbulente. Nous décrivons aussi les advections efficaces induites par cette vitesse non-résolue.

Résumé du chapitre 8: Transport conservatif sous incertitude de position de la vitesse d'un fluide et réduction de dimension

Dans une deuxième approche des ROM, nous ne supposons plus que la vitesse à grande échelle est différentiable par rapport au temps. Dès lors, la vitesse à grande échelle est transportée sous incertitude de position comme tous les autres traceurs. En plus de la correction d'advection et de la diffusion turbulente, le modèle grande échelle de Navier-Stokes possède un bruit multiplicatif. Cela permet des transferts d'énergie plus efficaces. En particulier, la conservation de l'énergie est assurée et l'UQ est de meilleure qualité. La structure algébrique du modèle est présentée. Ce modèle conduit à un ROM stochastique avec des modes temporels résolus qui sont aléatoires. En raison de la troncature des modes, une petite partie de l'énergie diffusée dans le ROM fuit vers les modes non-résolus et permet, ainsi, la dissipation visqueuse moléculaire dans les modes non-résolus. Le reste de l'énergie rétrodiffuse dans le système réduit grâce au forçage aléatoire multiplicatif. Cette analyse est valable pour toutes les projections de Galerkin sur une base orthonormée. Le résultat est donc aussi bien applicable aux ROMs qu'aux simulations numériques aléatoires de type LES. Comparativement aux chapitres précédents, des estimateurs plus complexes sont nécessaires pour

estimer les statistiques des bruits, sans recourir à l'énorme fonction de covariance spatiale de la vitesse petite échelle inhomogène. Les formules d'estimation pour POD ROM sont présentées. La méthode permet de prévoir de façon précise l'évolution des modes temporels de deux écoulements de sillage à Reynolds 300 et 3900. Les ROMs stabilisent les modes instables, empêchent la dégénérescence des modes stables et estiment la bonne quantité d'erreurs pour garder l'ensemble le plus proche possible de la référence.

Résumé du chapitre 9: Effets des flots lisses sur les gradients de traceur et sur les spectres de traceurs

La méthode d'advection Lagrangienne (Sutton et al., 1994; Desprès et al., 2011a; Berti and Lapeyre, 2014) augmente la résolution des images satellitaires de traceurs en utilisant l'advection par une vitesse lisse issue de l'altimétrie. Cette advection crée des structures mésoéchelle et sous-mésoéchelle. Néanmoins, en raison de la résolution spatio-temporelle grossière des champs impliqués, certaines de ces structures ne sont pas physiques. Ainsi, un lissage spatial supplémentaire du traceur ainsi qu'une préscription précise du temps d'advection sont nécessaires. Le but de ce chapitre est de mieux comprendre le processus de formation de ces structures et de donner une règle pour choisir le temps d'advection ainsi que la largeur du filtre.

Après un temps fini, les gradients d'un traceur lisse advecté ont évolué. Nous exprimons de façon analytique la moyenne de la norme au carré de ces gradients, en mettant en valeur les liens avec le tenseur de Cauchy-Green, les Exposants de Lyapunov a Temps Fini (FTLE) et la vitesse mésochronique (Mezić et al., 2010). L'angle relatif du gradient de traceur initial influence la rapidité de renforcement et d'affaiblissement des gradients du traceur. Cependant, lorsque le flot et le traceur initial sont faiblement corrélés, la norme quadratique moyenne des gradients du traceur ne peut qu'augmenter. De plus, le taux de croissance ne dépend pas du traceur. Ce renforcement du gradient est dû à l'action combinée de l'étirement et du repliement, aussi appelé mélange, des structures du traceur. Grâce à ce processus, ces structures cascadent vers des échelles de plus en plus petites. Cette vision du mélange s'applique à des traceurs lisses. En revanche, les simulations numériques qui sont sous-résolues en espace ou en temps, telles que les modèles sous l'incertitude de position, diluent les structures de traceur, lorsqu'elles sont plus petites que l'échelle de résolution. En plus du cas particulier des tracteurs passifs, une corrélation faible entre traceur et flot peut être également observée dans le cadre de la méthode d'advection Lagrangienne, à cause du lissage spatial du traceur. De plus, l'échelle de temps géostrophique associée aux variations du courant de surface est grande devant le temps d'évolution des gradients sous-mésoéchelles du traceur. Ainsi, pour cette application spécifique, la vitesse Eulérienne est quasi-stationnaire et crée des cisaillements localement uniformes et des cellules convectives stationnaires. Les premiers étirent le traceur le long d'une ligne droite alors que les secondes replient et enroulent le traceur autour des tourbillons. Lorsque le temps d'advection est trop long, cet enroulement fait apparaître des structures en spirales, dans le traceur, qui ne sont pas physiques. A partir de ces deux types de comportements, nous déduisons un modèle linéaire en temps pour le taux de croissance des gradients de traceurs. Ce modèle est seulement paramétré par la vitesse Eulerienne grande échelle. Après cela, nous montrons que la moyenne quadratique des gradients de traceur spécifie la longueur de corrélation de la covariance du traceur et donc la position de la queue du spectre. En contrôlant de façon analytique l'intensification des petites échelles dans l'espace de Fourier, nous pouvons spécifier la largeur du filtre de la méthode d'advection Lagrangienne, mais aussi la diffusivité turbulente dans le but de paramétrer des simulations numériques d'écoulements grande échelle comme les modèles sous incertitude de position. Des expériences numériques avec un flot jouet ainsi qu'avec des images satellites réelles illustrent chacun de nos résultats théoriques. En particulier, nous prédisons avec précision la répartition globale et locale de l'étirement pour une advection à temps fini. Ce travail met également en évidence l'effet important du repliement dans le mélange à temps fini.

Résumé du chapitre 10: Résultats supplémentaires

Ce chapitre présente d'abord les modèles sous incertitude de position avec les notations de Stratonovich au lieu de celle d'Itō. Cela permet notamment de mieux comprendre l'expression de la correction

d'advection dans ces modèles. Ensuite, nous comparons les modèles sous incertitude de position avec les travaux récents de Holm (2015).

Notations

Algebra

- Bold symbols refer to vectors or matrices,
- $\mathbb{I}_d$ is the identity matrix,
- $\mathrm{tr}(\boldsymbol{A}) = \sum_i < \boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_i >$ denotes the trace of the matrix or operator $\boldsymbol{A}$, where $(\boldsymbol{e}_i)_i$ denotes a complete orthonormal basis of the space,
- $\|\boldsymbol{A}\|^2 = \|\boldsymbol{A}\|_2^2 = \mathrm{tr}\left(\boldsymbol{A}\boldsymbol{A}^T\right) = \sum_{ij} A_{ij}^2$ is the square of the Frobenius norm of $\boldsymbol{A}$,
- $\|\boldsymbol{A}\|_{\boldsymbol{S}}^2 = \mathrm{tr}\left(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^T\right)$ is the square of the Frobenius norm, defined by the symmetric positive-definite matrix $\boldsymbol{S}$,
- $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the $\frac{\pi}{2}$ rotation in $\mathbb{R}^2$,
- $\boldsymbol{F}^\perp = \boldsymbol{J}\boldsymbol{F}$ is the orthogonal of the vector $\boldsymbol{F}$ in $\mathbb{R}^2$,

Probability

- $\breve{\Omega}$ is the sample space,
- $\mathbb{P}$ is the reference probability measure,
- $L^2 = L^2(\breve{\Omega}, \mathcal{P}(\breve{\Omega}), \mathbb{P}) = \{X \text{ measurable} \mid \mathbb{E}\|X\|_2^2 < \infty\}$ is the space of random variables with finite variance,
- $X \perp\!\!\!\perp Y$ means that X and Y are independent random variables,
- $\mathbb{E}\{X\}$ is the expectation of X,
- $\widehat{\mathbb{E}}\{X\} = \frac{1}{N}\sum_{i=1}^N X^{(i)}$ is the estimator of $\mathbb{E}\{X\}$,
- $Var(X)$ is the variance of X,
- $\widehat{Var}\{X\} = \frac{1}{N-1}\sum_{i=1}^N \left(X^{(i)} - \hat{\mathbb{E}}\{X\}\right)^2$ is the unbiased estimator of $Var(X)$,
- $\gamma_F(\boldsymbol{x}, \boldsymbol{y}) \triangleq \mathbb{E}\{F(\boldsymbol{x})F(\boldsymbol{y})\}$ is the covariance of F,
- $\Phi_{\boldsymbol{X}}(\boldsymbol{x}) \triangleq \mathbb{E}\{\exp\left(i\boldsymbol{x}\cdot\boldsymbol{X}\right)\}$ is the characteristic function of the random vector $\boldsymbol{X}$ at $\boldsymbol{x}$,
- $\Phi_{\boldsymbol{X}}[\boldsymbol{x}] \triangleq \mathbb{E}\{\exp\left(i\int \boldsymbol{x}\cdot\boldsymbol{X}\right)\}$ is the characteristic functional of the random function $\boldsymbol{X}$ at the function $\boldsymbol{x}$ of $\left(\mathcal{L}^2\right)^d$,
- $L^2 - \lim\limits_{\Delta t \to 0} X(\Delta t)$ is the mean square limit of $X(\Delta t)$,
- $\mathbb{P} - \lim\limits_{\Delta t \to 0} X(\Delta t)$ denotes the limit in probability of $X(\Delta t)$,

- $\boldsymbol{B}_t$ is depending on the context a vector of Brownian motions or a $\mathbb{I}_d$-cylindrical Wiener processes,
- $\dot{\boldsymbol{B}}$ abusively denotes $\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$,
- $< F, G >_t = \int_0^t \boldsymbol{f}_2(t') \cdot \boldsymbol{g}_2(t')\mathrm{d}t'$,
  where $\mathrm{d}F(t) = f_1(t)\mathrm{d}t + \boldsymbol{f}_2(t)\mathrm{d}\boldsymbol{B}_t$ and $\mathrm{d}G(t) = g_1(t)\mathrm{d}t + \boldsymbol{g}_2(t)\mathrm{d}\boldsymbol{B}_t$, corresponds to the quadratic cross-variation of F and G,
- $\mathrm{d}_t F(t, \boldsymbol{m}(t)) = F(\mathrm{d}t, \boldsymbol{m}(t)) = f_1(t, \boldsymbol{m}(t, \boldsymbol{y}))\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{m}(t, \boldsymbol{y}))\mathrm{d}\boldsymbol{B}_t$,
  where $\mathrm{d}F(t, \boldsymbol{x}) = f_1(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{x})\mathrm{d}\boldsymbol{B}_t$, stands for the Eulerian (Itō) time increment of a function F evaluated in the position $\boldsymbol{m}(t, \boldsymbol{y})$,
- $F(\circ\mathrm{d}t, \boldsymbol{m}(t)) = f_3(t, \boldsymbol{m}(t, \boldsymbol{y}))\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{m}(t, \boldsymbol{y})) \circ \mathrm{d}\boldsymbol{B}_t$,
  where $\mathrm{d}F(t, \boldsymbol{x}) = f_3(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{x}) \circ \mathrm{d}\boldsymbol{B}_t$, denotes the Eulerian Stratonovich time increment of a function F evaluated in the position $\boldsymbol{m}(t, \boldsymbol{y})$,
- $\mathrm{d}_t < F(t, \boldsymbol{m}(t)), G(t, \boldsymbol{m}(t)) >= \mathrm{d} < \int_0^t \mathrm{d}_{t'} F(t', \boldsymbol{m}(t')), \int_0^t \mathrm{d}_{t'} G(t', \boldsymbol{m}(t')) >,$
  $= \boldsymbol{f}_2(t, \boldsymbol{m}(t)) \cdot \boldsymbol{g}_2(t, \boldsymbol{m}(t))\mathrm{d}t,$
  if $\mathrm{d}F(t, \boldsymbol{x}) = f_1(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{x})\mathrm{d}\boldsymbol{B}_t$ and $\mathrm{d}G(t, \boldsymbol{x}) = g_1(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{g}_2(t, \boldsymbol{x})\mathrm{d}\boldsymbol{B}_t$ and denotes the Eulerian quadratic cross-variation of F and G,

Analysis

- Ω is a subset of $\mathbb{R}^d$,
- $\mathcal{L}^2\left(\mathbb{R}^d\right) = \left\{F : \mathbb{R}^d \to \mathbb{R} \mid \int_{\mathbb{R}^d} |F|^2 < \infty\right\}$ is the Hilbert space of finite-energy functions from $\mathbb{R}^d$ to $\mathbb{R}$,
- $\|F\|^2_{\mathcal{L}^2(\mathbb{R}^d)} = \int_{\mathbb{R}^d} |F|^2$ is the associated square norm,
- $\|\boldsymbol{F}\|^2_{(\mathcal{L}^2(\mathbb{R}^d))^n} = \int_{\mathbb{R}^d} \|\boldsymbol{F}\|_2^2$ is the square of the classical norm 2 in the space of finite-energy functions from $\mathbb{R}^d$ to $\mathbb{R}^n$,
- $\partial_i F = \partial_{x_i} F = \frac{\partial F}{\partial x_i}$ is the derivative of F with respect to the variable x_i,
- $\boldsymbol{\nabla} F$ is the gradient of F in $\mathbb{R}^2$ or in $\mathbb{R}^3$,
- $\boldsymbol{\nabla}_H F = \begin{pmatrix} \partial_x F \\ \partial_y F \end{pmatrix}$ is the horizontal gradient component of F,
- $\boldsymbol{\nabla}^\perp F = \begin{pmatrix} -\partial_y F \\ \partial_x F \end{pmatrix} = \boldsymbol{J}\boldsymbol{\nabla} F$ is the orthogonal gradient of F in $\mathbb{R}^2$,
- $\boldsymbol{\nabla} \cdot \boldsymbol{F}$ is the divergence of $\boldsymbol{F}$ $\left((\boldsymbol{\nabla} \cdot \boldsymbol{F})_j = \sum_{i=1}^d \partial_i F_{ij} \text{ if } \boldsymbol{F} \text{ is a matrix} \right)$,
- $\Delta \boldsymbol{F}$ is the Laplacian of $\boldsymbol{F}$,
- H_F is the Hessian of F,
- $\mu(\Omega)$ is the Lebesgue measure of Ω,
- $S = \mu(\Omega)$ is the surface of Ω if $d = 2$,
- $\overline{F}$ denotes either the time average, $\frac{1}{T}\int_0^T \mathrm{d}t\ F(t)$, (chapters 6, 7 and 8) or the ensemble and spatial average, $\frac{1}{\mu(\Omega)}\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\ F(\boldsymbol{x})$, (chapter 9) of F,
- $\gamma_F(\boldsymbol{x}) \triangleq \frac{1}{\mu(\Omega)} \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{y}\ F(\boldsymbol{y})F(\boldsymbol{y} + \boldsymbol{x})$ is the empirical covariance,
- $\hat{F}(\boldsymbol{k}) \triangleq \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{x}\ F(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}$ is the Fourier transform of F,

- $\Gamma_F(\boldsymbol{k}) \triangleq \frac{1}{\mu(\Omega)}|\hat{F}(\boldsymbol{k})|^2$ is the spectrum of F,
- $\overline{\Gamma}_F(\kappa) \triangleq \kappa \oint \mathrm{d}\theta_{\boldsymbol{k}}\, \Gamma_F(\boldsymbol{k})$ where $\boldsymbol{k} = \begin{pmatrix} \kappa\cos(\theta_{\boldsymbol{k}}) \\ \kappa\sin(\theta_{\boldsymbol{k}}) \end{pmatrix}$ is the omnidirectional spectrum of F i.e. the spectrum integrated over the wave vector's angles,

Physics

- $\boldsymbol{X}_t = \boldsymbol{\phi}(\bullet, t)$ is the flow,
- J is its Jacobian,
- $\boldsymbol{v}$ is the Eulerian velocity,
- $\boldsymbol{V}$ is the Lagrangian velocity,
- ψ is the two-dimensional streamfunction,
- $D_t F(t, \boldsymbol{x}) = (\mathrm{d}\,(F(t, \boldsymbol{X}_t)))_{|\boldsymbol{X}_t = \boldsymbol{x}}$ is the material derivative of F,
- f and $\boldsymbol{f}$ are scalar and vector Coriolis frequency,
- f_0 is the mean Coriolis frequency,
- β is the mean meridional gradient of the Coriolis frequency,
- g and $\boldsymbol{g}$ are scalar and vector gravity,
- $\boldsymbol{k}$ denotes either the vertical unitary vector or the wave vector depending on the context,
- Θ is a scalar tracer,
- T is the temperature,
- S is the salinity,
- ρ is the density,
- ρ_b is the mean density,
- ρ_0 is the horizontal average of $\rho - \rho_b$,
- $b = -g(\rho - \rho_b - \rho_0)/\rho_b$ referred to as the buoyancy variable,
- p is the pressure,
- $\phi = \frac{p}{\rho_b}$ is the density scaled pressure,
- $Q = \Delta\psi + f + \left(\frac{f_0}{N}\right)^2 \partial_z^2 \psi$ is the stratified QG potential vorticity,
- N is the stratification or Brunt-Väisälä frequency defined by $N^2 = -g\frac{1}{\rho_b}\partial_z \rho_0$,
- L is the horizontal length scale of the dynamics,
- h is the vertical length scale of the dynamics,
- $D = \frac{h}{L}$ is the aspect ratio,
- U is the horizontal velocity scale,
- $T = \frac{L}{U}$ is the advective time scale,
- $Ro = \frac{U}{f_0 L}$ is the Rossby number and expresses the ratio of horizontal advection to rotation (Coriolis terms),
- $Bu = \frac{NH}{f_0 L}$ is the Burger number and expresses the ratio of stratification to rotation,

Modeling under location uncertainty

- $\boldsymbol{w} = \begin{pmatrix} \boldsymbol{u} \\ w \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}$ is the large-scale velocity,
- $\boldsymbol{w}^\star = \boldsymbol{w} + \frac{1}{2}\sum_{i=1}^{d} \left(\boldsymbol{\sigma}\partial_i \boldsymbol{\sigma}_{i\bullet}^T - \partial_i \boldsymbol{\sigma}\boldsymbol{\sigma}_{i\bullet}^T\right)$ is the modified advective velocity,
- $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ is the small-scale velocity,
- $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$ is its horizontal component,
- $\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T = \partial_t \left\langle \int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s, \left(\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_r\right)^T \right\rangle$ referred to as the variance tensor,
- $\boldsymbol{a}_H$ is its horizontal component,
- a_H is the scalar variance tensor in the homogeneous case defined by $\boldsymbol{a}_H = a_H \mathbb{I}_d$,
- $k_c = \sqrt{\frac{2f_0}{a_H}}$ is a cutoff,
- $\mathbb{D}_t F = \mathrm{d}_t F + (\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla} F - \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla} F\right)$ for any function F,
- $\mathbb{D}_t^H F = \mathrm{d}_t F + (\boldsymbol{u}\mathrm{d}t + \boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla}_H F - \boldsymbol{\nabla}_H \cdot \left(\frac{\boldsymbol{a}_H}{2}\boldsymbol{\nabla}_H F\right)$ for any function F,
- A_u is the scaling of the horizontal variance tensor $\boldsymbol{a}_H$,
- $\Upsilon = \frac{UL}{A_u}$ expresses the ratio of horizontal advection to horizontal turbulent diffusion,

Reduced order models

- $\boldsymbol{\phi}_i$ is the i-th spatial mode,
- $\boldsymbol{\phi}_0 = \bar{\boldsymbol{v}}$ is the velocity averaged over time,
- b_i is the i-th temporal mode,
- λ_i is the variance of b_i over time.

Introduction

Context

Understanding, modeling and tracking geophysical fluid dynamics as well as other high Reynolds flows remain main challenges in current researches. Indeed, beyond economical applications related to weather forecasting and industrial flows, accurate climate projections have become a societal need. The complexity of such systems is mainly due to the non-linear and non-local nature of the evolution laws. It makes large-scale flow structures interact with smaller ones. As such, the large-scale flow components cannot be simulated alone (Palmer and Williams, 2008; Slingo and Palmer, 2011). However, the computational expense of Direct Numerical Simulations (DNS) – simulations solving all the hydrodynamical scales – are often beyond reach, especially in geophysics. As an example, the most accurate oceanic currents numerical simulations in the world use a horizontal mesh resolution of about 1 km (Klein et al., 2008; Gula et al., 2015), whereas solving the "real" equations of fluid dynamics, say the Navier-Stokes equation, would require a grid cell of about 1 cm.

The effects of the unresolved so-called turbulent small-scale fluctuations have to be modeled. Turbulent dissipations, advection corrections and backscatterings are the most famous ways to deal with them. They respectively reduce, move and increase energy of large-scale tracers. These different ingredients of the so-called subgrid models will be detailed in the following.

Even using the best subgrid model, the true dynamics is never exactly resolved by a large-scale model. There are always errors. The unresolved subgrid components of tracer or velocity are by definition unknown. In other words, they are *uncertain*. In this thesis, uncertainty and stochasticity does not mean noise but rather something we do not exactly know. Uncertainty can be very coherent. To better express this idea, let us introduce a comparison. When looking at a tree, you cannot see all the branches because they are either too small or hidden by leaves. The particular shapes of these branches are uncertain and could be modeled by random processes. However, you know that they respect some features. For instance, each of them is linked to the trunk through one and only one path. These branches are uncertain coherent structures in the same way as unresolved small-scale vortices are. Both must respect appropriate physical laws. Since resolved and subgrid fluid dynamics are coupled, the large-scale resolved dynamics has to be partially stochastic. Understanding, modeling and simulating this randomness or errors is the subject of Uncertainty Quantification (UQ). It will be the subject of the second part of the introduction.

For some applications, the resolution of deterministic or stochastic large-scale fluid dynamics models described by PDEs is too time consuming. An additional simplification is needed. To alleviate this issue, Reduced Order Models (ROM) can be a solution. They only resolve the most important modes. These models based on both physics and observations are explained in the third part.

Modes in ROM and large-scale models exchange energy. In particular, part of the energy is driven from unstable modes to stable ones. Geometrically, some structures of the flow are stretched and folded, becoming smaller and smaller (direct cascade). In other words, they are mixed. The study of mixing and mixing diagnoses has many theoretical and practical applications in the ocean, but it is also a good way to constrain the parameters of large-scale and reduced order fluid dynamics models. This will be discussed in the fourth and last part.

Let us first present the different modeling principles underlying the constitution of subgrid models.

A classical way to model effects of small-scale velocity on large-scale one relies on the Boussinesq "eddy viscosity" assumption (Kraichnan, 1987; Wilcox, 1988; Menter, 1992; Vallis, 2006). Like its equivalent for large-scale tracer, the so-called eddy diffusivity, this assumption consists in replacing the molecular diffusion coefficient by a higher value in the evolution law of the large-scale quantity. Eddy diffusivity models for the mean tracer are understood in idealized cases. Making use of many assumptions (ergodicity, scale gap, decorrelations, homogeneity, etc), Kraichnan (1987); Vallis (2006) exhibit an eddy diffusivity term. However, the diffusion coefficient or matrix may not be positive, which is a necessary condition for dissipation. The Kraichnan model – a passive scalar advected by a homogeneous and isotropic time-uncorrelated velocity – always shows a positive diffusion coefficient of the mean scalar (Gawędzki and Kupiainen, 1995). Reciprocally, some authors replace molecular diffusion and molecular viscosity respectively by a modification of the transport: a homogeneous and isotropic time-uncorrelated component is added to the real velocity. In such a case, the real, deterministic and diffused passive tracer is equal to the conditional expectation of the spurious randomly transported tracer (Klyatskin et al., 1996; Falkovich et al., 2001). For the Navier-Stokes equations, Constantin and Iyer (2008, 2011) show that the time evolution of the true velocity can be written as the expectation of a function of the random flow. Although eddy diffusivity may be justified, eddy viscosity is hardly supported by theory due to the pressure forcing (Kraichnan, 1987). Despite their lack of theoretical grounds, eddy viscosity models are widely used. They are introduced in Large Eddies Simulation (LES) techniques (Lesieur and Metais, 1996; Pope, 2010) and Reynolds Average Navier-Stokes (RANS) (Wilcox, 1988; Menter, 1992) to model the so-called Reynolds stress tensor. In any case, the new diffusion coefficient or matrix and its temporal and spatial dependence have to be determined. This is often done empirically and/or using scaling assumptions. It should be noticed that scaling assumptions for active and passive tracers, used since Kolmogorov (1941), are challenged by many evidences of anomalous scaling when the flow is non-smooth (Kraichnan, 1994; Gawędzki and Kupiainen, 1995; Chertkov et al., 1996). Only the Kolmogorov 4/5-law (Kolmogorov, 1941; Kupiainen, 2000; Falkovich et al., 2001) – the expression of the velocity third-order moment – is unanimously accepted by the fluid dynamics community, since it is supported both by theory and experiments.

In addition to diffusive terms, some authors suggest correcting the mean advective velocity. Such corrective drift is called noise-induced velocity in climate science and skew diffusion or bolus velocity in geophysics. The skew diffusion consists in decomposing the eddy diffusivity matrix into its symmetric and antisymmetric parts (Nakamura, 2001; Vallis, 2006). The symmetric part, in general positive-definite, leads to diffusion, whereas the antisymmetric one leads to advection by an effective free-divergence velocity. A famous and widely used example of this parametrization is the "bolus" velocity of Gent and Mcwilliams (1990). The corrective drift flattens isopycnals (isodensity surfaces) without any buoyancy dissipation. This reduces the Available Potential Energy (APE) as baroclinic instabilities do. Many improvements and similar methods exist (e.g. Tréguier et al., 1997). Wave literature also refers to a modified advective velocity as Stokes drift. The Stokes drift is the difference between the mean Lagrangian velocity and the mean Eulerian velocity (appearing in averaged transport equation of tracers). It is equal to the correlation between the gradient of the fluctuating component of the Eulerian velocity and the displacement resulting from the integration of the fluctuating Lagrangian velocity component (Leibovich, 1980). For divergence-free velocity it should be proportional to the divergence of an absolute diffusivity matrix. Penland (2003b) refers to a "noise induced drift" with a similar expression also from a Lagrangian point of view using Fokker-Planck equation. Finally, the dynamics of inertial particles embedded in a fluid are known to involve an effective drift. These particles trajectories deviate from the mean fluid velocity. They cluster near the minimum of turbulent kinetic energy (Reeks, 1983; MacInnes and Bracco, 1992), as long as those particles are not too heavy (Belan et al., 2015). This phenomenon is referred to as *turphoresis*. As for waves, the corrective drift is proportional to the divergence of turbulent kinetic energy. However, the time scale in factor is the particle relaxation time rather than the Lagrangian velocity correlation time.

As long as the correction of the advection is divergence-free, all the above models do not modify energy. In contrast, turbulent diffusion can only model direct energy cascade. In real systems, there are also intermittent back-scattering of energy from small scales toward larger scales.

Therefore, some authors proposed to include terms that artificially bring energy to the system in numerical simulations. Locally negative eddy viscosity is sometimes advocated (Protas et al., 2015). However random forcing is more commonly preconized in recent literature. This method is referred to as "stochastic backscaterring". Evolution laws no longer rely on bulk parametrizations *i.e.* deterministic models. They become Langevin equations expressed through Stochastic Differential Equations (SDE) or Stochastic Partial Differential Equations (SPDE). Note that the Langevin equations can be non-linear with complicated random forcing and memory terms. First, let us mention stochastic Lagrangian models of Berloff and McWilliams (2002); Veneziani et al. (2004) where fluid particles are uncorrelated. This means that each fluid particle evolves without taking into account the neighboring particles. The slightly non-linear resulting evolution models are defined on empirical grounds and Gaussian assumptions. A more famous approach is the Eddy-Damped Quasi Normal Markovian (EDQNM) model introduced by Orszag (1970) and Leith (1971). It closes the large-scale Navier-Stokes equations in the Fourier space by neglecting some phase correlations in non-linear terms but keeping energy constant. Chasnov (1991) uses this framework to set up a forced-dissipative Navier-Stokes model where the Eddy Viscosity is scale-dependant and the forcing is Gaussian, homogeneous and isotropic in space and uncorrelated in time. Replacing non-linear interactions by a damping term and a Gaussian forcing is now common practice. This in particular used to setup evolution laws of subgrid scales where accuracy is of lower concern. The solutions are in this case Gaussian processes. Structural Stability Theory (S3T) is one example in that spirit (Farrell and Ioannou, 2014). The Quasi-Linear (QL) approximation separates the non-linear deterministic dynamics of the mean field and the linearized randomly-forced dynamics of fluctuations. Then, stability analyses are applied to the augmented state-vector of mean and covariance in order to reveal and characterize various turbulent phenomenon. Stochastic superparametrization (SSP) proposed a similar model (Grooms and Majda, 2014). The point approximation separates the large-scale and the small-scale dynamics. The small-scale evolution law is linearized and corrected by the introduction of noise and damping terms. To constraint energy conservation in the stationary regime, noise variances and damping coefficients are related. Then, the second order moments of the solution are known analytically and can feed the subgrid tensor expression of the mean large-scale evolution law. Without involving any theoretical closure, Berloff (2005) considers a Gaussian forcing as well. Yet its noise is inhomogeneous in space and correlated in time. Well specified inhomogeneity brings phase information making the model more accurate and the forcing more efficient. Phase information can also be encoded by multiplicative noises. Besides, multiplicative noises are the most common non-Gaussian forcing. Leith (1990) multiplies the white Gaussian noise by a function of the resolved local strain rate. Schumann (1995) uses a quadratic function of a homogeneous Gaussian noise. Brankart (2013) adds at each time step a multiplicative noise to the active tracers (salinity and temperature) before computing the corresponding density. After this, the density is averaged over realizations. The non-linearity of the state-equation makes this transitional variability non-negligible and improves the simulation results. Rather than assuming delta correlation in space, Shutts (2005); Berner et al. (2009, 2011) introduce a spatially homogeneous and isotropic noise. Then, they multiply it by the dissipation rate. This method is called Stochastic Kinetic Energy Backscatter (SKEBS). As in Schumann (1995) and Brankart (2013), the noise is defined from an Ornstein-Uhlenbeck (OU) process (Gaussian process with an exponential covariance in the stationary regime) with a very small correlation time. Such a process corresponds to an Auto-Regressive (AR) process in discrete time. In contrast, Mana and Zanna (2014) use a non-Gaussian noise process uncorrelated in time and space whose variance depends on the resolved PV gradient. Except Schumann (1995), all these methods defined on empirical grounds have a common characteristic: the factor of the multiplicative noise is a function of the gradient of the transported quantity. This suggests a link between stochastic backscatter and turbulent dissipation, but the justification of this link is either unclear or arbitrary.

Shutts (2005); Berner et al. (2009, 2011) can be separated from the other works mentioned above since they focus more specifically on Uncertainty Quantification (UQ) and ensemble forecasting issues. In other words, they seek to quantify statistically the errors of the system dynamics.

To introduce this topic, let us first give a brief description about data assimilation and filtering method. Even beyond fluid dynamics applications, coupling numerical model simulations and measurements is of great interest. This is called data assimilation. Some of these methods derived

from variational principle and optimal control theory are deterministic (Le Dimet and Talagrand, 1986). A functional criterion is optimized in term of control variables to push the model closer to the observations. For instance, in 4D-Var algorithms solutions are functions of time and space but also of initial conditions. Variational optimization performed using the adjoint tangent dynamical model enables us to infer the initial condition which best explains the observation. This thesis will instead be placed in the framework of probabilistic data assimilation methods, such as smoothing and above all filtering (Doucet et al., 2001; Doucet and Johansen, 2009; Candy, 2011; Papadakis et al., 2010). Those methods allow both a random dynamical model as well as noisy and partial observations to guide the ensemble forecast – a set of realizations – along time. The randomness of the model represents error of the dynamical model whereas randomness of observations represents the measurement errors. If the model noise is prominent, the estimated filtered variable relies principally on the observations. Conversely, when the observation noise is dominant, the filtered variable trajectory is mainly driven by the model. For this reason, accurate design of the model errors is crucial in weather and climate communities (Allen and Stainforth, 2002; Penland, 2003a,b; Berner et al., 2015). Furthermore, due to the large dimensionality of the state space, the ensemble size is usually very small. Thus, the randomness of the dynamical model has to be very efficient. Ensemble members, also called particles, have to quickly spread in the phase space. At the same time, each particle should remain a "physically plausible realization".

The randomness can have several sources. Some authors inject randomness through the parameters. Indeed, parameters like viscosity, initial and boundaries conditions are often assumed random (Le Maitre et al., 2002; Sapsis and Lermusiaux, 2012). The chaotic nature of fluid dynamics increases quickly the infinitesimal error related to these parameters. Other authors study the uncertainty arising from forcings. In particular, CO_2 concentration conditions are difficult to specify in climate sciences. Lucarini et al. (2014) approach this problematic with Ruelle response theory. In fluid dynamics, random initial conditions have been widely used for both UQ and predictability studies (e.g. with EDQNM Métais and Lesieur, 1986). In the geophysical data assimilation communities, this method is known to be underdispersive, i.e. it underestimates errors (Berner et al., 2011; Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013; Franzke et al., 2015). As such the model is overconfident. When an observation – often far from the ensemble – is assimilated, only few particles – and in the worst case only one – are considered relevant. The others are discarded. This degeneracy is referred to as filter divergence. The weakness of the method can be explained by at least two facts. At the initial time, the random perturbations lives in a huge state space. Computational limitation leads to the setup of only a small-size ensemble, thus spanning only a small part of the state-space (Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013). As a consequence, without any adhoc compensation, the ensemble variance is underestimated. The compensation method is called covariance inflation (Anderson and Anderson, 1999). The ensemble covariance is multiplied by a carefully-tuned parameter. Furthermore, such a compensation may lead to unphysical behaviors. In any case, the ensemble covariance is usually strongly erroneous. Moreover, these random perturbations are injected at small scales and are hence quickly diffused by the subgrid tensor. From a dynamical system point of view, subgrid tensor makes small-scale components of the solution more stable. Without fully resolved non-linear mechanisms of energy redistribution, the particles tend to align asymptotically with the most unstable directions (Trevisan and Uboldi, 2004; Trevisan and Palatella, 2011; Mitchell and Gottwald, 2012; Sapsis, 2013; Gottwald and Harlim, 2013). Thus, as time evolves the ensemble spans a smaller and smaller space.

Other works preferably address the modeling of errors related to wrong dynamics. Indeed, as long as all the scales are not resolved, subgrid dynamics are modeled rather than resolved. This continuously introduces errors which also grow in time due to chaotic behaviors. A natural way to address this UQ is to continuously introduce noise in the dynamics. The simplest random models are defined from linear Langevin equations with additive Gaussian noise. This has already been discussed above for stochastic backscatter methods. We may add the linear inverse models (Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995). Keating et al. (2012) also uses a linear evolution model with Gaussian noise for a filtering purpose. However, the parameters of the models are themselves OU processes making the solution not Gaussian and more flexible. This method – well known in the filtering community – is referred to as Stochastic Parametrized Extended Kalman Filter (SPEKF) (Gottwald and Harlim, 2013). Among other empirical stochastic models, the Stochastic Perturbed Physics Tendency scheme (SPPT) introduces a correlated multiplicative

noise (Buizza et al., 1999). SPPT and SKEBS methods have been sucessfully applied in operational weather forecast centers (Franzke et al., 2015).

In fluid dynamics, due to non-linearities, the targeted probability density of the solution is highly non-Gaussian and relevant stochastic dynamical models are difficult to derive. In this perspective, an attractive path would be to infer randomness from physics (Berner et al., 2015). Yet as Navier-Stokes equations are deterministic, this path is not straightforward. To derive large-scale fluid dynamical model, the time-scale separation assumption is convenient. In the seventies, Hasselmann (1976) already relied on it for geophysical fluid dynamics. In his seminal work, the large-scale dynamics were encoded by both mean terms and noise terms. However, eventually only simple multidimensional OU process was considered. The time-scale separation assumption is also the foundation of the more rigorous averaging and homogenization theories (Kurtz, 1973; Papanicolaou and Kohler, 1974; Givon et al., 2004; Mitchell and Gottwald, 2012; Gottwald and Melbourne, 2013; Gottwald and Harlim, 2013; Franzke et al., 2015; Gottwald et al., 2015). As the time-scale separation goes to infinity, the large-scale dynamics will converge according to averaging or to homogenization depending on the structure of the global model. The global dynamics as well as the limit large-scale dynamics can be differential equations or SDEs. In the large-scale equation, terms which are only functions of the small-scale variable often tend to converge to a white-noise-in-time term in the Stratonovich sense with a covariance of the Green-Kubo type. Nevertheless, it is not always true for nonlinear dynamics. Sometimes, the noise has to be understood in the sense of Ito or Marcus stochastic integral. In the last case, the noise is a Levy process (Gottwald and Melbourne, 2013; Gottwald et al., 2015). A successful application of the homogenization theory in geophysics are the MTV algorithms (Majda et al., 1999, 2001; Franzke et al., 2005; Majda et al., 2008). MTV refers to the names of the three main authors: Majda, Timofeyev and Vanden-Eijnden. In practice, the non-linearity of the small-scale equation is empirically replaced by a noise term and a damping term before the homogenization procedure. The homogenized dynamics obtained are cubic with correlated additive and multiplicative (CAM) noises. Even without dealing with Levy processes, this structure is able to produce intermittency and extreme events especially because of the CAM noise. This specific form has also been used to infer data-driven models. Peavoy et al. (2015) proposed an example of such a model which uses energy constrained Bayesian estimators and artificial additional observations through Brownian bridge. Another method called invariant manifold theory also invokes a time-scale separation. Yet it relies directly on the SDE solution rather than on its probability density (Givon et al., 2004; Gottwald and Harlim, 2013). It has shown good UQ skills especially because it makes multiplicative noise appear in the limit dynamics. To conclude on these methods, for complicated non-linear dynamical systems, it is still not clear on how to perform homogenization and when this is possible. Moreover, the theory does not make the noise covariance explicit enough and it has to be estimated on data. During this step, some Gaussian approximation are usually done when estimating the coefficients of the model. Some homogenization methods like the MTV algorithm may suffer from energy-conservation issues. Nevertheless, workarounds exist (Frank and Gottwald, 2013; Jain et al., 2014). In addition, the homogenization methods have shown successful results in the context of reduced order models and suggest that geophysical stochastic fluid dynamic models should involve CAM noises. When there is no time-scale separation, the large-scale system can become non-Markovian. For deterministic dynamics, this is readily shown by the Mori-Zwanzig equation (Givon et al., 2004; Gottwald et al., 2015). Indeed, this explicit expression of the large-scale observables of interest involves a memory term.

The study of Mémin (2014) and this thesis follow another approach referred to as models under location uncertainty based on stochastic calculus and Ito-Wentzell formula (Kunita, 1997). This formula as well as basic tools of stochastic calculus will be recalled in Chapter 1. Brzeźniak et al. (1991) first introduces the idea but forgot the Wentzell term of the Ito-Wentzell formula. Then, Mikulevicius and Rozovskii (2004) and Flandoli (2011) introduced the right formula. Their works have focused on pure mathematical aims: existence and uniqueness of SPDE solutions. Neves and Olivera (2015) also studied the wellposedness of similar SPDE using the derivation of the previous authors. In this approach, we understand the large-scale point of view as a subsampling in time. This point of view makes the small-scale component of the velocity look uncorrelated in time. This changes the usual expression of the material derivative and most fluid dynamics equations. This approach referred to as modeling under location uncertainty will be discussed in chapters 2,

3, 4, 5, 6, 7 and 8. Recently, Holm (2015) derived similar evolution laws from the inviscid and adiabatic framework of Lagrangian mechanics. Compared to models under location uncertainty, the stochastic transport of scalars is identical. However, the momentum evolution of Holm (2015) involves an additional term which imposes the helicity conservation but may increase the kinetic energy.

When industrial constraints are such that no simulation of PDE or SPDE is possible – e.g. real time simulation constraint – one has to turn to Reduced Order Models (ROM). In fluid dynamics, they have a lot of similarities with large-scale simulations. Note that a large-scale simulation with periodic boundaries conditions correspond to a ROM where the resolved modes are the Fourier modes with frequencies smaller than the mesh-grid resolution. Obviously, they are better choices than the Fourier basis.

The most famous ROM methods are: Proper Orthogonal Decomposition (POD), Polynomial Chaos (PC), Proper Generalized Decomposition (PGD), Bi-Orthogonal (BO) or Dynamically Orthogonal (DO) methods. The POD (Holmes et al., 1998), also called Empirical Orthogonal Functions method (EOF), on which we will focus, separates the time and the space dependence. If the system is random, the associated POD ROM, which describes the time evolution of the solution, will be random as well (e.g. Franzke et al., 2005; Majda et al., 2008; Sapsis and Majda, 2013c). The PC (Le Maitre et al., 2002) separates the randomness and the space-time dependence. This method, widely used in UQ, especially addresses the modeling of random parameters. The PGD (Chinesta et al., 2011) separates all the coordinates of the solution, which can be spatial and/or temporal coordinates but also parameters or randomness. Finally, the DO (Sapsis, 2011; Sapsis and Lermusiaux, 2012; Sapsis, 2013; Sapsis and Majda, 2013b) and the BO (Choi et al., 2014) reduced order models, which can be considered as generalizations of POD and PC methods (Sapsis, 2011), rely on random temporal modes and time-dependent spatial modes. Some improvements can make the ROM dimension dynamical (Sapsis and Lermusiaux, 2012). Reduced Order Modified Quasilinear Gaussian (ROMQG) methods improve the ROM efficiency by simulating high order moments using few modes through MCMC (Markov Chain Monte Carlo) whereas first and second order moments of a larger set follow differential equations. Sapsis and Majda (2013c) applied this formalism to bases which are constant in time (*e.g.* Fourier modes, POD) whereas Sapsis and Majda (2013b) applied it to DO. The ROMQG method is based on the Modified Quasilinear Gaussian (MQG) method (Sapsis and Majda, 2013a) which is not restricted to dimensionally reduced system. This model approximates the third-order moment in the covariance evolution law in order to redistribute the right amounts of energy between modes. Based on stationary regime information, dampings and noises are specified in that way. Without any additional parametrization, models under location uncertainty naturally conserve the energy. As illustrated in chapters 2, 3 and 8, the combine effect of multiplicative noise and diffusion accurately redistributes the energy at every time instant.

In the POD framework, the velocity is assumed to live in a reduced subspace spanned by the so-called spatial modes. They are learned from data through a Karhunen-Loeve decomposition. The corresponding ROM is generally obtained through a Galerkin projection of the evolution laws onto the spatial modes. The ROM is hence composed of a finite set of coupled ordinary differential equations. They describe the time evolution of a reduced number of velocity modes. However, considering only a small number of modes ignores the small-scale contributions. This yields the same type of issues as in large-scale simulations. In particular, it usually destabilizes the system and shifts the temporal frequencies. To overcome this issue, some authors empirically introduce an eddy viscosity term in the reduced model (e.g. Aubry et al., 1988; Rempfer and Fasel, 1994; Östh et al., 2014; Protas et al., 2015). Other authors (Carlberg et al., 2011) perform non-linear Galerkin methods, with the same spatial modes. This leads to another form of the reduced model, that will not be investigated in this thesis. As explained earlier, MTV algorithms rely on homogenization theories (Franzke et al., 2005; Majda et al., 2008). This method also addresses the UQ issues in ROM. Similar to MTV algorithms, in this thesis we propose new models to be used in the POD framework. We will propose neither a new method of dimensional reduction, nor a new general algorithm improving the performances of a dimensional reduction method. We will show that the transport under location uncertainty enables us to derive deterministic and random POD-ROMs: quadratic systems with or without CAM noises. Coefficients of these ROMs are learned from re-

solved and unresolved modes. Yet the estimators are based on the physical structure of the models under location uncertainty rather than fitting of the evolution of resolved modes. This should make the estimators more robust. Moreover, their low complexity enables their use in large-dimensional systems.

ROMs and other models under location uncertainty express subgrid terms such as additional diffusion with second-order statistics of the unresolved velocity. Although attractive, these expressions cannot be always evaluated since unresolved velocity is hardly observed globally. In such cases, we have to find another way to parametrize subgrid terms in the models under location uncertainty and more generally in large-scale models. To constraint the parametrization, mixing has to be understood. When enough scales are resolved, mixing is the combination of stretching and folding. One can observe this beautiful process in the froth when stirring one's coffee. Stretching stretches pairs of points. Folding bends and folds a set of points, enabling mixing to remain in a finite volume. Mixing quickly generates smaller and smaller structures (direct cascade) until these structures are small enough to be dissipated by molecular viscosity or molecular diffusivity. From this point of view, mixing corresponds to the strengthening of gradients. In contrast, at a given scale, mixing is associated with the dissipation of gradients. Indeed, when structures become too small to be seen at this resolution scale or at larger scales, they are often assumed diluted or uncertain. This point of view is generally used in large-scale models and satellite observations of the ocean. The time and space resolutions involved are generally much larger than the ones associated with the true dynamics. This is the reason for replacing complex advection term by smoothing or other subgrid terms in numerical simulations. Satellite measurements also discard the sparse and possibly corrupted small-scale information. For instance, the processing of altimetry products introduce artificial nuggets in the covariance of the optimal interpolation.

This thesis focuses on the influence of large time step rather than large mesh-grid resolution even though both issues are linked. The models under location uncertainty explicitly use this assumption of large time step and link this point of view to diffusion. Other works focus on advection during a large time step with or without taking into account small-scale velocity influence, and relate it with mixing. Mezić et al. (2010) have introduced a new criterion to diagnose stretching and folding generated by smooth large-scale velocities after a finite time. Contrary to stretching, folding can only occurs after a finite time of advection. Many stretching diagnosis exist. Eulerian ones include in particular the Okubo-Weiss criterion (Okubo, 1970; Weiss, 1991; Shivamoggi and van Heijst, 2011) and improvements (e.g. Lapeyre et al., 1999) but also new criteria (e.g. Haller, 2005). The Okubo-Weiss criterion assumes the vorticity – quantifying rotation – and the strain-rate tensor eigenvectors – defining the directions of instantaneous compression and dilation – to be stationary along the flow. However, these directions often turn with the flow structures. Accordingly, Lapeyre et al. (1999) generalize the idea assuming that the rotation is stationary along the flow only in the local frame of these directions. Note that the objective vortex definition of Haller (2005) is not robust enough in its Eulerian version; in practice, an integration of the flow is necessary. Lagrangian diagnoses often rely on Finite Time Lyapunov Exponents (FTLE) (Pierrehumbert and Yang, 1993; Haller and Yuan, 2000; Thiffeault and Boozer, 2001; Haller and Sapsis, 2011; Keating et al., 2011), Finite Size Lyapunov Exponents (FSLE) (d'Ovidio et al., 2009; Peikert et al., 2014) or diffusivities (see e.g. Keating et al., 2011). Mendoza and Mancho (2010) also proposed a Lagrangian diagnosis but it does not directly measure mixing. In contrast, only few works deal with folding (Thiffeault, 2004; Mezić et al., 2010; Budišić and Thiffeault, 2015; Ma et al., 2016), and they all rely on a Lagrangian point of view.

The Lagrangian advection method is a good example of mixing after a finite time. Initially developed for atmospheric flows (Sutton et al., 1994; Mariotti et al., 1997; Orsolini et al., 2001; Legras et al., 2005), the method is now applied to oceanography (Desprès et al., 2011a,b; Berti and Lapeyre, 2014; Dencausse et al., 2014). The algorithm relies on the advection of a large-scale tracer by a smooth large-scale velocity, in order to increase the resolution of the tracer field. This type of tracer direct cascade is called chaotic advection (Pierrehumbert and Yang, 1993). The original fields are extracted from satellite images or in situ data, at mesoscales (about 100 km and 1 month). The advection is performed without small scales' influence for days or weeks and generates energetic mesoscales and submesoscales (about 10 km) structures. Nevertheless, some of these structures are unphysical. This drawback is mainly explained by the absence of submesoscale

velocities. Thus, spatial smoothing is also used in the method. A question remains: how do we choose the advection time and the filter width for the diffusion? Lagrangian advection algorithm have also been considered with active tracers (Berti and Lapeyre, 2014). Authors suggest to set up advection time with the root mean square of vorticity or velocity gradients. We address this parametrization issue in chapter 9. Beyond the Lagrangian advection method, we aim at better understanding mixing and at using satellite images to parametrize subgrid terms of large-scale models especially the models under location uncertainty.

Preview of chapters

The chapter 1 is an introduction which gathers theoretical results. Chapter 6 is a published paper. Chapters 2, 3, 4, 5 and 7 correspond to manuscripts that have been submitted. Chapters 8 and 9 are manuscripts in preparation. The chapter 10 brings complementary results.

Summary of chapter 1: Stochastic processes and stochastic calculus

In the first chapter, we recall several notions related to random processes and stochastic calculus. After presenting Gaussian processes and Brownian motions in finite and infinite dimensions, we briefly give the basic definitions and main properties of stochastic calculus: finite-variation processes, martingales, semimartingales and quadratic cross-variations. Then, several theorems are detailed; this concerns: Itō-Wentzell formula, Girsanov theorem as well as Kolmogorov's and Fokker-Planck equations. These results supply very powerful tools to handle random functions. In particular, the Itō-Wentzell formula expresses the time derivative of the composition of two random functions. The modeling under location uncertainty developed in chapters 2, 3, 4, 5, 6, 7 and 8 heavily relies on this theorem. We also present the multidimensional versions of two famous processes: the Ornstein-Uhlenbeck (OU) and the geometric Brownian motion. In this finite-dimensional framework, we give some insights on the algebraic structure of the evolution laws involved in the modeling under location uncertainty. After this, we review some methods to estimate diffusion coefficients. This point will be useful for chapters 6, 7 and 8. Before concluding, pros and cons of Stratonovich and Itō notations are discussed. Finally, we open the discussion with more general tools from statistical physics where the differential operators involved in the previously cited theorems are no longer restricted to the second order.

Summary of chapter 2: Geophysical flows under location uncertainty, Part I: Random transport and general models

The first part develops the general framework of modeling under location uncertainty introduced by Mémin (2014). This principle relies on the assumption of a small-scale velocity random and uncorrelated in time. This hypothesis originates from the observation that the time steps of both the simulations and the measurements of geophysical flows are much larger than the smallest hydrodynamical time scales. This assumption changes the material derivative expression introducing an advection correction, a multiplicative noise and an inhomogeneous and anisotropic diffusion. Under divergence-free conditions, this stochastic material derivative has the remarkable property of conserving the energy for each realization. The balanced effects of noise and diffusion enable our model to better take into account energy fluxes between the tracer's components than a deterministic material derivative. In particular, part of the energy of the tracer's (ensemble) mean is continuously transferred to the random tracer component. This random energy exchange constitutes a suitable property for uncertainty quantification. Finally, the Jacobian's evolution law and some similarities with the Kraichnan model are presented.

In the second part, general fluid dynamic conservations – mass, internal energy, salinity and momentum – are expressed in the location uncertainty framework. From them, a stochastic Boussinesq model can be derived. Under the presence of a stationary background gradient (*e.g.* stratification), stochastic forcings of randomly transported quantities are combinations of correlated additive and multiplicative noises. The additive noise transfers energy from the background to the anomaly

and increases its variance. Effects of the additive noise on the transport of buoyancy are study through a toy model of buoyancy oscillations. Assuming the large-scale velocity to be smooth in time enables us to split the momentum equation into two equations: one for the large-scale terms and one for the small-scale terms. This chapter also summarizes the main steps of derivation of models under location uncertainty as well as possible parametrizations.

Summary of chapter 3: Geophysical flows under location uncertainty, Part II: Quasi-geostrophy and efficient ensemble spreading

Following chapter 2, simplified mesoscale models are derived. In addition to the strong stratification and strong rotation assumptions, subgrid tensor terms have to be scaled. The scaling considered in this chapter leads to a Quasi-Geostrophic (QG) model referred to as under moderate uncertainty. The Potential Vorticity (PV) is conserved up to three source terms. The first one is due to inhomogeneities of the advection correction and turbulent diffusion of the velocity; the second one encodes interactions between Coriolis frequency and inhomogeneous subgrid eddies; The last one results from the joint action of resolved and unresolved velocity. Since this last term is uncorrelated in time, it increases the potential enstrophy. Assuming zero PV in the fluid interior yields the classical Surface Quasi-Geostrophic (SQG) relationship between large-scale velocity and buoyancy. As such, the transport under location uncertainty of buoyancy at the surface has been referred to as the SQG model under Moderate Uncertainty (SQG_{MU}).

As an illustration of the potential of the modeling under location uncertainty, numerical simulations of the SQG_{MU} model are performed with a homogeneous unresolved velocity. Two simulations of the deterministic SQG model are also described: one at the same coarse resolution for comparison and one at a higher resolution chosen as a reference. The simulation using the SQG_{MU} model is found to better resolve small-scale structures than the deterministic model at the same resolution. More precisely, the SQG_{MU} model resolves filament instabilities and the subsequent creation of small eddies called "pearl necklaces". Then, the balance between noise and diffusion prescribed by our framework is visually confirmed by spatial and spectral buoyancy fields. After this, an ensemble of simulations has been performed with the SQG_{MU} model. The ensemble accurately estimates the amplitudes and the positions of its own errors at each time step in both spatial and spectral spaces. On the contrary, an ensemble generated by the deterministic model with random initial conditions underestimates the errors by one order of magnitude. Moreover, an analysis of higher order moments of the SQG_{MU} ensemble correctly identifies "pearl necklaces" as cold or warm extreme events.

Summary of chapter 4: Geophysical flows under location uncertainty, Part III: SQG and frontal dynamics under strong turbulence conditions

This chapter explores another scaling of the random Boussinesq model derived in chapter chapter 2. Compared to chapter 3, subgrid dynamics have a stronger influence. However, others terms scale as in chapter 3. The strong rotation assumption leads to a modified geostrophic equilibrium involving a divergence of the horizontal velocity. Then, the strong rotation yields directly a zero PV in the fluid interior and a modified SQG relationship. This model has been referred to as the SQG model under Strong Uncertainty (SQG_{SU}).

The modified geostrophic balance provides a simple relationship between vorticity and divergence of the horizontal velocity. This diagnosis is successfully tested with realistic very-high-resolution model outputs (Gula et al., 2015). This test also enables us to estimate a realistic horizontal diffusion coefficient and to verify *a posteriori* our strong uncertainty assumption. The divergent (ageostrophic) velocity component expresses the frontolysis on warm sides of fronts and the frontogenesis on cold sides of fronts in the upper ocean.

Summary of chapter 5: Bifurcations and location uncertainty in geophysical fluid flows

This chapter is a natural follow-up of chapter 3. Modeling under location uncertainty is also studied through the example of the homogeneous SQG_{MU} model for a flow in a similar geometry.

Nevertheless, the focus here concerns a bifurcation appearing after a predictive time scale.

After briefly recalling the main aspects of modeling under location uncertainty and the SQG_{MU} model, the bifurcation due to a symmetry breaking is studied with the deterministic SQG model at high resolution. A slight modification of the initial conditions leads to a major change after one-month of advection: two scenarios are identified. Due to the unavoidable subgrid tensor, within the same initial condition, the deterministic SQG model with two different resolutions yields two different scenarios. A stochastic approach is hence needed at coarse resolution. Two ensembles are simulated: one with the classical evolution law and random initial conditions and one with the SQG_{MU} model. To visually analyze the ensembles, a Karhunen-Loève decomposition of the ensemble is performed at each time step. In this case study, the first mode of the decomposition is sufficient for a qualitative description. Its estimated probability density function is tracked for both ensembles. Only the ensemble associated with the SQG_{MU} model is found to be statistically converged and to correctly describe both scenarios. Furthermore, this ensemble is closer to the reference. For the method with random initial conditions, the need for a higher number of particles is related to its well-known underdispersive behavior. This is due to both the initial sampling and the badly resolved evolution law. In contrast, the modeling under location uncertainty involves an inhomogeneous random forcing even if the unresolved velocity is homogeneous; this property makes the ensemble spread more efficiently and enables us to use less realizations to represent the ensemble.

Summary of chapter 6: Reduced flow models from a stochastic Navier-Stokes representation

This chapter uses the modeling under location uncertainty to derive a deterministic Reduced Order Model (ROM) for fluid flows.

ROMs express solutions of PDEs or SPDEs on reduced bases. These bases are learned from data and generally optimize a specific criterion. This thesis focuses on Proper Orthogonal Decomposition (POD). The associated basis is a set of functions of space referred to as spatial modes which encodes the maximum of the solution's spatio-temporal variability. The coefficients of the velocity expression in that basis are called temporal modes. After a Galerkin projection of the evolution law onto the basis, additional terms are often needed to model the influence of unresolved modes. In this thesis, we propose two POD ROMs based on two different stochastic Navier-Stokes representations for two different applications: a deterministic model for LES-like simulation in chapters 6 and 7 and a stochastic model for UQ purpose in chapter 8. Both rely on the modeling under location uncertainty but with different assumptions.

In the first approach, the large-scale velocity and hence the resolved temporal modes are assumed to be differentiable with respect to (w.r.t.) time. As in Mémin (2014) and chapter 2, this leads in the context of modeling under location uncertainty to a large-scale Navier-Stokes representation without noise. An advection correction and an inhomogeneous and anisotropic diffusion appear both in the PDE and in the associated ROM. They are parametrized by the one-time one-point quadratic cross-variation matrix of the flow. This inhomogeneous matrix is referred to as the variance tensor. The chapter proposes two ways of estimating it based on the residual velocity. When the variance tensor is assumed to be stationary, a temporal averaging easily yields the result. For the non-stationary case, stochastic calculus still enables to estimate the variance tensor by decomposing it on a temporal basis (Genon-Catalot et al., 1992). By using the temporal modes of the POD as the temporal basis, the ROM remains autonomous. In addition, an optimal time step is estimated on the resolved modes to speed up the ROM simulation and improve the variance tensor estimation. Finally, the ROM successfully reconstructs the temporal modes of two wake flows at Reynolds 300 and 3900. The sub-grid tensor stabilizes the reduced system, even at high Reynolds, and the effective advection corrects the frequency shift.

Summary of chapter 7: Stochastic modeling and diffusion modes for POD models and small-scale flow analysis

This chapter is a follow-up of the previous one. We recall the principle of modeling under location uncertainty, the associated Navier-Stokes representation with smooth large-scale velocity;

the ensuing POD ROM and the two types of variance tensor estimation. The decomposition of the variance tensor on a temporal basis defines what we call the diffusion modes. This chapter proposes an estimation of modal characteristic times instead of a single time as in the previous chapter. For each resolved temporal mode, a characteristic time is associated. Since the variance tensor estimator depends on this characteristic time, the evolution equation of each resolved mode relies on a different variance tensor and hence a different subgrid tensor. This method strongly improves the reconstruction of temporal modes. Without any fit, our reconstruction becomes better than – or at least as good as – the POD ROM with modal eddy viscosity (Rempfer and Fasel, 1994) optimally fitted on the resolved modes' time series. Furthermore, the chapter shows how to analyze the residual velocity influence through the diffusion modes. As such, this chapter proposes new data analysis tools complementary to the usual POD decomposition (or Empirical Orthogonal Functions (EOF)). Such an analysis is impossible through the traditional eddy-viscosity assumption. These diffusion modes unveil the effects of the unresolved velocity component on the resolved components. We identify regions and directions of main turbulent diffusion. We describe in the same manner effective advections induced by this unresolved velocity.

Summary of chapter 8: Conservative transport of fluid velocity under location uncertainty and dimensional reduction

In a second ROM approach, we do not assume the large-scale velocity to be differentiable w.r.t. time anymore. As such, the large-scale velocity is transported under location uncertainty as every other tracer. In addition to advection correction and turbulent diffusion, the large-scale Navier-Stokes model involves a multiplicative noise. This enables more efficient energy transfers and in particular energy conservation as well as UQ. The algebraic structure of the model is presented. It leads to a stochastic ROM with random resolved temporal modes. Due to the mode truncation, a small part of the energy diffused in the ROM leaks to the unresolved modes and enables molecular viscous dissipation in unresolved modes. The rest backscatters in the reduced system through the multiplicative random forcing. This analysis holds for any Galerkin projections onto an orthonormal basis. The result is hence applicable to ROMs as well as random LES-like numerical simulations. Compared to previous chapters, more involved estimators are required to estimate the noises statistics, without resorting to a huge inhomogeneous small-scale velocity spatial covariance function. The estimation formulas for POD ROM will be presented. The method successfully forecasts ROMs of two wake flows at Reynolds 300 and 3900. The ROMs stabilize the unstable modes, prevent the degeneracy of stable modes and, quantify the right amount of errors to keep the ensemble as close as possible to the reference.

Summary of chapter 9: Effects of smooth flows on tracer gradients and tracer spectra

The Lagrangian advection method (Sutton et al., 1994; Desprès et al., 2011a; Berti and Lapeyre, 2014) increases the resolution of satellite images of tracers using advection by a smooth altimetry-derived velocity. This advection creates mesoscales and submesoscales structures. Notwithstanding, due to the coarse spatio-temporal resolution of the involved fields, some of these structures are unphysical. Thus, an additional spatial smoothing of the tracer as well as a careful monitoring of the advection time are necessary. The aim of this chapter is to better understand the formation process of these structures and to give a rule for choosing the advection time and the filter width.

After a finite time, the gradients of an advected smooth tracer have evolved from an initial configuration. We analytically express the final averaged squared norm of these gradients, highlighting the link with the Cauchy-Green tensor, Finite Time Lyapunov Exponents (FTLE) and mesochronic velocity (Mezić et al., 2010). The relative angle of the initial gradient tracer influences the rate of strengthening and decreasing of the tracer's gradients. Yet when the flow and the initial tracer are weakly correlated, the averaged squared norm of the tracer gradients can only increase. Moreover, the growth rate does not depend on the tracer. This gradient strengthening is due to the combined action of stretching and folding – so-called mixing – of tracer structures. By these process, those structures cascade to smaller and smaller scales. This point of view on mixing is associated with smooth tracers. In contrast, numerical simulations that are under-resolved in space or time such

as the models under location uncertainty dilute tracer structures when they are smaller than the resolution scale. Beside passive tracers, such a weak correlation with the flow is also observed in the Lagrangian advection method because of the tracer *prior* spatial smoothing. Furthermore, the geostrophic time scale of the involved velocity is large compared to the evolution of submesoscale tracer gradients. Thus, for this specific application, the Eulerian velocity is almost stationary and creates locally uniform shears and stationary convective cells. The first ones stretch the tracer along a straight line whereas the second ones fold and wrap the tracer around vortices. When the advection time is too long, this wrapping makes unphysical spiral structures appear in the tracer. From these two types of behaviors, we infer a model linear in time for the growth rate of tracer gradients. This model is only parametrized by the large-scale Eulerian velocity. After that, we show that the squared norm mean of these gradients specifies the correlation length of the tracer covariance and thus the position of its spectrum tail. By analytically monitoring the small scales' intensification in the Fourier space, we can specify the filter width of the Lagrangian advection method but also eddy diffusivity to parametrize numerical simulations of large-scale flows such as models under location uncertainty. Numerical experiments with a toy flow as well as real satellite images illustrate each of our theoretical results. In particular, we accurately predict the global and local distribution of stretching over the space of a finite-time advection. This work also highlights the important effect of folding in finite-time mixing.

Summary of chapter 10: Additional results

This chapter first presents the models under location uncertainty with the Stratonovich notations instead of Itō ones. This clarifies the expression of corrective drift in these models. Then, we compare the models under location uncertainty with the recent work of Holm (2015). Finally, we review the derivation of the geophysical models under location uncertainty.

Chapter 1

Stochastic processes and stochastic calculus

Abstract

In this thesis, we use intensively the stochastic calculus formalism, especially in infinite dimension. This formalism, which enables handling evolving random functions, will be interpreted physically in the next chapters. Therefore, in this chapter, we recall useful definitions and properties of this theory. We also present several of its powerful tools, such as the Itō-Wentzell formula. Examples of stochastic differential equations (SDE) are subsequently given. Then, since we intend to make the most of the Itō formalism, some estimation methods based on it are exposed. After that, we discuss the pros and cons of Stratonovich and Itō notations. Finally, we briefly present another stochastic processes theory, which goes beyond stochastic calculus for some results.

In all this section, we use a sample space $\breve{\Omega}$, a probability measure $\mathbb{P}$ and the associated probability space $(\breve{\Omega}, \mathcal{P}(\breve{\Omega}), \mathbb{P})$. Moreover, all the functions of space are assumed to be smooth.

1.1 Gaussian processes and Brownian motions

1.1.1 Definition and notations

Here, we recall quickly what is a Gaussian process. We also fix the notations used in this thesis. A Gaussian process is a function $\boldsymbol{x} \to f(\boldsymbol{x})$ such that, for all n and for all $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$, $(f(\boldsymbol{x}_1), ..., f(\boldsymbol{x}_n))$ is a Gaussian vector. The law of a real Gaussian process is defined by its mean $\boldsymbol{x} \mapsto m(\boldsymbol{x}) \triangleq \mathbb{E}(f(\boldsymbol{x}))$ and its covariance $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \gamma(\boldsymbol{x}, \boldsymbol{y}) \triangleq \mathbb{E}(f(\boldsymbol{x})\bar{f}(\boldsymbol{y}))$. We note $f \sim \mathcal{GP}(m, \gamma)$ to express that f follows a Gaussian distribution of mean m and covariance γ. In the complex case, we need to add a relation function $(\boldsymbol{x}, \boldsymbol{y}) \mapsto C(\boldsymbol{x}, \boldsymbol{y}) \triangleq \mathbb{E}(f(\boldsymbol{x})f(\boldsymbol{y}))$. In this case, the previous notation is extended to $f \sim \mathcal{CGP}(m, \gamma, C)$. Moreover, a function γ is a covariance function if and only if it is symmetric and positive semi-definite.

1.1.2 Finite dimensional Brownian motion

A real multidimensional Brownian motion is a centered Gaussian process with covariance $(t_1, t_2) \to \min(t_1, t_2)\mathbb{I}_d$. Its time increments are decorrelated: $(B_{t_1} - B_{t_2}) \perp\!\!\!\perp (B_{t_3} - B_{t_4})$ if $t_1 < t_2 \leqslant t_3 < t_4$. It is continuous w.r.t. (with respect to) time but not differentiable. In signal processing and in physics, we consider its derivative as a white noise. However, since the covariance of a white noise is a distribution (a Dirac), it is only a formal derivative.

1.1.3 Infinite dimensional Brownian motion: $\mathbb{I}_d$-cylindrical Wiener process

In this thesis, we will deal with stochastic partial differential equations. Therefore, we need a functional Brownian motion $t \mapsto (\boldsymbol{x} \mapsto B_t(\boldsymbol{x}))$ which lives in an infinite dimensional space such as $\mathcal{L}^2\left(\mathbb{R}^d\right)$. We need, similarly to the finite dimensional case, each component to be independent: $B_t(\boldsymbol{x}) \perp\!\!\!\perp B_t(\boldsymbol{y})$ if $\boldsymbol{x} \neq \boldsymbol{y}$. This would lead to a covariance $\mathbb{E}(B_{t_1}(\boldsymbol{x})B_{t_2}(\boldsymbol{y})) = \min(t_1, t_2)\delta(\boldsymbol{x} - \boldsymbol{y})$. This means that its variance is infinite, which is not suitable for a mathematical (or physical) definition. A workaround consists in defining the output of a linear application, when the argument is an infinite dimensional Brownian motion. We will define first a Q-Wiener process and then, and $\mathbb{I}_d$-cylindrical Wiener process, which is the interesting object. The following definitions and properties come from Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007).

If $Q : f \in \mathcal{L}^2\left(\mathbb{R}^d\right) \mapsto Q[f] \in \mathcal{L}^2\left(\mathbb{R}^d\right)$ is a symmetric, non negative, trace class operator, with $Q[f](\boldsymbol{x}) = \int_{\mathbb{R}^d} \breve{Q}(\boldsymbol{x}, \boldsymbol{y})f(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$, then a Q-Wiener process $t \mapsto B_t^Q$ is defined as a centered Gaussian process with covariance:

$$\mathbb{E}\left(B_{t_1}^Q(\boldsymbol{x})B_{t_2}^Q(\boldsymbol{y})\right) = \min(t_1, t_2)\breve{Q}(\boldsymbol{x}, \boldsymbol{y})$$

In this case, due to the trace class property of Q, $\mathbb{E}(B_t^Q(\boldsymbol{x}))^2$ and $\mathbb{E}\|B_t^Q\|^2_{\mathcal{L}^2(\mathbb{R}^d)}$ are bounded.
For the cylindrical Wiener process, the construction is more involved. We will denote by $(e_k)_{k\in\mathbb{N}}$ an orthonormal and complete basis of $\mathcal{L}^2\left(\mathbb{R}^d\right)$ and J the following Hilbert-Schmidt embedding:

$$f \in \mathcal{L}^2\left(\mathbb{R}^d\right) \longmapsto J(f) = \sum_{k\in\mathbb{N}} \frac{1}{k} < f, e_k > e_k \in \mathcal{L}^2\left(\mathbb{R}^d\right).$$

Since $tr(JJ^*) = \sum_{k\in\mathbb{N}} 1/k^2 < \infty$, a (JJ^*)-Wiener process can be defined in $\mathcal{L}^2\left(\mathbb{R}^d\right)$, through the previous definition. Then, an $\mathbb{I}_d$-cylindrical Wiener process, B_t, is defined as a process such as JB_t is an (JJ^*)-Wiener process. This definition does not depend on the embedding J. In practice, a general Hilbert-Schmidt embedding, σ, will be used instead of J defined above.

Finally, the interesting property to remember is the following. If $\boldsymbol{\sigma} : f \in \mathcal{L}^2\left(\mathbb{R}^d\right) \mapsto \boldsymbol{\sigma} f \in \left(\mathcal{L}^2\left(\mathbb{R}^d\right)\right)^d$, with $(\boldsymbol{\sigma} f)(\boldsymbol{x}) = \int_{\mathbb{R}^d} \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{y})f(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$, is a linear operator with good properties (an Hilbert-Schmidt operator), and if $t \to B_t$ is a $\mathbb{I}_d$-cylindrical Wiener process, then $\boldsymbol{\sigma} B_t$ is a centered Gaussian process with covariance:

$$\mathbb{E}\left(\boldsymbol{\sigma}(\boldsymbol{x})B_{t_1}\left(\boldsymbol{\sigma}(\boldsymbol{y})B_{t_2}\right)^T\right) = \min(t_1, t_2)\boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{\sigma}(\boldsymbol{y})^T = \min(t_1, t_2)\int_{\mathbb{R}^d} \breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{z})\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y}, \boldsymbol{z})\mathrm{d}\boldsymbol{z}. \quad (1.1.1)$$

In other words, the $\mathbb{I}_d$-cylindrical Wiener process has the properties of a white noise in space and of a Brownian motion in time.

Furthermore, functional analysis gives us another relevant result. It is possible to express an infinite dimensional Brownian motion as a linear combination of classical one-dimensional Brownian motions. Indeed, if we focus in the example above, the spatial covariance of $\boldsymbol{\sigma} B_t$, $\boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{\sigma}(\boldsymbol{y})^T$, is the kernel of a self-adjoint compact operator. The compactness comes from its trace-class structure. Therefore, this operator is diagonalizable in an orthonormal complete basis. We will denote by $(\boldsymbol{\phi}_k)_{k\in\mathbb{N}}$ this eigenbasis and $(\lambda_k)_{k\in\mathbb{N}}$ its eigenvalues:

$$\boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{\sigma}(\boldsymbol{y})^T = \sum_{k\in\mathbb{N}} \lambda_k \boldsymbol{\phi}_k(\boldsymbol{x})\boldsymbol{\phi}_k^T(\boldsymbol{y}).$$

Q-Wiener theory shows that $\boldsymbol{\sigma} B_t$ can be decomposed as follows:

$$\boldsymbol{\sigma}(\boldsymbol{x})B_t = \sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\beta_t^{(k)}\boldsymbol{\phi}_k(\boldsymbol{x}), \quad (1.1.2)$$

where $\left(\beta_t^{(k)}\right)_{k\in\mathbb{N}}$ are independent classical one-dimensional Brownian motions.

1.2 Basic notions of stochastic calculus

We recall here some basic definitions and properties of finite dimensional stochastic calculus. Here, for simplicity, we deal only with scalar functions of a compact set of time: $t \in [0, T]$ with $T \in \mathbb{R}_+^*$. However, everything can be generalized easily to functions of $\mathbb{R}_+ \times \Omega$ with $\Omega \subset \mathbb{R}^d$ (see Da Prato and Zabczyk, 1992; Prévôt and Röckner, 2007; Kunita, 1997). We use a sample space $\breve{\Omega}$, a probability measure $\mathbb{P}$, a Wiener process, $(B_t)_{t \geqslant 0}$, its filtration $(\mathcal{F})_{t \geqslant 0}$ (the set of σ-algebra generated by each B_t), the whole σ-algebra, $\mathcal{F} \triangleq \mathcal{F}_\infty \triangleq \bigcup_{t \geqslant 0} \mathcal{F}_t$, and the resulting filtered probability space $(\breve{\Omega}, \mathcal{F}, (\mathcal{F})_{t \geqslant 0}, \mathbb{P})$.

1.2.1 Finite variation process

We define first a finite-variations function and, then, a finite-variations process.

A function $t \to f(t)$ is a finite variation function if and only if for all $a < b$ and all partition $a = t_0 < ... < t_n = b$ of $[a, b]$, $\lim\limits_{\Delta t \to 0} \sum_{i=1}^n |f(t_i) - f(t_{i-1})| < \infty$.

$(t, \omega) \to f(t, \omega)$ has finite variation if and only if:

- f is adapted (i.e. $f(t, .)$ is $\mathcal{F}_t$ measurable),
- For each trajectory ω, $f(., \omega)$ is a finite variation function.

Characterization:
f is a finite variation process if and only if $\exists g, f(t, .) = f(0, .) + \int_0^t g(t', .) \mathrm{d}t'$. Classical deterministic functions constitute finite variation processes.

1.2.2 Martingale

$(t, \omega) \to f(t, \omega)$ is a martingale if and only if:

- f is adapted,
- $f(t, .) \in L^1_{\breve{\Omega}} \triangleq \{Y : \mathbb{E}|Y| < \infty\}$,
- $\forall s < t,\ \mathbb{E}(f(t, .)|\mathcal{F}_s) = f(s, .)$.

In particular, if $f = 0$ at $t = 0$, then f is a centered process.
Characterization:
f is a martingale if and only if $\exists g, f(t, .) = f(0, .) + \int_0^t g(t', .) \mathrm{d}B_{t'}$, where $\int_0^t g(t', .) \mathrm{d}B_{t'}$ denotes the Itō integral.

1.2.3 Continuous semimartingale, Itō and diffusion processes

A function f is a continuous semimartingale if and only if it is the sum of a finite variation process and a martingale.
Semimartingales are sometimes referred to as Itō processes. Stochastic calculus deals only with semimartingales. However, in our fluid dynamics representation, we also deal with time-decorrelated processes, formally, the differentiation along the time of a martingale.
A diffusion process is an Itō process which is a solution of an autonomous SDE.

1.2.4 Quadratic variation and quadratic cross-variation

If f and g are semimartingale and $f(t = 0) = g(t = 0) = 0$, then, their quadratic cross-variation, noted $< f, g >$, is the unique finite variation process such as $fg - < f, g >$ is a martingale and $< f, g >_{t=0} = 0$.
Characterization:

- If $f(t,.) = \int_0^t f_1(t')\mathrm{d}t' + \int_0^t f_2(t')\mathrm{d}B_{t'}$ and $g(t,.) = \int_0^t g_1(t')\mathrm{d}t' + \int_0^t g_2(t')\mathrm{d}B_{t'}$, then $< f, g >_t = \int_0^t f_2(t')g_2(t')\mathrm{d}t'$.
  One should notice that, if f_2 and g_2 are random, $< f, g >$ is also random.
- $< f, g >_t = \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_i=0}^{t_n=t} (f(t_i) - f(t_{i-1}))(g(t_i) - g(t_{i-1}))$.

Thus, $< f, g >$ may be interpreted as a kind of "covariance along the time" of the time increments. It should be noted that finite variation processes have a null quadratic variation.

1.2.5 Integral notations

In stochastic calculus, we use integral equations rather than differential equations. It is due to the fact that the Brownian motion is properly defined whereas, its derivative, the white noise, is defined only in the sense of distributions (Arnold, 1974). To simplify the notations, we write $\mathrm{d}f = g\mathrm{d}t + h\mathrm{d}B_t$ instead of $f(t) - f(0) = \int_0^t g(t')\mathrm{d}t' + \int_0^t h(t')\mathrm{d}B_{t'}$ and $\mathrm{d}y = s\mathrm{d}f$ instead of $y(t) - y(0) = \int_0^t s(t')\mathrm{d}f(t') = \int_0^t s(t')g(t')\mathrm{d}t' + \int_0^t s(t')h(t')\mathrm{d}B_{t'}$. However, it is important to keep in mind that we only write integral equations.

1.3 Important theorems

In this section, we cite powerful theorems of stochastic calculus theory. We also show a small property which is very useful in practice: the Itō Isometry. These theorems come from Kunita (1997) and Oksendal (1998).

1.3.1 Itō-Wentzell Formula

Assuming that $(t, \boldsymbol{x}) \to f(t, \boldsymbol{x}) \in \mathbb{R}$ and $(t, \boldsymbol{y}) \to \boldsymbol{g}(t, \boldsymbol{y}) \in \mathbb{R}^d$ are continuous semimartingales (as function of time) and f is smooth enough in space, the composition $(t, \boldsymbol{x}) \to f(t, \boldsymbol{g}(t, \boldsymbol{y}))$ is a continuous semimartingale and (Kunita, 1997, theorem 3.3.1 page 91):

$$\mathrm{d}\left[f(t, \boldsymbol{g}(t, \boldsymbol{y}))\right] = \mathrm{d}_t f + (\boldsymbol{\nabla} f)^T \mathrm{d}\boldsymbol{g} + \frac{1}{2} tr(\boldsymbol{H}_f \mathrm{d} < \boldsymbol{g}, \boldsymbol{g}^T >) + \mathrm{d}_t < (\boldsymbol{\nabla} f)^T, \boldsymbol{g} >, \tag{1.3.1}$$

where $\mathrm{d}_t h(t, \boldsymbol{m}(t)) = h(t + \mathrm{d}t, \boldsymbol{m}(t)) - h(t, \boldsymbol{m}(t))$ denotes the time increments of a function h and $\boldsymbol{H}_f$ the Hessian of f. The time increments d_t may be seen as the equivalent of partial derivative ∂_t in deterministic PDEs. More precisely, the time increments (denoted $h(\mathrm{d}t, \boldsymbol{m}(t))$ in Kunita (1997)) can be defined as $\mathrm{d}_t h(t, \boldsymbol{m}(t)) \triangleq f_1(t, \boldsymbol{m}(t, \boldsymbol{y}))\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{m}(t, \boldsymbol{y}))\mathrm{d}\boldsymbol{B}_t$ if $\mathrm{d}f(t, \boldsymbol{x}) = f_1(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{x})\mathrm{d}\boldsymbol{B}_t$.

The Itō formula is used intensively in stochastic calculus. It replaces the Chain rule of the classical differential calculus. It only applies when f is a deterministic function (or more generally a bounded variation process). In the specific case where f is itself a semimartingale, the term $\mathrm{d}_t < (\boldsymbol{\nabla} f)^T, \boldsymbol{g} >= \mathrm{d} < \int_0^t \mathrm{d}_{t'} (\boldsymbol{\nabla} f)^T, \int_0^t \mathrm{d}_{t'} \boldsymbol{g} >$ is added. The formula is referred to as the Itō-Wentzell formula (or generalized Itō formula) instead of Itō formula. The additional term can be calculated with the Eulerian semi-martingale expansion of f and $\boldsymbol{g}$.

1.3.2 Itō Isometry

If $\mathrm{d}\boldsymbol{f}(t) = \boldsymbol{g}(t)\mathrm{d}\boldsymbol{B}_t$, then:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}(\boldsymbol{f}\boldsymbol{f}^T) = \mathbb{E}(\boldsymbol{g}\boldsymbol{g}^T). \tag{1.3.2}$$

This property is straightforward from the Itō formula:

$$\mathrm{d}\ \mathbb{E}(\boldsymbol{f}\boldsymbol{f}^T) = \mathrm{d}_t \mathbb{E}(\boldsymbol{f}\boldsymbol{f}^T) = \mathbb{E}(\boldsymbol{f}(\boldsymbol{g}\mathrm{d}\boldsymbol{B}_t)^T) + \mathbb{E}(\boldsymbol{g}\mathrm{d}\boldsymbol{B}_t\boldsymbol{f}^T) + \frac{1}{2}\mathbb{E}(\boldsymbol{g}\boldsymbol{g}^T) + \frac{1}{2}\mathbb{E}(\boldsymbol{g}\boldsymbol{g}^T) = \mathbb{E}(\boldsymbol{g}\boldsymbol{g}^T).$$

In the last equality, we use the fact that martingales are centered processes.

1.3.3 Girsanov theorem

Assuming:

- $X_0 = Y_0 = x$,
- $\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t$,
- $\mathrm{d}Y_t = (\gamma(t) + b(Y_t))\,\mathrm{d}t + \sigma(Y_t)\mathrm{d}B_t$,
- b and σ respect the usual smoothness conditions for existence and uniqueness of solution,
- γ is adapted and almost surely in $L^2([0,T])$,
- there exists a process, u, adapted and almost surely in $L^2([0,T])$, such that $\sigma(Y_t)u(t) = \gamma(t)$,
- $\mathbb{E}\left[exp\left(\frac{1}{2}\int_0^T u^2(s)ds\right)\right] < \infty$ (Novikov's condition),
- $M_t \triangleq \exp\left(-\int_0^t u(s)\mathrm{d}B_s - \frac{1}{2}\int_0^t u^2(s)ds\right)$ and $\mathrm{d}\mathbb{Q}(\omega) = M_T(\omega)\mathrm{d}\mathbb{P}(\omega)$ on the filtration.

Then, the $\mathbb{Q}$-law of Y is the same as the $\mathbb{P}$-law of X.
This is a very powerful theorem for several reasons. First of all, it enables to find an attractive formulation for the log-likelihood, as we will see later. Moreover, it enables, at least theoretically, to simulate a SDE with a wrong drift b and to correct this error by changing the law of the solution directly. In MCMC framework like weighted Ensemble Kalman Filter or Particle Filter (Doucet et al., 2001; Doucet and Johansen, 2009; Candy, 2011; Papadakis et al., 2010), it would mean to multiply each particle weight by M_T at time T. Therefore, an equation with a quadratic drift, like in fluid dynamics, could be replaced by an SDE with a linear drift or no drift at all (Romito, 2016). To prevent misunderstanding, note that $\mathbb{Q}$ and $\mathbb{P}$ are the law of the processes $t \mapsto Y_t$ and $t \mapsto X_t$ on the whole interval $[0,T]$. The one-time probability distributions (which are marginal distributions of $\mathbb{Q}$ and $\mathbb{P}$) are not expressed explicitly by this theorem.

1.3.4 Kolmogorov's backward equation

If $\boldsymbol{X}$ is a diffusion process such as $\mathrm{d}\boldsymbol{X}_t(\boldsymbol{x}) = \boldsymbol{W}(\boldsymbol{X}_t(\boldsymbol{x}))\mathrm{d}t + \boldsymbol{\Sigma}(\boldsymbol{X}_t(\boldsymbol{x}))\mathrm{d}\boldsymbol{B}_t$ and $\boldsymbol{X}_{t_0}(\boldsymbol{x}) = \boldsymbol{x}$, then, its infinitesimal generator $\mathcal{L}$ is defined as:

$$\mathcal{L} = \boldsymbol{W} \cdot \boldsymbol{\nabla} + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^T\boldsymbol{\nabla}\boldsymbol{\nabla}^T\right), \tag{1.3.3}$$

and for any mean observable $h(\boldsymbol{x},t) = \mathbb{E}\left\{f\left(\boldsymbol{X}_t(\boldsymbol{x})\right)\right\}$, we have:

$$\partial_t h = \mathcal{L}h. \tag{1.3.4}$$

Hence, for any nonlinear dynamics, one can write down an equivalent closed linear differential equation of the mean observable of interest e.g. the mean large-scale tracer of a flow. At time t, if the solution of the original dynamical system lives in a finite-dimensional space, the equation is a partial differential equation. If the solution of the original dynamical system is a function, the variable of interest at time t is seen as a functional of the initial conditions at all scales. The equivalent differential equation is a Hopf equation, *i.e.* an equation with variational derivatives (Fréchet derivative) instead of partial derivatives (Da Prato and Zabczyk, 1992).

1.3.5 Fokker-Planck equation

Under the same assumptions, we have the evolution of the one-time probability density function of the solution replacing the generator $\mathcal{L}$ by its adjoint $\mathcal{L}^*$:

$$\partial_t f_X = \mathcal{L}^* f_X. \tag{1.3.5}$$

For deterministic dynamics, the equation is of order one and is referred to as the Liouville equation (Penland, 2003b). Furthermore, the semigroups of Kolmogorov and Fokker-Planck equations are in this case the Koopman and the Ruelle-Perron-Frobenius operators respectively (Koopman, 1931).

1.4 Example of classical processes

1.4.1 Ornstein-Uhlenbeck process

In $\mathbb{R}^n$, the SDE of a multidimensional Ornstein-Uhlenbeck (OU) process is:

$$\mathrm{d}\boldsymbol{X}_t = -\boldsymbol{A}\boldsymbol{X}_t\mathrm{d}t + \boldsymbol{C}\mathrm{d}\boldsymbol{B}_t,$$

where $\boldsymbol{B}_t$ is a multidimentional Brownian motion. Introducing $\boldsymbol{Z}_t = e^{\boldsymbol{A}t}\boldsymbol{X}_t$ (which is a martingale since $\mathrm{d}\boldsymbol{Z}_t = e^{\boldsymbol{A}t}\boldsymbol{C}\mathrm{d}\boldsymbol{B}_t$) yields:

$$\boldsymbol{X}_t = e^{-\boldsymbol{A}t}\boldsymbol{X}_0 + \int_0^t e^{\boldsymbol{A}(s-t)}\boldsymbol{C}\mathrm{d}\boldsymbol{B}_s = e^{-\boldsymbol{A}t}\boldsymbol{X}_0 + \int_0^t e^{-\boldsymbol{A}s}\boldsymbol{C}\mathrm{d}\tilde{\boldsymbol{B}}_s,$$

where $\tilde{\boldsymbol{B}}_s = \boldsymbol{B}_{t-s}$ is also a Brownian motion. $\boldsymbol{X}_t$ is a Gaussian process centered on $e^{-At}\mathbb{E}(X_0)$. Moreover, if $\boldsymbol{C}$ and $\boldsymbol{C}^T$ commute with $\boldsymbol{A}$, the Itō isometry on $\boldsymbol{Z}_t$ and the independence of Brownian motion increments leads to the covariance function of $\boldsymbol{X}_t$:

$$\boldsymbol{\gamma}_{OU}(t+\tau,t) = \boldsymbol{C}\boldsymbol{C}^T e^{-\boldsymbol{A}|\tau|}\left(\mathbb{I}_d - e^{-(\boldsymbol{A}+\boldsymbol{A}^T)t}\right)(\boldsymbol{A}+\boldsymbol{A}^T)^{-1} \underset{t\to\infty}{\longrightarrow} \boldsymbol{C}\boldsymbol{C}^T e^{-\boldsymbol{A}|\tau|}(\boldsymbol{A}+\boldsymbol{A}^T)^{-1}.$$

For the limit, we assume that $(\boldsymbol{A}+\boldsymbol{A}^T)$ has positive eigenvalues. Otherwise the variance diverges. Thus, after some time, the OU process becomes stationary, with exponential covariance function. Its finite correlation time is given by the inverse of eigenvalues of $\boldsymbol{A}$.

1.4.2 Geometric Brownian motion

The one-dimensional geometric Brownian motion is defined by the following SDE:

$$\mathrm{d}X_t = -aX_t\mathrm{d}t + X_t\alpha\mathrm{d}B_t,$$

where the constants a and α have real or complex values. Again the variable change, $Z_t = e^{at}X_t$, removes the finite-variations part of the SDE: $\mathrm{d}Z_t = Z_t\alpha\mathrm{d}B_t$. Then by the Itō formula:

$$Z_t = Z_0\exp\left(\alpha B_t - \frac{1}{2}\alpha^2 t\right) \text{ and } X_t = X_0\exp\left(\alpha B_t - \left(\frac{1}{2}\alpha^2 + a\right)t\right).$$

If α is real, X_t has a log-normal law. Since this law has a heavy tail, a geometric Brownian motion is usually more suitable than an OU process to model phenomena with extreme events, such as turbulent flows. If α is purely imaginary, X_t turns in the complex plane with a random Gaussian phase. The variance of this phase increases linearly with time. After a transient state, the phase modulo 2π overlaps and converges to a uniform distribution on $[-\pi,\pi]$. If a is small enough, its amplitude increases exponentially. For large a, it decreases exponentially. In any case, the probability density function of the solution can be written analytically.

In a n-dimensional case, a geometric Brownian motion equation can be formulated as follows:

$$\mathrm{d}\boldsymbol{X}_t = -\boldsymbol{A}\boldsymbol{X}_t\mathrm{d}t + \mathrm{d}\boldsymbol{C}\boldsymbol{X}_t \text{ with } \boldsymbol{C} = \sum_{k=1}^{n^2}\boldsymbol{\alpha}^{(k)}B_t^{(k)}, \tag{1.4.1}$$

where $\boldsymbol{X}\in\mathbb{R}^n$ and $\mathrm{d}\boldsymbol{C}$ the time differentiation of the Gaussian matrix $\boldsymbol{C}$. The coefficients $\left(B^{(k)}\right)_k$ denote independent one dimensional Brownian motions, whereas the matrices $\left(\boldsymbol{\alpha}^{(k)}\right)_k$ are assumed constant. If the $\boldsymbol{\alpha}^{(k)}$ commute and if $\boldsymbol{A}$ commutes as well with these matrices, we can then express the solution as in the one dimensional case through the exponential of a Gaussian matrix:

$$\boldsymbol{X}_t = \exp\left(\boldsymbol{C} - \left(\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} < \boldsymbol{C},\boldsymbol{C} > + \boldsymbol{A}\right)t\right)\boldsymbol{X}_0, \tag{1.4.2}$$

$$= \exp\left(\sum_{k=1}^{n^2}\boldsymbol{\alpha}^{(k)}B_t^{(k)} - \left(\frac{1}{2}\sum_{k=1}^{n^2}\left(\boldsymbol{\alpha}^{(k)}\right)^2 + \boldsymbol{A}\right)t\right)\boldsymbol{X}_0. \tag{1.4.3}$$

Nevertheless, as far as we know, in the general case, the expression of the solution cannot be written explicitly. However, the mean of $\boldsymbol{X}_t$ is explicitly known and, by Itō formula, the evolution equation of its one-time covariance matrix $\boldsymbol{S} \triangleq \mathbb{E}\left((\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t))(\boldsymbol{X}_t - \mathbb{E}(\boldsymbol{X}_t))^T\right)$ can be formulated as:

$$\mathbb{E}(\boldsymbol{X}_t) = e^{-\boldsymbol{A}t}\boldsymbol{X}_0, \tag{1.4.4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{S} = -(\boldsymbol{A}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{A}^T) + \sum_{k=1}^{n^2} \boldsymbol{\alpha}^{(k)}\boldsymbol{S}\left(\boldsymbol{\alpha}^{(k)}\right)^T + \sum_{k=1}^{n^2} \boldsymbol{\alpha}^{(k)}\mathbb{E}(\boldsymbol{X}_t)\mathbb{E}(\boldsymbol{X}_t)^T\left(\boldsymbol{\alpha}^{(k)}\right)^T. \tag{1.4.5}$$

As in the one-dimensional case, we expect extreme events and/or rotations depending on the eigenvalues of the matrices $(\boldsymbol{\alpha}^{(k)})_k$. Real eigenvalues should lead to extreme events with exponentially decreasing amplitudes. Purely imaginary eigenvalues should undergo rotations with random phases, and, if A is small, we get exponentially increasing amplitude.

Our stochastic fluid flow dynamics setup – the models under location uncertainty – developed in the chapters 2, 3, 5 and 8 deals naturally with similar types of SDEs. Yet, our setup deals with SPDE instead of SDE, qualitative behavior are expected to be similar after spatially discretizing the SPDE. The slight difference due to numerical truncation error will be discussed in chapter 8. In our setup, the matrix $\boldsymbol{A}$ may be state-dependent ($\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{X}_t)$) and the matrices $(\boldsymbol{\alpha}^{(k)})_k$ do not commute. However, $\boldsymbol{W} = \boldsymbol{A}(\boldsymbol{X}_t) + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} < \boldsymbol{C}, \boldsymbol{C} >$ and $\boldsymbol{C}$ are antisymmetric. Hence, they are diagonalizable in $\mathbb{C}$ with pure imaginary eigenvalues. This imposes in particular energy conservation for each realization:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{X}\|_2^2 = 0. \tag{1.4.6}$$

For $n = 2$ and $\boldsymbol{W}$ constant, the diagonalization of C reduces the system to a one-dimensional geometric Brownian motion in $\mathbb{C}$. This problem can easily be solved and the solution rotates randomly (see above). For higher dimensions, the matrix $\frac{\mathrm{d}}{\mathrm{d}t} < \boldsymbol{C}, \boldsymbol{C} >$ being symmetric negative, it can be diagonalized in $\mathbb{R}$ with negative eigenvalues and the SDE (1.4.1) can be further simplified. In the following chapters, $\frac{\mathrm{d}}{\mathrm{d}t} < \boldsymbol{C}, \boldsymbol{C} >$ will correspond to a Laplace-Beltrami operator encoding turbulent diffusion. Unfortunately, that simplification does not seem sufficient to explicitly write the semi-group as in (1.4.2) because of the correlations between coefficients of the matrix $\boldsymbol{C}$. Nevertheless, free probability in the random matrix theory deals with similar equations (Biane, 1997; Delyon and Yao, 2006; Demni, 2008; Lévy, 2008; Delyon, 2010; Demni and Hmidi, 2012; Cébron, 2014). The matrix $-\sqrt{-1}\boldsymbol{C}$ seems similar to a hermitian Brownian matrix and the semigroup of our equation seems similar to a unitary Brownian matrix when $\boldsymbol{W} = 0$. Many results exist especially when the dimension of the system goes to infinity: a very relevant limit for fluid dynamics. At the limit, the unitary Brownian matrix is called a free multiplicative Brownian matrix. The limit of the spectral distribution of the unitary Brownian matrix is well described. We remind that the spectral distribution of a random matrix $\boldsymbol{M}$ is the random measure:

$$\mu_M = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k}, \tag{1.4.7}$$

where $(\lambda_k)_k$ are the eigenvalues of $\boldsymbol{M}$. However, random matrix theory and stochastic differential geometry have not been further studied during the PhD due to lack of time.

1.5 Estimation methods for the noise

In this thesis, we try to benefit from the possibilities offered by stochastic calculus formalism. Therefore, we have attempted to base our estimation methods on it as well. The statistics based on stochastic calculus mostly used in finance are called the processes statistics and are at the heart of a huge literature (Genon-Catalot et al., 1992; Florens-Zmirou, 1993; Genon-Catalot and Jacod, 1993; Hofmann et al., 1999; Sørensen, 2004; Nourdin, 2004; Comte et al., 2007; van Waaij and van Zanten, 2016). See Rao (1999) for a review of existing methods. Hereafter, we present some of these methods as well as an estimation algorithm coming from statistical physics. We will consider

a general SDE

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(t, \boldsymbol{X}_t)\mathrm{d}t + \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{B}_t. \tag{1.5.1}$$

Let us note that $\boldsymbol{X}_t$ does not have to be the position of a particle. It can be the solution of a discretized SPDE.
In many cases, we can estimate $\boldsymbol{w}$ using for instance a low-pass filter. After removing it from the time series, the problem reduces to:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{\sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{B}_t. \tag{1.5.2}$$

Generally, only one realization of this semimartingale is available. In the following, $N+1$ will denote the number of time steps and Δt the time step of a path of the observed semimartingale realization.

1.5.1 Constant diffusion coefficient

A constant diffusion coefficient may be relevant for stationary Eulerian statistics:

$$\mathrm{d}\boldsymbol{X}_t(\boldsymbol{x}) = \boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t. \tag{1.5.3}$$

$\boldsymbol{\sigma}(\boldsymbol{x})$ can be estimated separately for each $\boldsymbol{x}$.

Local covariance matrix estimation

In equation (1.5.3), the process $\boldsymbol{X}$ is the multidimensional Brownian motion $\boldsymbol{\sigma}\boldsymbol{B}_t$. The estimation of $\boldsymbol{a}\Delta t = \boldsymbol{\sigma}\boldsymbol{\sigma}^T\Delta t$ is straightforward and stochastic calculus is not required. Indeed, the centered discrete-time process of small time increments of $\boldsymbol{X}$, $\Delta\boldsymbol{X}_{t_i} \triangleq \boldsymbol{X}_{(i+1)\Delta t} - \boldsymbol{X}_{i\Delta t} = \boldsymbol{\sigma}(\boldsymbol{B}_{(i+1)\Delta t} - \boldsymbol{B}_{i\Delta t})$, is decorrelated in time and Gaussian. Therefore, the elements of this set are independent and identically distributed. So, the variance, $\boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)^T\Delta t$, can be computed averaging the products of these time increments of $\boldsymbol{X}$. In other words, the process is ergodic.

$$\boldsymbol{a}\Delta t \quad = \quad \mathbb{P} - \lim_{N\to+\infty} \frac{1}{N}\sum_{i=1}^{N} \Delta\boldsymbol{X}_{t_i}(\Delta\boldsymbol{X}_{t_i})^T. \tag{1.5.4}$$

The convergence is almost sure by the strong law of large numbers, but, here, only the convergence in probability will be used. It can be noticed that $\boldsymbol{a}$ is the quadratic covariation of $\boldsymbol{X}$. Then, even if $\boldsymbol{w} \neq 0$,

$$\begin{aligned}
\boldsymbol{a}\Delta t &= \frac{1}{N}\boldsymbol{a}(N\Delta t) = \frac{1}{N} < \boldsymbol{X}, \boldsymbol{X}^T >= \frac{1}{N}\ \mathbb{P} - \lim_{\Delta t\to 0}\sum_{i=1}^{N} \Delta\boldsymbol{X}_{t_i}(\Delta\boldsymbol{X}_{t_i})^T, \\
&= \mathbb{P} - \lim_{\Delta t\to 0}\frac{1}{N}\sum_{i=1}^{N} \Delta\boldsymbol{X}_{t_i}(\Delta\boldsymbol{X}_{t_i})^T.
\end{aligned}$$

If the time interval of observation is fixed, the limit $\Delta t \to 0$ is equivalent to $N \to +\infty$ and we retrieve (1.5.4). This estimate will be used in chapters 6, 7 and 8. In practice, the process $\boldsymbol{X}$ is not a real Brownian motion and the time step Δt has to be chosen carefully to obtain $\boldsymbol{a} = \frac{1}{\Delta t}\left(\frac{1}{N} < \boldsymbol{X}, \boldsymbol{X}^T >\right)$. Although simple, this method needs an empirical estimate of Δt. Chapters 6 and 7 will propose a way to choose this time step from the smallest time scale of the large-scale velocity dynamics.

Absolute diffusivity

The Brownian motion, $\boldsymbol{X}$, may be seen as the limit of another process $\boldsymbol{Y}^L$ smooth in time, when the correlation time τ_Y of the associated velocity goes to zero. Generally, this picture is described with a stationary and homogeneous Lagrangian velocity, $\boldsymbol{w}_Y^L$, (Falkovich et al., 2001; Penland,

2003b; Klyatskin, 2005; Vallis, 2006; Keating et al., 2011). Denoting, $\boldsymbol{w}_Y^L$, the Lagrangian velocity and

$$\boldsymbol{\gamma}_{w_Y^L}(s) = \widetilde{\boldsymbol{\gamma}}_{w_Y^L}(s/\tau_Y) = \mathbb{E}\left\{\boldsymbol{w}_Y^L(t_1)\left(\boldsymbol{w}_Y^L(t_1+s)\right)^T\right\}, \tag{1.5.5}$$

the stationary covariance of this Lagrangian velocity, yields

$$\begin{aligned}
\boldsymbol{a} &= \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left\{\boldsymbol{X}_t\boldsymbol{X}_t^T\right\}, && (1.5.6)\\
&= \lim_{\tau_Y\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left\{\boldsymbol{Y}_t^L\left(\boldsymbol{Y}_t^L\right)^T\right\}, && (1.5.7)\\
&= \lim_{\tau_Y\to 0} 2\Pi_S\int_0^t \mathrm{d}s\ \mathbb{E}\left\{\boldsymbol{w}_Y^L(0)\left(\boldsymbol{w}_Y^L(s)\right)^T\right\}, && (1.5.8)\\
&= \lim_{\tau_Y\to 0} 2\Pi_S\int_0^{t/\tau_Y} \mathrm{d}s\ \widetilde{\boldsymbol{\gamma}}_{w_Y^L}(s), && (1.5.9)\\
&= 2\Pi_S\int_0^{\infty} \mathrm{d}s\ \widetilde{\boldsymbol{\gamma}}_{w_Y^L}(s), && (1.5.10)\\
&= 2\Pi_S\int_0^{\infty} \mathrm{d}s\ \boldsymbol{\gamma}_{w_Y^L}(s), && (1.5.11)
\end{aligned}$$

where $\Pi_S(\boldsymbol{M}) = \frac{1}{2}\left(\boldsymbol{M}+\boldsymbol{M}^T\right)$ is the projection on the space of symmetric matrices. The expectation is often defined from an average over initial conditions. This estimator does not depend on any time step Δt.

Usually , $\boldsymbol{w}$ and $\boldsymbol{\sigma}(s)$ are Lagrangian velocities. Accordingly, $\boldsymbol{Y}_t^L = \int_0^t \boldsymbol{w}_Y^L(s)\mathrm{d}s$ $\boldsymbol{X}_t = \int_0^t \boldsymbol{\sigma}(s)\mathrm{d}\boldsymbol{B}_s$ are fluid particle displacements labeled by their initial conditions $\boldsymbol{x}_0$. If $\boldsymbol{w}_Y^E(\boldsymbol{x},s)$ and $\boldsymbol{\sigma}^E(\boldsymbol{x},s)\dot{\boldsymbol{B}}$ are Eulerian velocities, $\boldsymbol{Y}_t^E = \int_0^t \boldsymbol{w}_Y^E(\boldsymbol{x},s)\mathrm{d}s$ and $\boldsymbol{X}_t^E = \int_0^t \boldsymbol{\sigma}^E(\boldsymbol{x},s)\mathrm{d}\boldsymbol{B}_s$ have no physical meanings as in (1.5.3). Yet, the same type of estimation is possible for each $\boldsymbol{x}$ of the Eulerian space. Spatial inhomogeneities of the quadratic variation, $\boldsymbol{a}$, can be estimated. Note that the expectation has too be replaced by a spatially or temporally averaging.

Diffusion semigroup

Many other methods allow us to estimate the quadratic variation. Some authors diagonalize numerically the semigroup $P_{\Delta t} = \exp\left(\mathcal{L}\Delta t\right)$ where $\mathcal{L}(\boldsymbol{x}) = \frac{1}{2}\mathrm{tr}\left(\boldsymbol{a}(\boldsymbol{x})\boldsymbol{\nabla}\boldsymbol{\nabla}^T\right)$ is the infinitesimal generator of the diffusion (Rao, 1999; Crommelin and Vanden-Eijnden, 2006; Gottwald et al., 2015). $P_{\Delta t}$ is evaluated using

$$(P_{\Delta t}h)(\boldsymbol{x}) = \mathbb{E}\left\{h(\boldsymbol{X}_{\Delta t})|\boldsymbol{X}_0 = \boldsymbol{x}\right\}. \tag{1.5.12}$$

The eigenvectors and eigenvalues of $P_{\Delta t}$ lead to the eigenvectors and eigenvalues of $\mathcal{L}$. Then, the quadratic covariations encoded in $\boldsymbol{a}$ are optimally chosen in order to make the operator $\mathcal{L}$ fit this diagonalization. Nonetheless, this method is computationally demanding.

Parametric estimation using likelihood

In the literature of processes statistics, a lot of parametric estimation methods rely on likelihood. Indeed, denoting $\boldsymbol{\theta}$ the parameters, even for a non-Gaussian diffusion process, such as

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t|\boldsymbol{\theta})\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t|\boldsymbol{\theta})\mathrm{d}\boldsymbol{B}_t, \tag{1.5.13}$$

there is a simple expression of the joint likelihood $p\left(\{\boldsymbol{X}_t'|0\leqslant t'\leqslant t\}|\theta\right)$. The Girsanov theorem (Oksendal, 1998) cited previously leads as explained in Rao (1999) to the following log-likelihood:

$$\begin{aligned}
l\left(\{\boldsymbol{X}_{t'}|0\leqslant t'\leqslant t\}|\boldsymbol{\theta}\right) &= \int_0^t \boldsymbol{w}(\boldsymbol{X}_{t'}|\boldsymbol{\theta})^T\boldsymbol{a}(\boldsymbol{X}_{t'}|\boldsymbol{\theta})^{-1}\mathrm{d}\boldsymbol{X}_{t'}\\
&\quad -\frac{1}{2}\int_0^t \boldsymbol{w}(\boldsymbol{X}_{t'}|\boldsymbol{\theta})^T\boldsymbol{a}(\boldsymbol{X}_{t'}|\boldsymbol{\theta})^{-1}\boldsymbol{w}(\boldsymbol{X}_{t'}|\boldsymbol{\theta})\mathrm{d}t'. && (1.5.14)
\end{aligned}$$

Note that the Girsanov theorem does not give us the conditional probability density function of $\boldsymbol{X}_t$ at time t but only the joint probability density function of $\{\boldsymbol{X}_{t'}|0 \leqslant t' \leqslant t\}$. The above formula is widely used to perform maximum likelihood estimations and Bayesian estimations (Rao, 1999; Sørensen, 2004; van Waaij and van Zanten, 2016). Nevertheless, maximum likelihood estimation in continuous time is hardy possible if $\boldsymbol{\sigma}$ depends on θ. In this case, the estimation method must be derived either from the discretized version of the stochastic differential equation (1.5.13) or from a quadratic version computation (1.5.4).

For geophysics, it could be a very powerful tool for at least two main applications. First, it could be used on Lagrangian data. Secondly, the transport under location uncertainty (see chapters 2, 3, 4, 5 and 8) provides SPDEs to describe geophysical tracers evolutions. After spatial discretization, these SPDEs reduce to the form (1.5.13). A sequence of satellite images of a tracer could hence be used to estimate a parametrization of the stochastic model. To directly deal with the SPDE, Da Prato and Zabczyk (1992) derived an infinite-dimensional version of the Girsanov theorem. A similar expression of the log-likelihood could hence be derived in infinite dimension.

1.5.2 The diffusion coefficient is time varying

Nonparametric method by projection using quadratic variation

Most processes statistics algorithms rely on the hypothesis that $\boldsymbol{\sigma}$ is either constant in time or in space (Genon-Catalot et al., 1992; Florens-Zmirou, 1993; Genon-Catalot and Jacod, 1993; Hofmann et al., 1999; Comte et al., 2007). Again, when we have access to a global Eulerian realization of the velocity, we can assume that $\boldsymbol{\sigma}$ do not depend on $\boldsymbol{X}$ in (1.5.2). This assumption enables us to estimate, for all functions h_k, the coefficients:

$$\begin{aligned}
\boldsymbol{c}_k &\triangleq \int_0^T h_k(t)\boldsymbol{a}(\boldsymbol{x},t)\mathrm{d}t, \\
&= \int_0^T h_k(t) d\left\langle \boldsymbol{X}, (\boldsymbol{X})^T \right\rangle_t, \\
&= \mathbb{P}-\lim_{\Delta t\to 0}\sum_{t_i=0}^{T} h^k(t_i)\Delta\boldsymbol{X}_{t_i}(\Delta\boldsymbol{X}_{t_i})^T.
\end{aligned}$$

Choosing the functions h^k as an orthonormal basis of $\mathcal{L}^2([0,T])$ allows to express the tensor $\boldsymbol{a}$ as:

$$\boldsymbol{a}(t) = \sum_{k\in\mathbb{N}} \boldsymbol{c}_k h_k(t).$$

In the case of a wavelet basis Genon-Catalot et al. (1992) show that such estimators have good statistical properties: local asymptotic normality of the integrated square errors, together with a known rate of convergence of its bias and variance. This estimate will be used in chapter 6 and 7.

1.6 Stratonovich versus Itō

The discussion of relying whether on Stratonovich or Itō convention for defining the stochastic integral is recurrent in physics. Here, we recall the definitions and the links between those two notations. This will highlight a first important point. Under appropriate assumptions, it is easy to switch from one integral notation to the other. As a consequence, the most convenient form can be used to tackle a given issue. Then, we describe the advantages and disadvantages of each formalism.

1.6.1 Definitions

The Stratonovich integral is defined as follows (Oksendal, 1998):

$$\begin{aligned}\int_0^t \boldsymbol{\Sigma}(t)\circ \mathrm{d}\boldsymbol{B}_t &\triangleq L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{t}\boldsymbol{\Sigma}\left(\frac{t_i+t_{i+1}}{2}\right)\left(\boldsymbol{B}_{t_{i+1}}-\boldsymbol{B}_{t_i}\right),\\ &= L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{t}\frac{\boldsymbol{\Sigma}\left(t_i\right)+\boldsymbol{\Sigma}\left(t_{i+1}\right)}{2}\left(\boldsymbol{B}_{t_{i+1}}-\boldsymbol{B}_{t_i}\right),\end{aligned}$$

whereas the Itō one is defined as:

$$\int_0^t \boldsymbol{\Sigma}(t)\mathrm{d}\boldsymbol{B}_t \triangleq L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{t}\boldsymbol{\Sigma}(t_i)\left(\boldsymbol{B}_{t_{i+1}}-\boldsymbol{B}_{t_i}\right).$$

1.6.2 Links between the two notations

If $\boldsymbol{\Sigma}$ does depend on $\boldsymbol{X}$, the two following SDEs are equivalents (Oksendal, 1998):

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{W}(t,\boldsymbol{X}_t)\mathrm{d}t+\boldsymbol{\Sigma}(t,\boldsymbol{X}_t)\circ\mathrm{d}\boldsymbol{Y}_t, \tag{1.6.1}$$

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{W}(t,\boldsymbol{X}_t)\mathrm{d}t+\frac{1}{2}\sum_{p,q=1}^{d}\left((\boldsymbol{\Sigma}_{\bullet p}\cdot\boldsymbol{\nabla})\boldsymbol{\Sigma}_{\bullet q}\right)(t,\boldsymbol{X}_t)\mathrm{d}<Y_p,Y_q>_t+\boldsymbol{\Sigma}(t,\boldsymbol{X}_t)\mathrm{d}\boldsymbol{Y}_t. \tag{1.6.2}$$

More generally, the theorem 3.2.5 page 60 of Kunita (1997) states that:

$$\mathrm{d}\boldsymbol{X}_t \triangleq \boldsymbol{\Sigma}(t)\circ\mathrm{d}\boldsymbol{Y}_t=\boldsymbol{\Sigma}(t)\mathrm{d}\boldsymbol{Y}_t+\frac{1}{2}\mathrm{d}<\boldsymbol{\Sigma},\boldsymbol{Y}>_t. \tag{1.6.3}$$

This can be reformulated as follows. Let $\boldsymbol{Z}$ be a process possibly related to $\boldsymbol{Y}$, we have:

$$\begin{aligned}\boldsymbol{W}(t,\boldsymbol{Z}_t)\mathrm{d}t+\boldsymbol{\Sigma}(t,\boldsymbol{Z}_t)\circ\mathrm{d}\boldsymbol{Y}_t &= \boldsymbol{W}(t,\boldsymbol{Z}_t)\mathrm{d}t+\frac{1}{2}\left[\mathrm{d}_t<\boldsymbol{\Sigma},\boldsymbol{Y}>_t\right](t,\boldsymbol{Z}_t)\\ &\quad+\frac{1}{2}\sum_{i,j=1}^{d}\partial_i\boldsymbol{\Sigma}_{\bullet j}(t,\boldsymbol{Z}_t)\mathrm{d}<Z_i,Y_j>_t+\boldsymbol{\Sigma}(t,\boldsymbol{Z}_t)\mathrm{d}\boldsymbol{Y}_t.\end{aligned} \tag{1.6.4}$$

Sketch of proof:

$$\int_0^T \boldsymbol{\Sigma}(t,\boldsymbol{Z}_t)\circ\mathrm{d}\boldsymbol{Y}_t = L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{T}\frac{\boldsymbol{\Sigma}\left(t_i,\boldsymbol{Z}_{t_i}\right)+\boldsymbol{\Sigma}\left(t_{i+1},\boldsymbol{Z}_{t_{i+1}}\right)}{2}\left(\boldsymbol{Y}_{t_{i+1}}-\boldsymbol{Y}_{t_i}\right), \tag{1.6.5}$$

$$= L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{T}\frac{\boldsymbol{\Sigma}\left(t_{i+1},\boldsymbol{Z}_{t_{i+1}}\right)-\boldsymbol{\Sigma}\left(t_i,\boldsymbol{Z}_{t_i}\right)}{2}\left(\boldsymbol{Y}_{t_{i+1}}-\boldsymbol{Y}_{t_i}\right) \tag{1.6.6}$$

$$+L^2-\lim_{\Delta t\to 0}\sum_{t_i=0}^{T}\boldsymbol{\Sigma}\left(t_i,\boldsymbol{Z}_{t_i}\right)\left(\boldsymbol{Y}_{t_{i+1}}-\boldsymbol{Y}_{t_i}\right), \tag{1.6.7}$$

$$= \frac{1}{2}<\widetilde{\boldsymbol{\Sigma}},\boldsymbol{Y}>_t+\int_0^T\boldsymbol{\Sigma}(t,\boldsymbol{Z}_t)\mathrm{d}\boldsymbol{Y}_t, \tag{1.6.8}$$

where $\widetilde{\boldsymbol{\Sigma}}(t)=\boldsymbol{\Sigma}(t,Z_t)$. Its differential can be computed through the Ito-Wentzell formula (1.3.1), which yields:

$$\mathrm{d}<\widetilde{\boldsymbol{\Sigma}}_{p\bullet},\boldsymbol{Y}>_t = \mathrm{d}_t<\boldsymbol{\Sigma}_{p\bullet},\boldsymbol{Y}>_t+\sum_{p,q=1}^{d}\left(\boldsymbol{\nabla}\Sigma_{pq}\right)^T\mathrm{d}<\boldsymbol{Z},Y_q>_t. \tag{1.6.9}$$

Re-injecting this expression in the above integral equation yields the result.

Therefore, depending on the needs, one can use the Itō or the Stratonovich formalism alternatively.

1.6.3 Advantages of each setups

The Stratonovich integral possesses properties which are advantageous for some issues. First, if we measure $\mathrm{d}\boldsymbol{X}_t = \boldsymbol{W}(t, \boldsymbol{X}_t)\mathrm{d}t + \boldsymbol{\Sigma}(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{B}_t$ by $\frac{\boldsymbol{X}_{t_{i+1}} - \boldsymbol{X}_{t_{i-1}}}{2}$, the estimated drift may be interpreted, as a time centered estimate, with the Stratonovich convention. Moreover, if the noise or the fast dynamic of deterministic or stochastic equation is a continuous process with infinitesimal correlation time, the solution of this SDE generally converges to the solution of another SDE. In this other SDE, the noise term is often replaced by a Stratonovich integral (Arnold, 1974). Nevertheless, it is not always true for nonlinear dynamics. Sometimes, it has to be understood in the sense of Ito or Marcus stochastic integrals. In the last case, the noise is a Levy process (Gottwald and Melbourne, 2013; Gottwald et al., 2015). The Stratonovich framework involves formulae close to the deterministic differential calculus. For instance, the Itō-Wentzell formula with Stratonovich integrals reads (theorem 3.3.2 Kunita, 1997):

$$\mathrm{d}\left(f(t, \boldsymbol{g}(t, \boldsymbol{y}))\right) \quad = \quad f(\circ\, \mathrm{d}t, \boldsymbol{g}(t, \boldsymbol{y})) + (\boldsymbol{\nabla} f)^T \circ \mathrm{d}\boldsymbol{g}, \tag{1.6.10}$$

where $f(\circ\, \mathrm{d}t, \boldsymbol{g}(t, \boldsymbol{y})) = f_3(t, \boldsymbol{g}(t, \boldsymbol{y}))\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{g}(t, \boldsymbol{y})) \circ \mathrm{d}\boldsymbol{B}_t$ if $\mathrm{d}f(t, \boldsymbol{x}) = f_3(t, \boldsymbol{x})\mathrm{d}t + \boldsymbol{f}_2(t, \boldsymbol{x}) \circ \mathrm{d}\boldsymbol{B}_t$ (Chow, 2014) and $(\boldsymbol{\nabla} f)^T \circ \mathrm{d}\boldsymbol{g} = (\boldsymbol{\nabla} f)^T \boldsymbol{g}_3(t, \boldsymbol{y})\mathrm{d}t + (\boldsymbol{\nabla} f)^T \boldsymbol{g}_2(t, \boldsymbol{y}) \circ \mathrm{d}\boldsymbol{B}_t$ if $\mathrm{d}\boldsymbol{g}(t, \boldsymbol{y}) = \boldsymbol{g}_3(t, \boldsymbol{y})\mathrm{d}t + \boldsymbol{g}_2(t, \boldsymbol{y}) \circ \mathrm{d}\boldsymbol{B}_t$. Nonetheless, these notations need to be handled very carefully. Indeed, $f(\circ\, \mathrm{d}t, \boldsymbol{g}(t, \boldsymbol{y}))$ has to be used with Stratonovich notations for f and g whereas $f(\mathrm{d}t, \boldsymbol{g}(t, \boldsymbol{y}))$ has to be used with Itō notations for f and g.

The Itō formalism provides also several advantages. When it comes to the numerical simulation of a SDE, only this formalism can be used. Moreover, in Itō SDEs, the "noise terms" $f(t, \boldsymbol{X}_t)\mathrm{d}\boldsymbol{B}_t$ are always centered, since it is a martingale time increment. Conversely, $f(t, \boldsymbol{X}_t) \circ \mathrm{d}\boldsymbol{B}_t$ is not always centered. For instance, $B_t \circ \mathrm{d}B_t = \frac{1}{2}\mathrm{d}t + B_t \mathrm{d}B_t$ (by (1.6.3)) is not centered. The Itō notation more explicitly identifies and separates the "real noise terms" and the other effects induced by randomization of equations (e.g. diffusion, noise-induced drift, ...). An other consequence is a much easier derivation of the evolution law of moments as illustrated in the next chapter. Finally, some measurements like Particle Image Velocimetry (PIV) often estimate velocity with $\boldsymbol{X}_{t_{i+1}} - \boldsymbol{X}_{t_i}$ leading to an Itō convention.

In the following, we often refer to the Itō drift of the fluid flow as "the large-scale component". Moreover, as explained in the next chapters, we assume that the large-scale conserved momentum involved this Itō drift. Although these physical interpretations are debatable, we chose to rely on Itō formalism for the reasons evoked above. Nevertheless, it remains a choice.

1.7 Generalization to Markovian processes

To conclude this section, we present briefly some formulae which can be seen as a generalization of some tools of stochastic calculus. We also show good practices to describe random functions of space and time. This point is also interesting for this thesis as we intensively deal with such functions in the next chapters. The following comes from the statistical physics work of Klyatskin (2005).

Even though the stochastic calculus is a very powerful tool, it deals only with semimartingales. If we do not consider jumps (as we have done so far), the associated infinitesimal generators (1.3.3) only involve second order terms. In a general analysis of Markovian processes, the infinitesimal generator can involve an infinite number of orders.

1.7.1 Analog of the Fokker-Planck equation: one-time density

If $t \to X(t)$ is Markovian and $f_X \triangleq \mathbb{E}\{\delta(X(t)-x)\}$ denotes its one-time probability density function, then:

$$\partial_t f_X = \mathcal{L}^* f_X, \tag{1.7.1}$$

$$\text{with } \forall f, (\mathcal{L}^* f)(\boldsymbol{x},t) \triangleq \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{\partial^n}{\partial x^n} \left(B_n(x,t) f(x)\right), \tag{1.7.2}$$

$$\text{and } B_n(x,t) \triangleq \lim_{\Delta t \to 0} \mathbb{E}\left\{ \frac{(X(t+\Delta t)-X(t))^n}{\Delta t} \middle| X(t) \right\}. \tag{1.7.3}$$

$\mathcal{L}^*$ is the adjoint of the infinitesimal generator $\mathcal{L} = \sum_{n=1}^{\infty} \frac{B_n}{n!} \frac{\partial^n}{\partial x^n}$. In stochastic calculus, $B_n = 0$, for all $n > 2$, and we retrieve the Fokker-Planck equation (1.3.5).

1.7.2 Multi-time density

To study a joint probability distribution of $\{X(t')|t' \leqslant t\}$, we do not rely on the Fokker-Planck equation. Instead, we describe the evolution of the functional

$$\Psi_X[t, x_0; x] \triangleq \mathbb{E}\left\{ \delta\left(X(t)-x_0\right) \exp\left(i \int_0^t \mathrm{d}t' \; x(t') X(t') \right) \right\}, \tag{1.7.4}$$

which is a functional of any function x, with the following equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi = (i x_0 x(t) + \mathcal{L}^*)\Psi. \tag{1.7.5}$$

Then, we retrieve the multi-time characteristic functional of interest

$$\Phi_X[t; x] \triangleq \mathbb{E}\left\{ \exp\left(i \int_0^t \mathrm{d}t' \; x(t') X(t') \right) \right\}, \tag{1.7.6}$$

by integrating Ψ:

$$\Phi_X[t; x] = \int_{\mathbb{R}} \mathrm{d}x_0 \; \Psi_X[t, x_0; x]. \tag{1.7.7}$$

Note that the time integral of (1.7.6) ends at t for causality reasons. The finite-dimensional law of X is exhaustively described by Ψ, Indeed, the Fourier transform of $p_{X(t_1),\ldots X(t_n)}(x_1, ..., x_n)$ evaluated at the point $(f_1, ..., f_n)$ is:

$$\mathbb{E}\left\{ \exp\left(i \sum_{k=1}^{n} f_k X(t_k) \right) \right\} = \left[\exp\left(\sum_{k=1}^{n} f_k \frac{\delta}{\delta x(t_k)} \right) \Psi_X[t; x] \right]_{x=0}, \tag{1.7.8}$$

where the exponential of the variational differential operator in the right-hand side is defined by the Taylor series of the exponential. Note that the Girsanov theorem 1.3.3 is another way of describing the law of $\{X(t')|t' \leqslant t\}$.

1.7.3 Analog of the mean of Itō formula

We will denote by f, a smooth enough function and $R[t; X(t')]$, a functional that depends both on time t and $\{X(t')|0 \leqslant t' \leqslant t\}$. From (1.7.1), one can show that:

$$\frac{\partial}{\partial t}\mathbb{E}\left\{ f(X(t)) R[t; X(t')] \right\} = \mathbb{E}\left\{ f(X(t)) \frac{\partial}{\partial t} R[t; X(t')] \right\} + \mathbb{E}\left\{ (\mathcal{L}f)(X(t)) R[t; X(t')] \right\}. \tag{1.7.9}$$

This equation can be seen as a generalization of the Itō formula applied to $f(X(t))$. However, the result is only expressed through a mean.

1.7.4 Multi-point density

When $X = X(t, \boldsymbol{y})$ is a function of $\boldsymbol{y} \in \Omega$, the operators $\mathcal{L}$ and $\mathcal{L}^*$ involve variational derivatives $\frac{\delta}{\delta x(\boldsymbol{y})}$ instead of classical derivatives $\frac{\partial}{\partial x}$. This kind of equation is called a Hopf equation instead of a partial differential equation. When the spatial derivatives of the stochastic evolution law have orders higher than 1, there are no closed form for the indicator function $\psi_X(\boldsymbol{y}, x_0, t) = \delta(X(t, \boldsymbol{y}) - x_0)$ and accordingly no closed form for its mean: the one-time one-point probability density function. Hence, we do not uses the Fokker-Planck equation to describe a probability density function. Instead, we use a Hopf equation to describe the one-time multi-point characteristic functional:

$$\Phi_X[t; x] = \mathbb{E}\left\{\exp\left(i \int_\Omega \mathrm{d}\boldsymbol{y}\ x(\boldsymbol{y}) X(t, \boldsymbol{y})\right)\right\}, \tag{1.7.10}$$

which is a functional of any deterministic function x. The Hopf equation which describes Φ_X generally involves an infinite number of variational derivatives and the mutli-point multi-time cumulants of the random forcing terms.

In this section, we have exposed the main mathematical tools that will be used within this study. Definitions, properties and theorems have been briefly explained. After some examples of SDEs, estimation methods based on stochastic calculus have been developed. Then, the choice between Stratonovich and Itō calculus has been discussed. Finally, we have presented briefly more general tools to deal with Markovian processes in finite and infinite dimensions which are not semimartingales. In the next section of the thesis, we will apply these tools to fluid dynamics equations.

Chapter 2

Geophysical flows under location uncertainty, Part I Random transport and general models

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part I: Random transport and general models. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017a

Abstract

A stochastic flow representation is considered with the Eulerian velocity decomposed between a smooth large scale component and a rough small-scale turbulent component. The latter is specified as a random field uncorrelated in time. Subsequently, the material derivative is modified and leads to a stochastic version of the material derivative to include a drift correction, an inhomogeneous and anisotropic diffusion, and a multiplicative noise. As derived, this stochastic transport exhibits a remarkable energy conservation property for any realizations. As demonstrated, this pivotal operator further provides elegant means to derive stochastic formulations of classical representations of geophysical flow dynamics.
Keywords: stochastic flows, uncertainty quantification, ensemble forecasts, upper ocean dynamics

2.1 Introduction

Despite the increasing power of computational resources and the availability of high quality observations, a precise description of geophysical flows over their whole dynamical scales is today completely beyond reach. Challenges appear as unlimited as the variety of dynamics and boundary conditions with their broad range of spatial and temporal scales across the globe. To face these challenges, numerous efforts are taking place to build an ever-increasing quality, quantity, duration and integration of all observations, in situ and satellite. In parallel, simulation capabilities largely improved, *i.e.*, analysis can now be routinely carried out to more precisely characterize the variability in the global ocean, at scales of ten to hundreds of kilometers and one to hundreds of days. Yet, for these ocean models, the unresolved small scales and associated fluxes are always accounted for by simple mathematical models, *i.e.* parameterizations.

Although the development of more efficient sub-grid representations remains a very active research area, the possible separation between relatively low-frequency, large scale patterns and transient, small-scale fluctuations, strongly invites to consider stochastic representations of the geophysical dynamics (e.g. Hasselmann, 1976; Allen and Stainforth, 2002; Penland, 2003b; Berner et al.,

2011; Franzke et al., 2015). As derived, such developments are meant to better describe the system's variability, especially including a mean drift , called "bolus" velocity (Gent and Mcwilliams, 1990) or skew-diffusion (Nakamura, 2001; Vallis, 2006) in oceanography, and noise-induced velocity in climate sciences.

In that context, several different strategies have been proposed (Franzke et al., 2015). Among them, techniques motivated by physics have been devised. Those schemes aim to overcome a bad representation of the small scale forcing and of their interactions with the large scale processes. Two of such schemes have been carried out at ECMWF. The first one, the stochastic perturbation of the physical tendencies – SPPT – (Buizza et al., 1999) implements a multiplicative random perturbation of parameterized physical tendencies. The random variables involved are correlated in space and time, and their characteristics set from fine grid simulations. The second one, the stochastic kinetic-energy backscatter – SKEB – (Shutts, 2005) introduces a perturbation of the stream function and potential temperature. This scheme is based on earlier works on energy backscattering modelling through the introduction of random variables (Mason and Thomson, 1992). Numerous works showed a beneficial impact of the injected randomness on weather and climate forecasts mean and variability (see (Berner et al., 2015) and references therein) or in oceanography (Brankart, 2013; Mana and Zanna, 2014). However, the amplitude of the perturbations to apply is difficult to specify. The non-conservative and the variance-creating nature of those schemes is also problematic in that prospect. A too large amplitude, while increasing significantly the ensemble spread, may lead to unstable schemes for simulations that go beyond short-term forecast applications. A balance between the large-scale sub-grid diffusive tensor and the noise amplitude must thus be found to stabilize the system.

Also based on a separation of the state variables between slow and fast components, a mathematical framework – refereed to as MTV algorithms – has been proposed to derive stochastic reduced-order dynamical systems for weather and climate modelling (Franzke et al., 2005; Franzke and Majda, 2006; Majda et al., 1999, 2001, 2003). Considering a linear stochastic equation to describe the fast modes, derivations have been rigorously studied (Gottwald and Melbourne, 2013; Melbourne and Stuart, 2011; Pavliotis and Stuart, 2008). As demonstrated, the continuous fast dynamics converges in continuous time towards a Stratonovich noise, leading to a diffusion term when expressed in a corresponding Ito stochastic integral form.

As well, stochastic superparametrization assumes a scale separation (Grooms and Majda, 2013, 2014). The point approximation and Reynolds decompositions replace homogenization techniques. As for MTV methods, the small-scale evolution law is linearized and corrected with the introduction of noise and damping terms. The second order moments of the solution are then known analytically and can feed the sub-grid tensors expression of the mean deterministic large-scale evolution law. For such developments, the direct use of the Reynolds decomposition implicitly assumes that small-scale components are differentiable. This theoretically prevents the use of Langevin type equations for the small-scale evolution. Furthermore, in such a derivation, each scalar evolution law involves a different sub-grid tensor. Similarly to the definition of eddy viscosity and diffusivity models for Large-Eddy simulation, the noise expression of most stochastic fluid dynamic models are hardly inferred from physics. So, instantaneous diffusion and randomness may not be consistently related; even though some careful parametrizations of stationary energy fluxes couple them (Grooms and Majda, 2013; Sapsis and Majda, 2013b; Grooms and Majda, 2014; Sapsis and Majda, 2013c).

To overcome these difficulties, we propose to dwell on a different strategy. As previously initiated (Mémin, 2014), the large-scale dynamics is not prescribed from a deterministic representation of the system's dynamics. Instead, a random variable, referred to as location uncertainty, is added to the Lagrangian expression of the flow. The resulting Eulerian expression then provides stochastic extensions of the material derivative and of the Reynolds transport theorem. An explicit expression of a noise-induced drift is further obtained. As also derived, a sub-grid stress tensor, describing the small-scale action on the large scales, does not resort to the usual Boussinesq eddy viscosity assumption, and further, consistently appears throughout all the conservation equations of the system. Moreover, the advection by the unresolved velocity acts as a random forcing. As such, this framework provides a direct way to link the resulting material transport and the underlying dynamics. The well-posedness of these equations has been studied by Mikulevicius and Rozovskii (2004) and Flandoli (2011). Recently, Holm (2015) derived similar evolution laws from the inviscid and adiabatic framework of Lagrangian mechanics. Compared to models under location uncer-

tainty, the stochastic transport of scalars is identical. However, the momentum evolution of Holm (2015) involves an additional term which imposes the helicity conservation but may increase the kinetic energy.

Starting with the description of the transport under location uncertainty (section 2), developments are then carried out to explore this stochastic framework for different classical geophysical dynamical models (section 3).

2.2 Transport under location uncertainty

2.2.1 A 2-scale random advection-diffusion description

As often stated, ocean and atmospheric dynamics can be assumed to be split into two contributions with very distinct correlation times. This assumption can especially hold for the top layer of the ocean. For example, the larger ocean geostrophic component generally varies on much slower time scales than motions at smaller spatial scales. From an observational perspective, current generation satellite altimeter instruments are capable of resolving only the largest eddy scales, and the measurements can depend sensitively on the local kinetic energy spectrum of the unresolved flow (Poje et al., 2010; Keating et al., 2011). Satellite observations of the upper-ocean velocity field at higher resolution can also be obtained (e.g. Chapron et al., 2005) but are certainly too sparse and possibly noisy.

Accordingly, without loss of generality, observations of an instantaneous Eulerian velocity field are likely coarse-grained in time, and can be interpreted under a 2-scale framework. As such, the instantaneous Eulerian velocity is decomposed between a well resolved smooth component, denoted $\boldsymbol{w}$, continuous in time, and a rough small-scale one, rapidly decorrelating in time. This badly-resolved contribution, expressed as $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is then assumed Gaussian, correlated in space, but uncorrelated in time. This contribution can be inhomogeneous and anisotropic in space. Due to the irregularity of the flow, the transport of a conserved quantity, Θ, by the whole velocity, defined as

$$\Theta(\boldsymbol{X}_{t+\Delta t}, t+\Delta t) \quad = \quad \Theta(\boldsymbol{X}_t, t) \tag{2.2.1}$$

corresponds to a random mapping. In this setup the large-scale velocity possibly depends on the past history of the small-scale component. This latter being white in time, the two components are uncorrelated. Hence, the above conservation shall lead to a classical advection-diffusion evolution, with the introduction of an inhomogeneous and anisotropic diffusion coefficient matrix, $\boldsymbol{a}$, solely defined by the one-point one-time covariance of the unresolved displacement per unit of time:

$$\boldsymbol{a} = \frac{\mathbb{E}\left\{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}}{\mathrm{d}t}. \tag{2.2.2}$$

The inhomogeneous structure of the small-scale variance motions shall create inhomogeneous spreading rates. More agitated fluid parcels spread faster than those over quiescent regions. Overall, the latter can be seen as "attracting" the large-scale gradients. This effect leads to invoke a drift correction, anti-correlated with the variance gradient, or, in a multi-dimensional point of view, anti-correlated with the covariance matrix divergence. Accordingly, the random advection under a 2-scale description can be expected to be expressed as:

$$\partial_t\Theta + \underbrace{\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}\Theta}_{\text{Corrected advection}} = \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}} - \underbrace{\boldsymbol{\sigma}\dot{\boldsymbol{B}}\cdot\boldsymbol{\nabla}\Theta}_{\text{Random forcing}}, \tag{2.2.3}$$

with a modified velocity given by

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T. \tag{2.2.4}$$

We note the conserved quantity is diffused by the small-scale random velocity. The random forcing expresses the advection by the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$, and continuously backscatters

random energy to the system. Because of this white-noise forcing term, the Eulerian conservation equation (2.2.3) (that will be formally expressed in the following sections) intrinsically concerns a random non-differentiable tracer. Finally, the conserved quantity is also advected by an "effective" velocity, $\boldsymbol{w}^\star$, taking into account the possible spatial variation of the small-scale velocity variance, as well as the possible divergence of this velocity component.

Considering the unresolved velocity and this effective drift, $\boldsymbol{w}^\star$, divergent-free, we shall see that this 2-scale development establishes an exact balance between the amount of diffusion and the random forcing. Subsequently, essential properties related to energy conservation and mean/variance tracer evolution directly result from this balance.

2.2.2 Uncertainty formalism

In a Lagrangian stochastic form, the infinitesimal displacement associated with a particle trajectory $\boldsymbol{X}_t$ is:

$$\mathrm{d}\boldsymbol{X}_t \quad = \quad \boldsymbol{w}(\boldsymbol{X}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t,t)\mathrm{d}\boldsymbol{B}_t. \tag{2.2.5}$$

Formally, this is defined over the fluid domain, Ω, from a d-dimensional Brownian function $\boldsymbol{B}_t$. Such a function can be interpreted as a white noise process in space and a Brownian process in time[1]. The time derivative of the Brownian function, in a distribution sense, is denoted $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$, and is a white noise distribution. The spatial correlations of the flow uncertainty are specified through the diffusion operator $\boldsymbol{\sigma}(.,t)$, defined for any vectorial function, $\boldsymbol{f}$, through the matrix kernel $\breve{\boldsymbol{\sigma}}(.,.,t)$:

$$\boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{f} \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z},t)\boldsymbol{f}(\boldsymbol{z},t)\mathrm{d}\boldsymbol{z}. \tag{2.2.6}$$

This quantity is assumed to have a finite norm[2] and to have a null boundary condition on the domain frontier[3]. The resulting d-dimensional random field, $\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t$, is a centered vectorial Gaussian function, correlated in space and uncorrelated in time with covariance tensor:

$$\boldsymbol{Cov}(\boldsymbol{x},\boldsymbol{y},t,t') \quad \triangleq \quad \mathbb{E}\left\{\left(\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\sigma}(\boldsymbol{y},t')\mathrm{d}\boldsymbol{B}_{t'}\right)^T\right\}, \tag{2.2.7}$$

$$= \quad \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z},t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y},\boldsymbol{z},t)\mathrm{d}\boldsymbol{z}\ \delta(t-t')\mathrm{d}t. \tag{2.2.8}$$

For sake of thoroughness, the uncertainty random field has a (mean) bounded norm[4]: $\mathbb{E}\|\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\|^2_{L^2(\Omega)} < \infty$ for any bounded time $t \leqslant T < \infty$. Hereafter, the diagonal of the covariance tensor, $\boldsymbol{a}$, will be referred to as the variance tensor:

$$\boldsymbol{a}(\boldsymbol{x},t)\delta(t-t')\mathrm{d}t = \boldsymbol{Cov}(\boldsymbol{x},\boldsymbol{x},t,t').$$

By definition, it is a symmetric positive definite matrix at all spatial points, $\boldsymbol{x}$. This quantity, also denoted $\boldsymbol{\sigma}\boldsymbol{\sigma}^T$, corresponds to the time derivative of the so-called quadratic variation process:

$$\boldsymbol{\sigma}\boldsymbol{\sigma}^T \triangleq \boldsymbol{a} = \partial_t\left\langle \int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_s, \left(\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_r\right)^T\right\rangle.$$

with $\langle f,g\rangle$ to stand for the quadratic cross-variation process of f and g (see Appendix 2.A).

Given this strictly defined flow, the corresponding material derivative expression of a given quantity can be introduced.

[1]Formally it is a cylindrical I_d-Wiener process (see Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007) for more information on infinite dimensional Wiener process and cylindrical I_d-Wiener process).

[2]More precisely, the operator $\boldsymbol{\sigma}$ is assumed to be Hilbert-Schmidt.

[3]Note that periodic boundary conditions can also be envisaged.

[4] This norm is finite since $\boldsymbol{\sigma}$ is Hilbert-Schmidt, ensuring the boundness of the trace of operator Q – defined by the kernel $(\boldsymbol{x},\boldsymbol{y}) \mapsto \boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{\sigma}^T(\boldsymbol{y},t)$ –, and $\forall t \leqslant T < \infty$, $\mathbb{E}\|\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\|^2_{L^2(\Omega)} = \int_0^t\int_\Omega \|\breve{\boldsymbol{\sigma}}(\bullet,\boldsymbol{z})\|^2_{L^2(\Omega)}\mathrm{d}\boldsymbol{z}\mathrm{d}t' = \int_0^t \|\boldsymbol{\sigma}\|^2_{HS,L^2(\Omega)}\mathrm{d}t' = \int_0^t tr(\boldsymbol{Q})\mathrm{d}t' < \infty$, where the index HS refers to the Hilbert-Schmidt norm.

2.2.3 Material derivative

To derive the expression of the material derivative $D_t\Theta \triangleq \left(\mathrm{d}\left(\Theta\left(\boldsymbol{X}_t,t\right)\right)\right)_{|\boldsymbol{X}_t=x}$, also quoted as the Ito-Wentzell derivative or generalized Ito derivative in a stochastic flow context (Kunita, 1997, theorem 3.2.2), let us introduce an operator, hereafter referred to as the stochastic transport operator:

$$\mathbb{D}_t\Theta \quad \triangleq \quad \underbrace{\mathrm{d}_t\Theta}_{\substack{\triangleq\ \Theta(\boldsymbol{x},t+\mathrm{d}t)-\Theta(\boldsymbol{x},t)\\ \text{Time increment}}} + \underbrace{\left(\boldsymbol{w}^\star\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}}\mathrm{d}t \tag{2.2.9}$$

This operator corresponds to a strict formulation of (2.2.3). More specifically, it involves a time increment term $\mathrm{d}_t\Theta$ instead of a partial time derivative as Θ is non differentiable. Contrary to the material derivative, the transport operator has an explicit expression (equation (2.2.9)). However, the material derivative is explicitly related to the transport operator (see proof in Appendix 2.B)

$$\left\{\begin{array}{lcl}\mathbb{D}_t\Theta &=& f_1\mathrm{d}t+\boldsymbol{h}_1^T\mathrm{d}\boldsymbol{B}_t,\\ D_t\Theta &=& f_2\mathrm{d}t+\boldsymbol{h}_2^T\mathrm{d}\boldsymbol{B}_t,\end{array}\right. \iff \left\{\begin{array}{lcl} f_2 &=& f_1+\mathrm{tr}\big(\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)\boldsymbol{h}_1^T\big),\\ \boldsymbol{h}_1 &=& \boldsymbol{h}_2.\end{array}\right. \tag{2.2.10}$$

Note, the material derivative, D_t, has a clear physical meaning but no explicit expression whereas the explicit expression of the transport operator offers elegant means to derive stochastic Eulerian evolution laws. Most often both operators coincide and can interchangeably be used. As a matter of fact, in most cases, we deal with null Brownian function $\boldsymbol{h}_1$ in (2.2.10). This corresponds, for instance, either to the transport of a scalar $\mathbb{D}_t\Theta=0$ or to the conservation of an extensive property $\left(\int_{\mathcal{V}(t)} q\right)$ when the unresolved velocity component is solenoidal ($\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t=0$), which leads, as we will see it, to $D_tq=-\boldsymbol{\nabla}\cdot\boldsymbol{w}^*q\mathrm{d}t$ ((2.2.28)). In such a case, it is straightforward to infer from the system (2.2.10), that $\mathbb{D}_t$ and D_t coincide. For this precise case, those operators lead to

$$\mathbb{D}_t\Theta(\boldsymbol{X}_t,t)=D_t\Theta(\boldsymbol{X}_t,t)=\mathrm{d}\left(\Theta(\boldsymbol{X}_t,t)\right)=f_1(\boldsymbol{X}_t,t)\mathrm{d}t. \tag{2.2.11}$$

Going back to the Eulerian space, the classical calculus rules apply to operator $\mathbb{D}_t$, e.g. the product rule

$$\mathbb{D}_t(fg)(\boldsymbol{x},t)=\left(\mathbb{D}_tf\ g+f\ \mathbb{D}_tg\right)(\boldsymbol{x},t), \tag{2.2.12}$$

and the chain rule:

$$\mathbb{D}_t\big(\varphi\circ f\big)(\boldsymbol{x},t)=\mathbb{D}_tf(\boldsymbol{x},t)(\varphi'\circ f)(\boldsymbol{x},t). \tag{2.2.13}$$

Given these properties, an expression for the stochastic advection of a scalar quantity can be derived.

2.2.4 Scalar advection

The advection of a scalar Θ thus reads:

$$\mathbb{D}_t\Theta=D_t\Theta=0. \tag{2.2.14}$$

To analyze this stochastic transport equation, let us first consider that the effective drift and the unresolved velocity are both divergence-free. As shown later, these conditions ensure an isochoric stochastic flow (see (2.2.33)). With these conditions, the stochastic transport equation exhibits remarkable conservation properties.

Energy conservation

From (2.2.9-2.2.14) and Ito lemma, the scalar energy evolution is given by:

$$\mathrm{d}\int_\Omega\frac{1}{2}\Theta^2=\int_\Omega\left(\Theta\mathrm{d}_t\Theta+\frac{1}{2}\mathrm{d}_t\langle\Theta,\Theta\rangle\right), \tag{2.2.15}$$

$$=-\int_\Omega\frac{1}{2}\left(\boldsymbol{w}^*\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\left(\Theta^2\right)+\underbrace{\int_\Omega\Theta\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)\mathrm{d}t}_{\text{Loss by diffusion}}+\underbrace{\int_\Omega\frac{1}{2}\left(\boldsymbol{\nabla}\Theta\right)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\mathrm{d}t}_{\text{Energy intake from noise}}. \tag{2.2.16}$$

For suitable boundary conditions, the two last terms cancel out after integration by part. The diffused energy is thus exactly compensated by the energy brought by the noise. With divergent-free conditions for $\boldsymbol{w}^\star$ and $\boldsymbol{\sigma}$, another integration by part gives

$$\mathrm{d}\int_\Omega \frac{1}{2}\Theta^2 \quad = \quad \int_\Omega \frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{w}^*\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\Theta^2 = 0. \tag{2.2.17}$$

The energy is thus conserved for all scalar random realizations. The expectation of the energy – the energy (ensemble) mean – is therefore also conserved. Moreover, from the decomposition $\Theta = \mathbb{E}(\Theta) + \big(\Theta - \mathbb{E}(\Theta)\big)$ into the mean and the random anomaly component, we obtain a partition of this constant energy mean:

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\|\Theta\|^2_{\mathcal{L}^2(\Omega)} = \frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{E}(\Theta)\|^2_{\mathcal{L}^2(\Omega)} + \frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega Var(\Theta). \tag{2.2.18}$$

A decrease of the mean energy – the energy of the (ensemble) mean – is always associated with an (ensemble) variance increase. Similar energy mean budgets have recently been discussed by several authors. Majda (2015) refers to this energy mean as the statistical energy. The author derives the evolution law of this energy by adding the evolution equations of the mean energy and of the integrated variance, whereas our energy budget is obtained by evaluating the mean of the evolution law of the total energy, $\|\Theta\|^2_{\mathcal{L}^2(\Omega)}$. However, Majda (2015) does not specify the random forcing. This is why the latter does not *a priori* balance the turbulent diffusion. Farrell and Ioannou (2014) also studied the energy mean of stochastic fluid dynamics systems especially under quasi-linear approximations and with an additive Gaussian forcing.

By the chain rule, all the tracer moments are also conserved:

$$\mathbb{D}_t\Theta^p = p\ \Theta^{p-1}\mathbb{D}_t\Theta = 0. \tag{2.2.19}$$

Yet, the energy of statistical moments are in general not conserved, as detailed in the following section.

Mean and variance fields of a passive scalar

Consider now that the expectation corresponds to a conditional expectation given the effective drift. This applies to passive scalar transport for which the drift does not depend on the tracer. Terms in $\mathrm{d}\boldsymbol{B}_t$ have zero-mean, and the mean passive scalar evolution can be immediately derived taking the conditional expectation of the stochastic transport:

$$\partial_t\mathbb{E}(\Theta) + \underbrace{\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}\mathbb{E}(\Theta)}_{\text{Advection}} = \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\mathbb{E}(\Theta)\right)}_{\text{Diffusion}}. \tag{2.2.20}$$

Since $\boldsymbol{w}^*$ is divergent-free, it has no influence on the energy budget. The mean field energy decreases with time due to diffusion. As for the variance, its evolution equation, derived in Appendix 2.C, reads:

$$\partial_t Var(\Theta) + \underbrace{\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}Var(\Theta)}_{\text{Advection}} = \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}Var(\Theta)\right)}_{\text{Diffusion}} + \underbrace{\left(\boldsymbol{\nabla}\mathbb{E}(\Theta)\right)^T\boldsymbol{a}\boldsymbol{\nabla}\mathbb{E}(\Theta)}_{\text{Variance intake}}. \tag{2.2.21}$$

This is also an advection-diffusion equation, with an additional source term. Integrating this equation on the whole domain, with the divergent-free condition, and considering the divergence form of the first right-hand term, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega Var(\Theta) = \int_\Omega \left(\boldsymbol{\nabla}\mathbb{E}(\Theta)\right)^T\boldsymbol{a}\boldsymbol{\nabla}\mathbb{E}(\Theta) \geqslant 0. \tag{2.2.22}$$

It shows that the stochastic transport of a passive scalar creates variance. The dissipation that occurs in the mean-field energy equation is exactly compensated by a variance increase. This

mechanism is very relevant for ensemble-based simulations. The uncertainty modeling directly incorporates a large-scale dissipating sub-grid tensor, and further encompasses a variance increase mechanism to balance the total energy dissipation. Such a mechanism is absent in ensemble-based data assimilation development (Berner et al., 2011; Gottwald and Harlim, 2013; Snyder et al., 2015). An artificial inflation of the ensemble variance is usually required in consequence to avoid filter divergence (Anderson and Anderson, 1999).

Active tracers

For the more general case of an active tracer, the velocity depends on the tracer distribution, additional energy transfers occurs between the mean and the random tracer components (Sapsis, 2013; Sapsis and Majda, 2013b,c; Ueckermann et al., 2013; Majda, 2015). Though a complete analytical description is involved, these energy transfers are mainly due to the nonlinearity of the flow dynamics, and are hence more familiar. The models under location uncertainty involve both types of interactions: the "usual" nonlinear interactions and the random energy transfers previously described. As such, these two energy fluxes analyzes are complementary. In deterministic turbulent dynamics with random initial conditions, energy is drained from the mean tracer toward several modes (*e.g.* Fourier modes) of the tracer random component, and is backscattered from other modes. The energy fluxes toward (from) random modes increases (decreases) the variance. In the case of the deterministic Navier-Stokes equations, Sapsis (2013) analytically expressed the integrated variance. The molecular or turbulent diffusion decreases the variance whereas the mean velocity may increases or decreases the random energy, by triad interactions. The modes receiving energy become unstable, whereas those giving energy are over-stabilized (Sapsis and Majda, 2013b). In ensemble data assimilation of large-scale geophysical flows, the solution is defined by a manifold sampled by a small ensemble of realizations. Those stabilizations and destabilizations are the reason for the alignment of ensembles along unstable directions (Trevisan and Uboldi, 2004; Ng et al., 2011). It can lead to filter divergence (Gottwald and Harlim, 2013; Bocquet et al., 2016). In the absence of any modes truncation, the nonlinear interactions redistribute the energy between those modes. Otherwise, the missing energy fluxes can be parametrized with additional random terms (Sapsis and Majda, 2013b,c).

To further describe the energy exchanges involved in the dynamics under location uncertainty of active tracers, we introduce the decomposition $\Theta = \widetilde{\Theta} + \Theta'$ in terms of a slow component $\widetilde{\Theta}$ and a highly oscillating component Θ'. The first one is time-differentiable whereas the second is only continuous with respect to time. Both components are random. This decomposition, the so-called semi-martingale decomposition, is unique (Kunita, 1997). For each component, the following coupled system of transport equations is:

$$\partial_t \widetilde{\Theta} + \boldsymbol{w}^\star \cdot \boldsymbol{\nabla}\Theta = \boldsymbol{\nabla} \cdot \left(\frac{1}{2} \boldsymbol{a} \boldsymbol{\nabla}\Theta \right), \tag{2.2.23}$$

$$\mathrm{d}_t \Theta' + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t \cdot \boldsymbol{\nabla}\Theta = 0. \tag{2.2.24}$$

At the initial time, the first component is deterministic (given the initial conditions) and the second one is zero. The large-scale component becomes random through the oscillating component, which is characterized by a gradually increasing energy along time:

$$\mathbb{E}\|\Theta'\|^2_{L^2(\Omega)} = \mathbb{E} \int_\Omega \langle \Theta', \Theta' \rangle = \mathbb{E} \int_0^t \int_\Omega (\boldsymbol{\nabla}\Theta)^T \boldsymbol{a} \boldsymbol{\nabla}\Theta \, \mathrm{d}t \geqslant 0. \tag{2.2.25}$$

Note, the expectation is taken with respect to the law of the Brownian path. The energy mean of the non-differentiable component Θ' is the mean of the energy intake provided by the noise (2.2.16). The same amount of energy mean is removed from the system by the diffusion (2.2.16). Once diffused, this energy is fed back to the small-scale tracer Θ', the white noise velocity acting here as an energy bridge. Such an energy redistribution is a main issue in sub-grid modeling. Indeed, as explained above, large-scale flow simulations often miss to capture the energy fluxes between the mean and the random components but also the energy redistribution from the unstable modes to the stable modes. Note that, even though the two components are orthogonal as functions of time (in a precise sense), they are not, in general, as functions of space: $\int_\Omega \widetilde{\Theta}\Theta' \neq 0$. In particular, it can be shown that those two components are indeed anti-correlated when the tracer is passive.

The homogeneous case and the Kraichnan model

A divergent-free isotropic random field for the small-scale velocity component corresponds to the Kraichnan model (Kraichnan, 1968, 1994; Gawędzki and Kupiainen, 1995; Majda and Kramer, 1999). The variance tensor, $\boldsymbol{a}$, becomes a constant diagonal matrix $\frac{1}{d}tr(\boldsymbol{a})\mathbb{I}_d$, where d stands for the dimension of the spatial domain Ω. The tracer evolution now involves a Laplacian diffusion

$$\mathrm{d}_t\Theta + \left(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta = \frac{tr(\boldsymbol{a})}{2d}\Delta\Theta\mathrm{d}t. \tag{2.2.26}$$

Additionally, the original Kraichnan model considers a small molecular diffusion, ν, and an external Gaussian forcing, $f\mathrm{d}B'_t$, defined as an homogeneous random field uncorrelated in time and independent of the velocity component $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ (Gawędzki and Kupiainen, 1995). In our framework, the Kraichnan model, which does not involve any large-scale drift term, reads:

$$\mathrm{d}_t\Theta + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\Theta = \left(\nu + \frac{tr(\boldsymbol{a})}{2d}\right)\Delta\Theta\mathrm{d}t + f\mathrm{d}B'_t. \tag{2.2.27}$$

As compared to the original model, this derivation directly identifies the eddy diffusivity contribution, only implicitly termed in the Kraichnan model (Gawędzki and Kupiainen, 1995; Majda and Kramer, 1999). The Ito calculus further offers means to infer the evolution of the tracer moments, (2.2.20) and (2.2.21). The proposed development introduces an additional non-linearity through $\boldsymbol{w}$ and possible non-uniform turbulence conditions.

2.2.5 Transport of extensive properties

Hereafter, all fundamental conservation laws are formulated for extensive properties.

Stochastic Reynolds transport theorem

Similar to the deterministic case, the stochastic Reynolds transport theorem shall describe the time differential of a scalar function, $q(\boldsymbol{x},t)$, within a material volume, $\mathcal{V}(t)$, transported by the random flow (2.2.5):

$$\mathrm{d}\int_{\mathcal{V}(t)} q = \int_{\mathcal{V}(t)}\left[D_t q + \boldsymbol{\nabla}\cdot\left(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)q + \mathrm{d}\left\langle\int_0^t D_{t'}q, \int_0^t \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\right\rangle\right]. \tag{2.2.28}$$

This expression, rigorously derived in Appendix 2.D, was first introduced in a slightly different version by Mémin (2014). In most cases, the unresolved velocity component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is divergence-free and, the source of variations of the extensive property $\int_{\mathcal{V}(t)} q$ is time-differentiable, *i.e.* with a differential of the form $\mathrm{d}\int_{\mathcal{V}(t)} q = \mathcal{F}\mathrm{d}t$. In such a case, for an arbitrary volume, the transport theorem takes the form $D_t q = f\mathrm{d}t$, and according to equation (2.2.10) the material derivative can be replaced by the stochastic transport operator, $\mathbb{D}_t q$, to provide an intrinsic expression of this stochastic transport theorem.

Jacobian

Taking $q = 1$ characterizes the volume variations through the flow Jacobian, J:

$$\begin{aligned}
\int_{\mathcal{V}(t_0)}\mathrm{d}(J(\boldsymbol{X}_t(\boldsymbol{x}_0),t))\mathrm{d}\boldsymbol{x}_0 &= \mathrm{d}\int_{\mathcal{V}(t)}\mathrm{d}\boldsymbol{x}, && (2.2.29)\\
&= \int_{\mathcal{V}(t)}\boldsymbol{\nabla}\cdot\left(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}, && (2.2.30)\\
&= \int_{\mathcal{V}(t_0)}\left[J\boldsymbol{\nabla}\cdot\left(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\right](\boldsymbol{X}_t(\boldsymbol{x}_0),t)\,\mathrm{d}\boldsymbol{x}_0. && (2.2.31)
\end{aligned}$$

Valid for an arbitrary initial volume $\mathcal{V}(t_0)$, it leads to a familiar form for the Lagrangian flow Jacobian evolution law:

$$D_t J - J\boldsymbol{\nabla}\cdot\left(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0. \tag{2.2.32}$$

Incompressibility condition

The Jacobian evolution (2.2.32) ensures a necessary and sufficient condition for the isochoric nature of the stochastic flow:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = 0 \text{ and } \boldsymbol{\nabla} \cdot \boldsymbol{w}^* = 0. \tag{2.2.33}$$

If the large-scale flow component, $\boldsymbol{w}$, is solenoidal, this reduces to:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = 0 \text{ and } \boldsymbol{\nabla} \cdot \boldsymbol{w} = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{a})^T = 0. \tag{2.2.34}$$

Note that for an isotropic unresolved velocity, the last condition is naturally satisfied, as this unresolved velocity component is associated with a constant variance tensor, $\boldsymbol{a}$.

2.2.6 Summary

An additional Gaussian and time-uncorrelated velocity modifies the expression of the material derivative. In most cases, the resulting stochastic transport operator, $\mathbb{D}_t$, coincides with the material derivative, D_t. Yet, possible differences between $\mathbb{D}_t$ and D_t have simple analytic expressions. This stochastic transport operator leads to an Eulerian expression of the tracer transport. As obtained, the tracer is forced by a multiplicative noise and mixed by an inhomogeneous and anisotropic diffusion. Moreover, the advection drift is possibly modified with a correction term related to the spatial variation of the small-scale velocity variance. The random forcing, the dissipation and the effective drift correction are all linked. Accordingly, the energy is conserved for each realization, as the tracer energy dissipated by the diffusion term is exactly compensated by the energy associated with the random velocity forcing. For a passive tracer, the evolution laws for the mean and variance precise these energy exchanges. The unresolved velocity transfers energy from the mean part of the tracer to its random part. For an active tracer, this velocity component bears energy from the whole tracer field to its random non-differentiable component.

2.3 Stochastic versions of geophysical flow models

The stochastic version of the Reynolds transport theorem provides us the flow Jacobian evolution law, as well as the rate of change expression of any scalar quantity within a material volume. Together with the fundamental conservation laws of classical mechanics, it provides us a powerful tool to derive in a systematic way stochastic flow models. Thanks to the bridge between the material derivative and the stochastic transport operator, this derivation closely follows the usual deterministic derivations.

All along the following development, the small-scale random flow component will be assume incompressible, *i.e.* associated with a divergence-free diffusion tensor:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = 0. \tag{2.3.1}$$

This assumption remains realistic for the geophysical models considered in this study, and does not prevent the resolved velocity component (and therefore the whole field) to be compressible.

2.3.1 Mass conservation

Mass conservation for arbitrary volumes rules the stochastic transport of the fluid density, denoted ρ:

$$\mathbb{D}_t \rho + \rho \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t = 0. \tag{2.3.2}$$

A suggested in 2.2.5, the material derivative, D_t, is now replaced by $\mathbb{D}_t$, defined by Eq. (2.2.9). Indeed, the mass variation is zero and thus time-continuous, and the stochastic operator coincides with the material derivative.

2.3.2 Active scalar conservation law

The transport theorem (2.2.28) applied to a quantity $\rho\Theta$ describes the rate of change of the scalar Θ and is generally balanced by a production/dissipation term, as:

$$\mathbb{D}_t(\rho\Theta) + \rho\Theta\boldsymbol{\nabla}\cdot\boldsymbol{w}^*\mathrm{d}t = \rho\mathcal{F}_\Theta(\Theta)\mathrm{d}t. \tag{2.3.3}$$

Again, the stochastic transport operator, $\mathbb{D}_t$, is used instead of the material derivative, D_t, since the source of variation $\int_0^t \left(\int_{\mathcal{V}(t)} \rho\mathcal{F}_\Theta\right)\mathrm{d}t$ of the extensive property, $\int_{\mathcal{V}(t)} \rho\Theta$, is time-differentiable (integral in $\mathrm{d}t$), as explained in 2.2.5. Considering the product rule (2.2.12) and mass conservation (2.3.2), the transport evolution model for the scalar writes:

$$\mathbb{D}_t\Theta = \mathcal{F}_\Theta(\Theta)\mathrm{d}t. \tag{2.3.4}$$

For a negligible production/dissipation term, the scalar is conserved by the stochastic flow and follows properties highlighted in section 2.2 – e.g. the energy conservation of each realization and the dissipation of the mean field. As in the deterministic case, the 1st law of thermodynamics implies both temperature conservation ($\Theta = T$) and conservation of the amount of substance – e.g. the conservation of salinity ($\Theta = S$):

$$\mathbb{D}_t T = \mathcal{F}_T(T)\mathrm{d}t, \tag{2.3.5}$$

$$\mathbb{D}_t S = \mathcal{F}_S(S)\mathrm{d}t. \tag{2.3.6}$$

The term $\mathcal{F}_\Theta(\Theta)$ corresponds to diabatic terms such as the molecular diffusion process or the radiative heat transfer.

2.3.3 Conservation of momentum

To derive a stochastic representation of the Navier-Stokes equations, pressure forcing is decomposed into continuous component, p, and white-noise term $\dot{p}_\sigma = \frac{\mathrm{d}_t p_\sigma}{\mathrm{d}t}$. The smooth component of the velocity is not only assumed continuous but also time-differentiable (Mémin, 2014). As demonstrated in Appendix 2.E, the flow dynamics for an observer in an uniformly rotating coordinate frame writes:

Navier-Stokes equations under location uncertainty in a rotating frame

Momentum equations

$$\partial_t\boldsymbol{w} + (\boldsymbol{w}^*\cdot\boldsymbol{\nabla})\,\boldsymbol{w} - \frac{1}{2\rho}\sum_{i,j}\partial_i\Big(\rho a_{ij}\partial_j\boldsymbol{w}\Big) + \boldsymbol{f}\times\boldsymbol{w} = \boldsymbol{g} - \frac{1}{\rho}\boldsymbol{\nabla}p + \frac{1}{\rho}\mathcal{F}(\boldsymbol{w}), \tag{2.3.7a}$$

Effective drift

$$\boldsymbol{w}^* = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T, \tag{2.3.7b}$$

Random pressure contribution

$$\boldsymbol{\nabla}\mathrm{d}_t p_\sigma = (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla})\,\boldsymbol{w} - \rho\boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t + \mathcal{F}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{2.3.7c}$$

Mass conservation

$$\mathbb{D}_t\rho + \rho\boldsymbol{\nabla}\cdot\boldsymbol{w}^*\mathrm{d}t = 0, \quad \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = 0. \tag{2.3.7d}$$

Similarly to the Reynolds decomposition, the dynamics associated with the drift component includes an additional stress term, and the large-scale velocity component is advected by an eddy effective drift velocity. The density is driven by a stochastic mass conservation equation or alternatively through the stochastic transport of temperature and salinity (2.3.5-2.3.6), together with a state law. The random density constitutes a random forcing in the large-scale momentum equation.

For incompressible flows, the pressure is then recovered from a modified Poisson equation;

$$-\Delta p = \boldsymbol{\nabla}\cdot\left(\rho(\boldsymbol{w}^*\cdot\boldsymbol{\nabla})\boldsymbol{w} + \rho\boldsymbol{f}\times\boldsymbol{w} - \frac{1}{2}\sum_{ij}\partial_i(\,\rho a_{ij}\partial_j\boldsymbol{w})\right). \tag{2.3.8}$$

The pressure acts as a Lagrangian penalty term to constrain the large scale component to be divergent-free.

This formalization can be compared to another stochastic framework based on scale gap: Stochastic Super-Parametrization (SSP) (Grooms and Majda, 2013, 2014). Both modeling enable separating the large-scale velocity (2.3.7a) and the small-scale contribution (2.3.7c). This is done by a differentiability assumption on the large-scale drift, $\boldsymbol{w}$, in the modeling under location uncertainty, and through the Reynolds decomposition and a point approximation assumption in SSP. However, it can be pointed out that no averaging procedure is settled in the modeling under location uncertainty. Furthermore, the transports of density, temperature and salinity involve random forcings. Unlike SSP, the whole system to be simulated is thus random. This randomness is of main importance for Uncertainty Quantification (UQ) aplications as illustrated theoretically in section 2.2 and numerically in the part II of this set of papers (Resseguier et al., 2017b). Another main difference between the two methods lies in the subgrid tensors parametrization. Each SSP scalar evolution law involves a different subgrid tensor whereas there is a single one (related to the small-scale velocity) for every transports under location uncertainty. For both model it can be noted that the small-scale velocity component is Gaussian conditionally on the large-scale properties. Unlike our models, the SSP proposes a simple evolution model for this unresolved velocity and hence for its statistics. This type of linear forced-dissipative evolution laws, introduced by Eddy-Damped Quasi Normal Markovian (EDQNM) models (Orszag, 1970; Leith, 1971; Chasnov, 1991), could be as well used to specify the diffusion operator $\boldsymbol{\sigma}$ and close the models under location uncertainty. Yet, such closure also need to be parametrized.

2.3.4 Atmosphere and Ocean dynamics approximations

Ocean and atmosphere dynamical models generally rely on several successive approximations. In the following, we review these approximations within the uncertainty framework.

For ocean and atmosphere flows, a partition of the density and pressure is generally considered:

$$\rho = \rho_b + \rho_0(z) + \rho'(x, y, z, t), \tag{2.3.9a}$$

$$p = \widetilde{p}(z) + p'(x, y, z, t). \tag{2.3.9b}$$

Fields $\widetilde{\rho}(z) = \rho_b + \rho_0(z)$ and $\widetilde{p}(z)$ correspond to the density and the pressure at equilibrium (without any motion), respectively; they are deterministic functions and depend on the height only. The pressure and density departures, p' and ρ', are random functions, depending on the uncertainty component. From the expression of the vertical velocity component (2.3.7a), the equilibrium fields are related through an hydrostatic balance:

$$\frac{\partial \widetilde{p}}{\partial z} = -g\widetilde{\rho}(z). \tag{2.3.10}$$

Traditional approximation

This approximation helps to neglect the deflecting rotation forces associated with vertical movements. Considering the first moment conservation along the vertical direction of (2.3.7), with the hydrostatic balance (2.3.10), it writes:

$$\partial_t w + (\boldsymbol{w}^* \cdot \boldsymbol{\nabla})\, w - \frac{1}{2}\sum_{i,j} \partial_i \left(a_{ij} \partial_j w \right) + f_x v - f_y u = -\frac{1}{\rho}\left[\rho' g + \frac{\partial p'}{\partial z} \right] + \mathcal{F}(w). \tag{2.3.11}$$

This approximation is justified when an hydrostatic assumption is employed.

Boussinesq approximation

Within small density fluctuations (*i.e.* the Boussinesq approximation) as observed in the ocean, the stochastic mass conservation reads

$$0 = \mathbb{D}_t \rho + \rho \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t \approx \rho_b \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t. \tag{2.3.12}$$

This implies that the flow is volume-preserving. In an anelastic approximation, density variations dominate. It can be shown we get the weaker constraint, associated with an horizontal uncertainty:

$$\boldsymbol{\nabla}\cdot\boldsymbol{w}-\frac{1}{2}\boldsymbol{\nabla}_H\cdot(\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H)^T=\frac{g}{c^2\rho}(w\tilde{\rho}) \tag{2.3.13}$$

where c^{-2} denotes the velocity of the acoustic waves and subscript H indicates the set of horizontal coordinates. The classical anelastic constraint implicitly assumes a divergence-free condition on the variance tensor divergence (as obtained for homogeneous turbulence).

According to equations (2.3.5) and (2.3.6), temperature and salinity are transported by the random flow. If those tracers do not oscillate too much, the density anomaly, $\rho-\rho_b$, can be approximated by a linear combination of these two properties. And thus, in the Boussinesq approximation, this anomaly is transported:

$$0=D_t(\rho-\rho_b)=\mathbb{D}_t(\rho-\rho_b). \tag{2.3.14}$$

Using the same approximation, the contribution of the momentum material derivative associated with the density variation can be neglected. The Navier-Stokes equations coupling the Boussinesq and traditional approximations then read:

Simple Boussinesq equations under location uncertainty

Momentum equations

$$\partial_t\boldsymbol{w}+(\boldsymbol{w}^*\cdot\boldsymbol{\nabla})\,\boldsymbol{w}-\frac{1}{2}\sum_{i,j}\partial_i\Big(a_{ij}\partial_j\boldsymbol{w}\Big)+f\boldsymbol{k}\times\boldsymbol{u}=b\,\boldsymbol{k}-\frac{1}{\rho_b}\boldsymbol{\nabla}p'+\mathcal{F}(\boldsymbol{w}), \tag{2.3.15a}$$

Effective drift

$$\boldsymbol{w}^*=\begin{pmatrix}\boldsymbol{u}^*\\ w^*\end{pmatrix}=\boldsymbol{w}-\frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T, \tag{2.3.15b}$$

Buoyancy equation

$$\mathbb{D}_t b+N^2\left(w^*\mathrm{d}t+(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right)=\frac{1}{2}\boldsymbol{\nabla}\cdot\left(\boldsymbol{a}_{\bullet z}N^2\right)\mathrm{d}t, \tag{2.3.15c}$$

Random pressure fluctuation

$$\boldsymbol{\nabla}\mathrm{d}_t p_\sigma=-\rho_b\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{w}^*-f\boldsymbol{k}\times(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H+\mathcal{F}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{2.3.15d}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w}=\boldsymbol{\nabla}\cdot\big(\boldsymbol{\sigma}\dot{\boldsymbol{B}}\big)=\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a}=0. \tag{2.3.15e}$$

For this system, the thermodynamics equations are expressed through the buoyancy variable $b=-g\rho'/\rho_b$, and the stratification (Brunt-Väisälä frequency) $N^2(z)=-g\frac{1}{\rho_b}\partial_z\rho_0(z)$ is introduced. The buoyancy term constitutes a random forcing of the vertical large-scale velocity component. Since the density anomaly, $\rho-\rho_b$, has been decomposed into a constant background slope and a residual, the multiplicative noise of equation (2.3.14) is split into an additive and a multiplicative noise in (2.3.15c). The additive noise drains random energy from the stratification toward the buoyancy. Therefore, the buoyancy energy is not conserved due to the background stratification.

Buoyancy oscillations

To illustrate the effect of this additive noise in simple cases, we consider here constant-along-depth buoyancy anomaly and stratification ($\partial_z b=0$ and $\partial_z N=0$) and only a vertical motion component (*i.e* $\boldsymbol{u}=0$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=0$) with no dependence on depth (due to the divergence constraint). Note that this latter constraint on the diffusion tensor, implies that only a_{zz} is non null with no dependence on depth as well. Then, the Boussinesq equations read

$$\partial_t w=b \text{ and } \mathrm{d}_t b=-N^2(w\mathrm{d}t+(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z). \tag{2.3.16}$$

Similarly to the deterministic case, we recognize an oscillatory system if $N^2 > 0$ and a diverging system if $N^2 < 0$ (*i.e.* when lighter fluid is below heavier fluid). The velocity and buoyancy are coupled by gravity and transport. However, in our stochastic framework, the density anomaly is also transported by a random velocity. This highly oscillating velocity may be interpreted as the action of wind on the surface of the ocean. The interaction between this unresolved velocity component and the stratification acts has a random forcing on the oscillator:

$$\mathrm{d}_t \partial_t w + N^2 w \mathrm{d}t = -N^2 (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_z. \tag{2.3.17}$$

To solve this equation, one can note that:

$$\mathrm{d}_t \left(e^{-2iNt} \partial_t (e^{iNt} w) \right) = -N^2 e^{-iNt} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_z. \tag{2.3.18}$$

Then, by integrating twice, we get the solutions of the stochastic system (2.3.16):

$$w(t) = \underbrace{w(0)\cos(Nt) + \partial_t w(0)/N \sin(Nt)}_{=\mathbb{E}(w(t))} - N \int_0^t \sin\big(N(t-r)\big)(\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_r)_z, \tag{2.3.19}$$

$$b(t) = \underbrace{\partial_t w(0)\cos(Nt) - w(0) N \sin(Nt)}_{=\mathbb{E}(b(t))} - N^2 \int_0^t \cos(N(t-r))(\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_r)_z. \tag{2.3.20}$$

The ensemble means are the traditional deterministic solutions whereas the random parts are continuous summations of sine wave with uncorrelated random amplitudes. At each time r, the additive random forcing introduces an oscillation. Without dissipative processes, the latter remains in the system. But, the influence of the past excitations are weighed by sine wave due to the phase change. The buoyancy and the velocity are Gaussian random variables (as linear combinations of independent Gaussian variables). Therefore, their finite dimensional law (*i.e.* the multi-time probability density function) are entirely defined by their mean and covariance functions. The variances can be computed through the Ito isometry (Oksendal, 1998). Then, the velocity covariance can be inferred from the SDE (2.3.17):

$$Cov_w(t, t+\tau) = \frac{a_{zz} N}{4} \cos(N\tau)\left(2Nt - \sin(2Nt)\right) + \frac{a_{zz} N}{4} \sin(N\tau)\left(1 - \cos(2Nt)\right). \tag{2.3.21}$$

The covariance of the buoyancy is similar. Since the interaction between the unresolved velocity component and the background density gradient cannot be resolved deterministically, uncertainties of the dynamics accumulate. Each time introduces a new random uncorrelated excitation. This is why the buoyancy and velocity variances increase linearly with time. In contrast, in a deterministic oscillator with random perturbations of the initial conditions, the variance remains constant and depends solely on the initial velocity variance. This growing also illustrates in a very simple case the possible destabilization effects of the unresolved velocity in the models under location uncertainty.

The first term of the covariance (2.3.21) modulates the variance with a sine wave. The randomness of w is generated by a set of sine wave which have coherent phases and interfere. When $N\tau = 0[2\pi]$ the noises with correlated amplitudes, $(\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_r)_z$, in $w(t)$ and $w(t+\tau)$ are in phase, and thus the velocity covariance is large. When $N\tau = \pi[2\pi]$ these correlated noises have opposite phases, and yields a negative velocity covariance. When $N\tau$ is close to $\frac{\pi}{2}[\pi]$, the noises are in quadrature and the first term of the velocity covariance is zero.

2.3.5 Summary

The fundamental conservation laws (mass, momentum and energy) have been interpreted within the proposed stochastic framework. Usual approximations of fluid dynamics are considered, leading to a stochastic version of Boussinesq equations. As developed, the buoyancy is transported by a smooth large-scale velocity component and a small-scale random field, delta-correlated in time. Consequently, the buoyancy is forced by an additive and a multiplicative noises, uncorrelated in time but correlated in space. The additive noise encodes the interaction between the unresolved velocity and the background stratification. The resulting random buoyancy then appears as an

additive time-correlated random forcing in the vertical momentum equation. Both momentum and thermodynamic equations then involve an inhomogeneous and anisotropic diffusion, and a drift correction that both depend on the unresolved velocity variance tensor, $\boldsymbol{a}$. Assuming hydrostatic equilibrium in this stochastic Boussinesq model directly provides a stochastic version of the primitive equations. A solvable model is also derived from this Boussinesq model. This toy model exemplifies how the random forcing continually increases the variance of the solution.

2.3.6 Guidelines for the derivation of models under location uncertainty

The main steps of the derivation of dynamics under location uncertainty are sketched out below.

(i) The conservation laws of classical mechanics describe variation of some extensive properties. As illustrated in Appendix 2.E for the stochastic Navier-Stokes model, if the extensive property of interest (linear momentum in this Appendix) has a component uncorrelated in time, the variations of this component must be balanced by a very irregular forcing, and can be discarded.

(ii) The stochastic Reynolds transport theorem (2.2.28) enables us to interpret the variation of the time-correlated component of the extensive property. The expression of the stochastic material derivative of an associated intensive quantity follows.

(iii) The formulas (2.2.10) relate this material derivative, D_t, to the stochastic transport operator, $\mathbb{D}_t$. In most cases, these operators coincide.

(iv) Gathering the equations from (ii) and (iii) provides an explicit Eulerian evolution law.

(v) Additional regularity assumptions can be used to separate the large-scale and small-scale components of the evolution law. As an example, the velocity component, $\boldsymbol{w}$, has been assumed to be differentiable with respect to time in this section *i.e.* the acceleration component, $\partial_t \boldsymbol{w}$, is correlated in time. Thus, there is no time-uncorrelated noise in the large-scale momentum evolution law and the random pressure fluctuations appear in a separate equation. This separation is of great interest for deterministic LES-like simulations. However, by this approximation, we lose the conservation of the kinetic energy (2.2.17). For Uncertainty Quantification (UQ) purposes, this separation is not necessary.

(vi) With or without regularity assumptions, usual approximations (*e.g.* the Boussinesq approximation) can be done to simplify further the stochastic model.

Let us point out that the corresponding models involve subgrid terms which generally cannot be neglected. When non-dimentionalized, those subgrid terms are weighted by an additional adimentional number whose value depends on the noise magnitude. For a low noise the approximate dynamical models take a random form that remains similar to their deterministic counterparts. At the opposite, the system is generally significantly changed when considering a strong noise.

A second companion paper (part II) (Resseguier et al., 2017b) describes random versions of Quasi-Geostrophic (QG) and Surface Quasi-Geostrophic (SQG) models with a moderate influence of the subgrid terms, whereas the third one (part III) (Resseguier et al., 2017c) focuses on the same models with a stronger influence of subgrid terms. The two dynamics are significantly different.

To close the stochastic system, the operator $\boldsymbol{\sigma}$ needs to be fully specified. Several solutions can be proposed to that purpose. The simplest specification consists in resorting to a homogeneous parametrization such as the Kraichnan model (Kraichnan, 1968, 1994; Gawędzki and Kupiainen, 1995; Majda and Kramer, 1999). The companion paper Resseguier et al. (2017b) relies on this type of random field with a parameterization fixed from an ideal spectrum. When the small-scale velocity is observable or at least partially observable the structure of that operator can then be estimated. For instance, in Resseguier et al. (2015) a nonparametric and inhomogeneous variance tensor $\boldsymbol{a}(\boldsymbol{x}) = \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{\sigma}(\boldsymbol{x})^T$ is estimated from a sequence of observed velocity. Parametric and/or homogeneous models could also be specified. If no small-scale statistics are available, the choice of a closure can expressed $\boldsymbol{\sigma}$ as a function of large-scale quantities and similarity assumption (Kadri-Harouna

and Mémin, 2016; Chandramouli et al., 2016). The unresolved velocity can be defined as the solution of a simple linearized equations subject to advection by large-scale components, damping and additive random forcing as in *e.g.* quasi-linear approximations (Farrell and Ioannou, 2014) or stochastic super-parameterizations (Grooms and Majda, 2013, 2014). Existing methodologies of data assimilation literature would also be of great interest in this context. Several authors define models from observed correlation length or correlation deformation estimation (Pannekoucke and Massart, 2008; Mirouze and Weaver, 2010; Weaver and Courtier, 2001). Others specify the correlation matrices by diffusion equations (Michel, 2013a,b; Pannekoucke et al., 2014).

2.4 Conclusion

In this paper, a random component is added to the smooth velocity field. This helps model a coarse-graining effect. The random component is chosen Gaussian and uncorrelated in time. Nevertheless, it can be inhomogeneous and anisotropic in space. With such a velocity, the expression of the material derivative is changed. To make this change explicit, we introduce the stochastic transport operator, $\mathbb{D}_t$. The material derivative, D_t, generally coincides with this operator, especially for tracer transports. Otherwise, the difference between these operators has a simple analytic expression. The stochastic transport operator involves an anisotropic and inhomogeneous diffusion, a drift correction and a multiplicative noise. These terms are specified by the statistics of the sub-grid velocity. The diffusion term generalizes the Boussinesq assumption. Moreover, the link between the three previous terms ensures many desired properties for tracers, such as energy conservation and continuous variance increasing. For passive tracer, the PDEs of mean and variance field are derived. The unresolved velocity transfers energy from the small-scale mean field to the variance. This is very suitable to quantify the uncertainty associated with sub-grid dynamics. This randomized dynamics has been called transport under location uncertainty. A stochastic version of the Reynolds transport theorem is then derived. It enables us to compute the time differentiation of extensive properties to interpret the conservation laws of classical mechanics in a stochastic sense.

Applied to the conservation of linear momentum, amount of substance and first principle of thermodynamics, a stochastic version of the Navier-Stokes equations is obtained. Similarly to the deterministic case, a small buoyancy assumption leads to random Boussinesq equations. The random transport of buoyancy involves both a multiplicative and an additive noises. The additive noise encodes the interaction between the unresolved velocity and the background stratification. We schematically presented the action of this last forcing through a solvable model of fluid parcels vertical oscillations.

Under strong rotation and strong stratification assumptions, the stochastic Boussinesq representation simplifies to different mesoscale models depending on the scaling of the subgrid terms. The companion papers part II (Resseguier et al., 2017b) and part III (Resseguier et al., 2017c) describe such models. For a moderate influence of noise-driven subgrid terms, the Potential Vorticity (PV) is randomly transported up to three source terms (Resseguier et al., 2017b). Assuming zero PV in the fluid interior yields the usual Surface Quasi-Geostrophic (SQG) relationship. The stochastic transport of buoyancy, yields a stochastic SQG model referred to as SQG model under Moderate Uncertainty (SQG_{MU}). This two-dimensional nonlinear dynamics enables Resseguier et al. (2017b) to numerically unveil advantages of the models under location uncertainty in terms of small-scale structures restoration (in a single realization) and ensemble model error prediction (with an improvement compared to perturbed deterministic models of one order of magnitude).

To go beyond the framework of this paper, larger-scale random dynamics can be inferred by averaging the models under location uncertainty using singular perturbation or stochastic invariant manifold theories (Gottwald and Harlim, 2013). Finally, a delta-correlated process and stochastic calculus may seem insufficient to model the smallest velocity scales. Ito formulas deal with white-noise forcing and contains only second-order terms. For higher order terms, such as hyperviscosity, more complete theories exist (Klyatskin, 2005).

Acknowledgments

The authors thank Aurélien Ponte, Jeroen Molemaker, Guillaume Roulet and Jonathan Gula for helpful discussions. We also acknowledge the support of the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project.

Appendix

2.A Quadratic variation

The quadratic co-variation process denoted $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle_t$, is defined as the limit in probability over a partition $\{t_1, \ldots, t_n\}$ of $[0, t]$ with $t_1 < t_2 < \cdots < t_n$, and a partition spacing $\delta t_i = t_i - t_{i-1}$, noted as $|\delta t|_n = \max_i \delta t_i$ and such that $|\delta t|_n \to 0$ when $n \to \infty$:

$$\langle \boldsymbol{X}, \boldsymbol{Y} \rangle_t = \lim_{|\delta t|_n \to 0}^{P} \sum_{i=0}^{n-1} \big(\boldsymbol{X}(t_{i+1}) - \boldsymbol{X}(t_i)\big)\big(\boldsymbol{Y}(t_{i+1}) - \boldsymbol{Y}(t_i)\big)^T.$$

For Brownian motions, it follows $\langle B, B \rangle_t = t$, $\langle B, h \rangle_t = \langle h, B \rangle_t = \langle h, h \rangle_t = 0$, where h is a deterministic function (or a random time-differentiable function) and B a scalar Brownian motion. The quadratic co-variation of the uncertainty component reads

$$\begin{aligned} \left\langle \int_0^t \left(\boldsymbol{\sigma}(\boldsymbol{x}, t) \mathrm{d}\boldsymbol{B}_t\right)^i, \int_0^t \left(\boldsymbol{\sigma}(\boldsymbol{y}, t) \mathrm{d}\boldsymbol{B}_t\right)^j \right\rangle &= \int_0^t \sum_k \int_\Omega \breve{\sigma}^{ik}(\boldsymbol{x}, \boldsymbol{z}, s)\, \breve{\sigma}^{jk}(\boldsymbol{y}, \boldsymbol{z}, s) \mathrm{d}s \mathrm{d}\boldsymbol{z}, \\ &\triangleq \int_0^t a^{ij}(\boldsymbol{x}, \boldsymbol{y}, s) \mathrm{d}s. \end{aligned} \tag{2.A.1}$$

Its time derivative corresponds to the spatial covariance tensor. The diagonal of this tensor, denoted the variance tensor, corresponds to $\boldsymbol{x} = \boldsymbol{y}$. For isotropic random fields, $\breve{\boldsymbol{\sigma}}(\boldsymbol{x}, \boldsymbol{z}) = \breve{\boldsymbol{\sigma}}(\|\boldsymbol{x} - \boldsymbol{z}\|_2)$, the quadratic variation is a constant diagonal matrix.

2.B Link between the material derivative D_t and the operator $\mathbb{D}_t$

Let us assume:

$$\mathbb{D}_t \Theta = f \mathrm{d}t + \boldsymbol{h}^T \mathrm{d}\boldsymbol{B}_t. \tag{2.B.1}$$

By definition of $\mathbb{D}_t$ (Eq. (2.2.9)),

$$\mathbb{D}_t \Theta = \mathrm{d}_t \Theta + (\boldsymbol{w}^* \mathrm{d}t + \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla} \Theta - \frac{1}{2} \boldsymbol{\nabla} \cdot (\boldsymbol{a} \boldsymbol{\nabla} \Theta)\, \mathrm{d}t. \tag{2.B.2}$$

It yields:

$$\mathrm{d}_t \Theta = \left(f - \boldsymbol{w}^* \cdot \boldsymbol{\nabla} \Theta + \frac{1}{2} \boldsymbol{\nabla} \cdot (\boldsymbol{a} \boldsymbol{\nabla} \Theta)\right) \mathrm{d}t + \boldsymbol{h}^T \mathrm{d}\boldsymbol{B}_t - (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla} \Theta. \tag{2.B.3}$$

Denoting $\boldsymbol{H}_\Theta$ the Hessian of the function Θ, we have:

$$\mathrm{d}_t \boldsymbol{\nabla} \Theta = \boldsymbol{\nabla} \left(f - \boldsymbol{w}^* \cdot \boldsymbol{\nabla} \Theta + \frac{1}{2} \boldsymbol{\nabla} \cdot (\boldsymbol{a} \boldsymbol{\nabla} \Theta)\right) \mathrm{d}t + \boldsymbol{\nabla} \boldsymbol{h}^T \mathrm{d}\boldsymbol{B}_t - \boldsymbol{\nabla} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)^T \boldsymbol{\nabla} \Theta - \boldsymbol{H}_\Theta (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t). \tag{2.B.4}$$

As Θ is a random function, its material derivative, *i.e.* the differential of $\Theta(t, \boldsymbol{X}_t)$, involves the composition of two stochastic processes. Its evaluation requires the use of a generalized Ito formula, referred to as the Ito-Wentzell formula (see theorem 3.3.1, Kunita, 1997). In the same way as the

classical Ito formula[5], it incorporates quadratic variation terms related to the process $\boldsymbol{X}_t$, but also co-variation terms between $\boldsymbol{X}_t$ and the gradient of the random function Θ, as:

$$
\begin{aligned}
(D_t\Theta)(t,\boldsymbol{X}_t) &\triangleq \mathrm{d}\left(\Theta\left(t,\boldsymbol{X}_t\right)\right), && (2.B.5)\\
&= \mathrm{d}_t\Theta + \mathrm{d}\boldsymbol{X}_t\cdot\boldsymbol{\nabla}\Theta + \frac{1}{2}\mathrm{tr}\left(\mathrm{d}<\boldsymbol{X}_t,\boldsymbol{X}_t^T>\boldsymbol{H}_\Theta\right) + \mathrm{d}<\boldsymbol{X}_t^T,\boldsymbol{\nabla}\Theta>, && (2.B.6)\\
&= \mathrm{d}_t\Theta + \left(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{a}\boldsymbol{H}_\Theta\right)\mathrm{d}t + \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t\\
&\quad -\sum_{k=1}^{d}\boldsymbol{\sigma}_{\bullet k}^T\boldsymbol{\nabla}\boldsymbol{\sigma}_{\bullet k}^T\boldsymbol{\nabla}\Theta\mathrm{d}t - \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{H}_\Theta\boldsymbol{\sigma}\right)\mathrm{d}t, \qquad \text{(using (2.B.4))} && (2.B.7)\\
&= \mathrm{d}_t\Theta + \left(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta - \frac{1}{2}\mathrm{tr}\left(\boldsymbol{a}\boldsymbol{H}_\Theta\right)\mathrm{d}t\\
&\quad +\mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t - \left(\boldsymbol{\nabla}\cdot\boldsymbol{a} - \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\boldsymbol{\sigma}^T\right)\boldsymbol{\nabla}\Theta\mathrm{d}t, && (2.B.8)\\
&= \mathrm{d}_t\Theta + \left(\left(\boldsymbol{w} - \frac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)^T + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T\right)\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta\\
&\quad -\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\Theta)\,\mathrm{d}t + \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t, && (2.B.9)\\
&= \mathbb{D}_t\Theta + \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t. \qquad \text{(by definition of } \mathbb{D}_t) && (2.B.10)
\end{aligned}
$$

Finally, taking this Lagrangian formulation at $\boldsymbol{X}_t = \boldsymbol{x}$ leads to the (Eulerian) expression of the material derivative:

$$
D_t\Theta \triangleq \left(\mathrm{d}_t\left(\Theta\left(t,\boldsymbol{X}_t\right)\right)\right)_{|\boldsymbol{X}_t=\boldsymbol{x}} = \mathbb{D}_t\Theta + \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t. \tag{2.B.11}
$$

Conversely, assuming that the explicit expression (2.B.1) is unknown whereas the expression of the material derivative is known:

$$
D_t\Theta = \tilde{f}\mathrm{d}t + \tilde{\boldsymbol{h}}^T\mathrm{d}\boldsymbol{B}_t. \tag{2.B.12}
$$

Using the equation (2.B.11)

$$
\mathbb{D}_t\Theta = D_t\Theta - \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t = \left(f - \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\right)\mathrm{d}t + \boldsymbol{h}^T\mathrm{d}\boldsymbol{B}_t. \tag{2.B.13}
$$

By uniqueness of the martingale decomposition (term in $\mathrm{d}t$ and term in $\mathrm{d}\boldsymbol{B}_t$), we can identify $\tilde{\boldsymbol{h}} = \boldsymbol{h}$. Then, using again (2.B.11) yields:

$$
\mathbb{D}_t\Theta = D_t\Theta - \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T\right)\mathrm{d}t = D_t\Theta - \mathrm{tr}\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\tilde{\boldsymbol{h}}^T\right)\mathrm{d}t. \tag{2.B.14}
$$

2.C The evolution of the variance of a passive tracer

For a passive scalar Θ, we denote $Y \triangleq \Theta - \mathbb{E}(\Theta)$ and $Z \triangleq Y^2$. The goal is to find the evolution of $Var(b) = \mathbb{E}(Z)$. The conservation of the tracer, says $\mathbb{D}_t\Theta = 0$, gives the evolution equation of Y:

$$
\mathrm{d}_tY = -\left(\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}\right)Y\mathrm{d}t + \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}Y\right)\mathrm{d}t - \left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\Theta. \tag{2.C.1}
$$

And, by the Ito formula,

$$
\begin{aligned}
\mathrm{d}_tZ &= 2Y\mathrm{d}_tY + \mathrm{d}_t<Y,Y>, && (2.C.2)\\
&= -\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}Z\mathrm{d}t + Y\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}Y)\,\mathrm{d}t - 2Y\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\Theta + \left(\boldsymbol{\nabla}\Theta\right)^T\boldsymbol{a}\boldsymbol{\nabla}\Theta\mathrm{d}t. && (2.C.3)
\end{aligned}
$$

Taking the expectation of this expression and using $\Theta = \mathbb{E}(\Theta) + Y$, yields

$$
\partial_tVar(b) = -\boldsymbol{w}^\star\cdot\boldsymbol{\nabla}Var(b) + \mathbb{E}\left\{Y\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}Y)\right\} + \left(\boldsymbol{\nabla}\mathbb{E}(\Theta)\right)^T\boldsymbol{a}\boldsymbol{\nabla}\mathbb{E}(\Theta) + \mathbb{E}\left\{\left(\boldsymbol{\nabla}Y\right)^T\boldsymbol{a}\boldsymbol{\nabla}Y\right\}. \tag{2.C.4}
$$

[5] relevant only to express the differential of a time-differentiable function of a stochastic process.

Expanding the second term of the right-hand side makes appear the diffusion of the variance

$$\begin{aligned}
\mathbb{E}\left\{Y\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}Y)\right\} &= \sum_{i,j}\partial_i a_{ij}\mathbb{E}\left\{Y\partial_j Y\right\} + \sum_{i,j} a_{ij}\mathbb{E}\left\{Y\partial^2_{ij}Y\right\}, && (2.C.5)\\
&= \sum_{i,j}\partial_i a_{ij}\mathbb{E}\left\{Y\partial_j Y\right\} + \sum_{i,j} a_{ij}\mathbb{E}\left\{\frac{1}{2}\partial^2_{ij}(Y^2) - \partial_i Y\partial_j Y\right\}, && (2.C.6)\\
&= \frac{1}{2}\sum_{i,j}\partial_i a_{ij}\partial_j\mathbb{E}\left\{Z\right\} + \frac{1}{2}\sum_{i,j} a_{ij}\partial_{ij}\mathbb{E}\left\{Z\right\} - \mathbb{E}\left\{(\boldsymbol{\nabla}Y)^T\boldsymbol{a}\boldsymbol{\nabla}Y\right\}, && (2.C.7)\\
&= \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}Var(b)\right) - \mathbb{E}\left\{(\boldsymbol{\nabla}Y)^T\boldsymbol{a}\boldsymbol{\nabla}Y\right\}. && (2.C.8)
\end{aligned}$$

Finally, the evolution law of the variance writes

$$\partial_t Var(b) + \boldsymbol{w}^\star\cdot\boldsymbol{\nabla}Var(b) = \boldsymbol{\nabla}\cdot\left(\frac{1}{2}a\boldsymbol{\nabla}Var(b)\right) + (\boldsymbol{\nabla}\mathbb{E}(\Theta))^T\boldsymbol{a}\boldsymbol{\nabla}\mathbb{E}(\Theta). \quad (2.C.9)$$

2.D Stochastic extension of the Reynolds transport theorem

In the following, we consider a scalar function ϕ transported by the stochastic flow $\boldsymbol{x}_0 \mapsto \boldsymbol{x} = \boldsymbol{X}_t(\boldsymbol{x}_0)$ (2.2.5). Its initial time value g:

$$\phi(\boldsymbol{X}_t(\boldsymbol{x}_0), t) = g(\boldsymbol{x}_0). \quad (2.D.1)$$

We will assume that the initial function $g : \Omega \to \mathbb{R}$ has bounded spatial gradients and vanishes outside the initial volume $\mathcal{V}(t_0)$ and on its boundary. The material derivative of ϕ is:

$$(D_t\phi)(t, \boldsymbol{X}_t(\boldsymbol{x}_0)) \triangleq \mathrm{d}\left(\phi(t, \boldsymbol{X}_t(\boldsymbol{x}_0))\right) = \mathrm{d}g(\boldsymbol{x}_0) = 0. \quad (2.D.2)$$

With equation (2.2.10), it writes in the Eulerian space:

$$0 = \mathbb{D}_t\phi \triangleq \mathrm{d}_t\phi + (\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\phi - \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\phi\right)\mathrm{d}t, \quad (2.D.3)$$

with

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T. \quad (2.D.4)$$

Thus,

$$\mathrm{d}_t\phi = \mathcal{L}\phi\mathrm{d}t - \boldsymbol{\nabla}\phi\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t, \quad (2.D.5)$$

$$\mathcal{L}\phi = -\boldsymbol{\nabla}\phi\cdot\boldsymbol{w}^\star + \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}\phi). \quad (2.D.6)$$

Denoting J the Jacobian corresponding to the change of variables $\boldsymbol{x}_0 \mapsto \boldsymbol{x} = \boldsymbol{X}_t(\boldsymbol{x}_0)$, the differential of the integral over a material volume of the product $q\phi$ is given by

$$\begin{aligned}
\mathrm{d}\int_{\mathcal{V}(t)}(q\phi)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} &= \mathrm{d}\int_{\mathcal{V}(0)}(Jq\phi)(\boldsymbol{X}_t(\boldsymbol{x}_0),t)\mathrm{d}\boldsymbol{x}_0, && (2.D.7)\\
&= \mathrm{d}\int_{\Omega}(Jq\phi)(\boldsymbol{X}_t(\boldsymbol{x}_0),t)\mathrm{d}\boldsymbol{x}_0, && (2.D.8)\\
&= \mathrm{d}\int_{\Omega}(q\phi)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}, && (2.D.9)\\
&= \int_{\Omega}\big(\mathrm{d}_t q\phi + q\mathrm{d}_t\phi + \mathrm{d}_t\langle q,\phi\rangle\big)(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x}, && (2.D.10)
\end{aligned}$$

where the second line comes from $\phi(\boldsymbol{X}_t(\boldsymbol{x}_0),t) = g(\boldsymbol{x}_0) = 0$ if $\boldsymbol{x}_0 \in \Omega\backslash\mathcal{V}(t_0)$ and the last line from the Ito's formula. To compute the quadratic covariation $\mathrm{d}_t\langle q,\phi\rangle$, we introduce a notation for the non-differentiable part (*i.e.* the integral in $\mathrm{d}\boldsymbol{B}_t$) of $\int_0^t \mathbb{D}_t q$:

$$\mathbb{D}_t q = f\mathrm{d}t + \boldsymbol{h}^T\mathrm{d}\boldsymbol{B}_t. \tag{2.D.11}$$

Together with the stochastic operator, $\mathbb{D}_t$, this relation determines the form of the time differential of q:

$$\mathrm{d}_t q = m\ \mathrm{d}t + (-\boldsymbol{\nabla}q^T\boldsymbol{\sigma} + \boldsymbol{h}^T)\,\mathrm{d}\boldsymbol{B}_t. \tag{2.D.12}$$

Hence, from (2.D.5), we have

$$\mathrm{d}\int_\Omega q\phi = \int_\Omega\Big[\mathrm{d}_t q\phi + q\big(\mathcal{L}\phi\mathrm{d}t - \boldsymbol{\nabla}\phi\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big) - \boldsymbol{\nabla}\phi^T\boldsymbol{\sigma}\,(-\boldsymbol{\sigma}^T\boldsymbol{\nabla}q + \boldsymbol{h})\,\mathrm{d}t\Big]. \tag{2.D.13}$$

Introducing $\mathcal{L}^*$ the (formal) adjoint of the operator $\mathcal{L}$ in the space $L^2(\Omega)$ with Dirichlet boundary conditions, this can be written as

$$\int_\Omega\Big[\mathrm{d}_t q + \big(\mathcal{L}^*q - \boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla}q) + \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{h})\big)\mathrm{d}t + \boldsymbol{\nabla}\cdot\big(q\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)\Big]\phi. \tag{2.D.14}$$

With the complete expression of $\mathcal{L}^*$ (the second right-hand term of 2.D.6 is self-adjoint), the condition $\phi(\boldsymbol{x},t) \to 1\!\!1_{\mathcal{V}(t)/\partial\mathcal{V}(t)}$, where $1\!\!1$ stands for the characteristic function, leads to the following form of this differential:

$$\begin{aligned}
\mathrm{d}\int_{\mathcal{V}(t)} q &= \int_{\mathcal{V}(t)}\Big[\mathrm{d}_t q + \Big(\boldsymbol{\nabla}\cdot\big(q\boldsymbol{w}^\star\big) + \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\boldsymbol{h})\Big)\mathrm{d}t + \boldsymbol{\nabla}\cdot\big(q\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)\Big], && (2.D.15)\\
&= \int_{\mathcal{V}(t)}\Big[\mathbb{D}_t q + \mathrm{tr}\,(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{h}^T)\,\mathrm{d}t + (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\boldsymbol{h}\ \mathrm{d}t + \boldsymbol{\nabla}\cdot(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\,q\Big], && (2.D.16)\\
&= \int_{\mathcal{V}(t)}\Big[D_t q + \boldsymbol{\nabla}\cdot(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\,q + (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\boldsymbol{h}\ \mathrm{d}t\Big], && (2.D.17)\\
&= \int_{\mathcal{V}(t)}\Big[D_t q + \boldsymbol{\nabla}\cdot(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\,q + \mathrm{d}\left\langle\int_0^t D_{t'}q, \int_0^t \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\right\rangle\Big], && (2.D.18)
\end{aligned}$$

where the third line comes from the explicit link (2.2.10), between the stochastic transport operator $\mathbb{D}_t$ and the material derivative D_t.

2.E Stochastic Navier-Stokes model

From the conservation of linear momentum, the balance between the momentum variation and the forces can be expressed as:

$$\mathrm{d}\int_{\mathcal{V}(t)}\rho\big(\boldsymbol{w} + \boldsymbol{\sigma}\dot{\boldsymbol{B}}\big) = \int_{\mathcal{V}(t)}\mathrm{d}_t\boldsymbol{F}. \tag{2.E.1}$$

The left-hand term must be interpreted in a distribution sense, the small-scale velocity, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, being non-continuous. For every test function $h \in C_0^\infty(\mathbb{R}_+)$, we have:

$$\int_{\mathbb{R}_+} h(t)\mathrm{d}\int_{\mathcal{V}(t)}\rho\boldsymbol{w} - \int_{\mathbb{R}_+}\frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)}\rho\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}\mathrm{d}_t\boldsymbol{F}. \tag{2.E.2}$$

Both sides of this equation must have the same structure, and the forces can be written as:

$$\int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}\mathrm{d}_t\boldsymbol{F} = -\int_{\mathbb{R}_+}\frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)}\rho\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t + \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}(\boldsymbol{\eta}\mathrm{d}t + \boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t)\,. \tag{2.E.3}$$

The right-hand first term must compensate the white-noise distributional differentiation of (2.E.2), whereas the last term of (2.E.3) provides the structure of the forces under location uncertainty. The forces are due to the gravitation potential Φ_a within the absolute frame, pressure and friction forces, $\mathrm{d}_t\mathcal{F}(\boldsymbol{w},\boldsymbol{\sigma})$. A direct stochastic extension of the deterministic forces expression reads:

$$\int_{\mathcal{V}(t)}(\boldsymbol{\eta}\mathrm{d}t+\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t)=\int_{\mathcal{V}(t)}(\rho\boldsymbol{\nabla}\Phi_a\mathrm{d}t-\boldsymbol{\nabla}(p\mathrm{d}t+\mathrm{d}_tp_\sigma)+\mathrm{d}_t\mathcal{F}(\boldsymbol{w},\boldsymbol{\sigma}))\,. \tag{2.E.4}$$

The pressure term p denotes the continuous contribution of the pressure. The other term, $\dot{p_\sigma}$, is a zero-mean non-continuous stochastic process (the white noise part of the pressure). It describes the pressure fluctuations due to the random velocity component. Note that the gravity force is continuous in time, whereas the friction force applies both on the deterministic and stochastic velocity components. For a fixed observer in a rotating frame, the rate of change of the fluid velocity incorporates (considering the rotation, $\boldsymbol{f}$, constant in time) the centripetal acceleration and the Coriolis acceleration as additional terms. The centrifugal force is included within an effective gravity, $\boldsymbol{g}=-\boldsymbol{\nabla}\Phi$. The Coriolis term applies both to the large-scale component of the velocity and to the random small-scale field.

The transport equation applied to the linear momentum gives:

$$\mathrm{d}\int_{\mathcal{V}(t)}\rho\boldsymbol{w} \quad = \quad \int_{\mathcal{V}(t)}\mathbb{D}_t\left(\rho\boldsymbol{w}\right)+\rho\boldsymbol{w}\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star\mathrm{d}t. \tag{2.E.5}$$

With $\mathbb{D}_t$ given by (2.2.9), the equation (2.E.5) can be expressed in terms of ρ, $\boldsymbol{w}$ and $\mathrm{d}_t(\rho\boldsymbol{w})$. The large-scale velocity $\boldsymbol{w}$ is assumed to be differentiable in time,

$$\mathrm{d}_t(\rho\boldsymbol{w})=\mathrm{d}_t\rho\boldsymbol{w}+\rho\partial_t\boldsymbol{w}\mathrm{d}t. \tag{2.E.6}$$

The density time derivative, $\mathrm{d}_t\rho$ uses (2.2.9) and the mass conservation equation:

$$\mathbb{D}_t\rho+\rho\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star=0. \tag{2.E.7}$$

From equations (2.E.5), (2.2.9), (2.E.6) and (2.E.7), the variation of the large-scale linear momentum reads:

$$\mathrm{d}\int_{\mathcal{V}(t)}\rho w_i \quad = \quad \int_{\mathcal{V}(t)}\Big(\rho\big(\partial_tw_i\mathrm{d}t+\rho\left(\boldsymbol{w}^*\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}w_i-\frac{1}{2}\boldsymbol{\nabla}\cdot\left(\rho\boldsymbol{a}\boldsymbol{\nabla}w_i\right)\mathrm{d}t\Big). \tag{2.E.8}$$

From the balance between the forces (2.E.4) and the momentum variation (2.E.8), the expression of the flow dynamics for an observer in an uniformly rotating coordinate frame is then obtained by considering the slow temporal bounded variation terms and the Brownian terms.

Chapter 3

Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part II: Quasi-geostrophic models and efficient ensemble spreading. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017b

Abstract

Models under location uncertainty are derived assuming that a component of the velocity is uncorrelated in time. The material derivative is accordingly modified to include an advection correction, inhomogeneous and anisotropic diffusion terms and a multiplicative noise contribution. In this paper, simplified geophysical dynamics are derived from a Boussinesq model under location uncertainty. Invoking usual scaling approximations and a moderate influence of the subgrid terms, stochastic formulations are obtained for the stratified Quasi-Geostrophy (QG) and the Surface Quasi-Geostrophy (SQG) models. Based on numerical simulations, benefits of the proposed stochastic formalism are demonstrated. A single realization of models under location uncertainty can restore small-scale structures. An ensemble of realizations further helps to assess model error prediction and outperforms perturbed deterministic models by one order of magnitude. Such a high uncertainty quantification skill is of primary interests for assimilation ensemble methods. MATLAB® code examples are available online.
Keywords: stochastic sub-grid parameterization, uncertainty quantification, ensemble forecasts.

3.1 Introduction

Ensemble forecasting and filtering are widely used in geophysical sciences for forecasting and climate projection. In practice, dynamical models are randomized through their initial conditions and a Gaussian error model, and are generally found to be underdispersive (Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013; Berner et al., 2011; Snyder et al., 2015) with a low variance. As a consequence, errors are underestimated and observations are hardly taken into account. Corrections are considered by incorporating inflation procedures or hyperprior to increase the variance of ensemble Kalman filters (Anderson and Anderson, 1999; Bocquet et al., 2015). However, such corrections do not provide an accurate spatial localization of the errors.

Another difficulty of ensemble methods lies in the huge dimensions of the involved state spaces. For obvious computational reasons, ensembles for geophysical applications appear constrained and

limited to small sizes. It thus becomes primordial to build strategies to best track the most likely dynamical events. From this point of view, ensemble simulations and stochastic dynamics have clear advantages over the deterministic models.

The simplest random models are defined from Langevin equations with linear damping and additive isotropic Gaussian noise, as, for instance, the linear inverse models (e.g. Penland and Matrosova, 1994; Penland and Sardeshmukh, 1995), or the Eddy-Damped Quasi Normal Markovian (EDQNM) models (e.g. Orszag, 1970; Leith, 1971; Chasnov, 1991). Among other empirical stochastic models, the Stochastic Kinetic Energy Backscatter (SKEBS) (Shutts, 2005; Berner et al., 2009, 2011) and the Stochastic Perturbed Physics Tendency scheme (SPPT) (Buizza et al., 1999) introduce correlated multiplicative noises. SPPT and SKEBS methods have been successfully applied in operational weather forecast centers (Franzke et al., 2015). To target highly non-Gaussian distribution of fluid dynamics properties, an attractive path is to infer randomness from physics (Berner et al., 2015). For this purpose, the time-scale separation assumption is convenient. Hasselmann (1976) already relied on it for geophysical fluid dynamics. This assumption is the foundation of averaging and homogenization theories (Kurtz, 1973; Papanicolaou and Kohler, 1974; Givon et al., 2004; Gottwald and Melbourne, 2013; Mitchell and Gottwald, 2012; Gottwald and Harlim, 2013; Franzke et al., 2015; Gottwald et al., 2015). A successful application of homogenization theory in geophysics is the MTV algorithms (Majda et al., 1999, 2001; Franzke et al., 2005; Majda et al., 2008). The homogenized dynamics is cubic with correlated additive and multiplicative (CAM) noises. This noise structure is able to produce intermittency and extreme events. In practice, the non-linearity of the small-scale equation (fast dynamics) is conveniently replaced by a noise and a damping terms before the homogenization procedure. Noise statistics are estimated from data, with Gaussian assumptions.

In Resseguier et al. (2017a), following Mémin (2014), another approach has been considered to help derive models under location uncertainty based on stochastic calculus and the Ito-Wentzell formula (Kunita, 1997). Mikulevicius and Rozovskii (2004) and Flandoli (2011) already introduce this methodology. Yet, their works mostly focused on pure mathematical aims: existence and uniqueness of SPDE solutions. For our more practical purpose, the large-scale is understood as sub-sampled in time, and the remaining small-scale velocity component is then considered as uncorrelated in time.

Starting with the definition of the revised transport under location uncertainty (section 2), developments are then carried out to derive and analyze the stochastic versions of Quasi-Geostrophy (QG) and Surface Quasi-Geostrophy (SQG) models with a moderate influence of sub-grid terms (section 3). Numerical results highlight the potential of these models under location uncertainty, especially for ensemble forecast (Section 4).

3.2 Models under location uncertainty

This section briefly outlines main theoretical results discussed in Resseguier et al. (2017a). The velocity is decomposed between a possibly random large-scale component, $\boldsymbol{w}$, and a time-uncorrelated component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. The latter is Gaussian, correlated in space with possible inhomogeneities and anisotropy. Hereafter, this unresolved velocity component will further be assumed to be solenoidal. To parameterize those spatial correlations, we apply an infinite-dimensional linear operator, $\boldsymbol{\sigma}$, to a d-dimensional space-time white noise[1], $\dot{\boldsymbol{B}}$.

In time, the velocity is irregular. The material derivative, D_t, is then changed. In most cases, it coincides with the stochastic transport operator, $\mathbb{D}_t$, defined for every field, Θ, as follows:

$$\mathbb{D}_t\Theta \triangleq \underbrace{\mathrm{d}_t\Theta}_{\substack{\triangleq\ \Theta(\boldsymbol{x},t+\mathrm{d}t)-\Theta(\boldsymbol{x},t)\\ \text{Time increment}}} + \underbrace{(\boldsymbol{w}^\star\mathrm{d}t+\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\Theta}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}}\mathrm{d}t, \tag{3.2.1}$$

where the time increment term $\mathrm{d}_t\Theta$ is used in place of the partial time derivative, as Θ is in general non-differentiable. The diffusion coefficient matrix, $\boldsymbol{a}$, is solely defined by the one-point one-time

[1]Formally each coefficient of $(t \mapsto \boldsymbol{B}_t)$ is a cylindrical I_d-Wiener process (see Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007) for more information on infinite dimensional Wiener processes and cylindrical I_d-Wiener processes).

covariance of the unresolved displacement per unit of time:

$$\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T = \frac{\mathbb{E}\left\{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}}{\mathrm{d}t}, \tag{3.2.2}$$

and the modified drift is given by

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T. \tag{3.2.3}$$

For a divergent small-scale velocity, this drift would involve an additional component (Resseguier et al., 2017a). With this modified material derivative (3.2.1), the transport equations under location uncertainty involve three new terms: a modification of the large-scale advection ($\boldsymbol{w}^\star$ instead of $\boldsymbol{w}$), an inhomogeneous and anisotropic diffusion and a multiplicative noise. This random forcing is directly related to the advection by the unresolved velocity.

For incompressible flows ($\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star = 0$), the energy of any tracer, Θ, is conserved for each realization:

$$\mathrm{d}\int_\Omega \Theta^2 = 0, \tag{3.2.4}$$

where Ω is the spatial domain. This still holds for active tracers. The diffusion dissipates as much energy as the multiplicative noise is injecting it in the system. In particular, the (ensemble) mean of the energy, $\mathbb{E}\int_\Omega \Theta^2$, is conserved. This results ensures a constant balance between the energy of the mean and the (ensemble) variance. The energy fluxes in these stochastic models are more thoroughly described in Resseguier et al. (2017a).

A random version of the Reynolds transport theorem can further be derived (Mémin, 2014; Resseguier et al., 2017a). From this theorem, usual conservation of mechanics (mass, linear momentum, energy and amount of substance) can be expressed in a stochastic sense. Random Navier-Stokes and Boussinesq models can then be derived. This last model describes the stochastic transports of velocity and density anomaly, as well as incompressibility conditions.

3.3 Mesoscales under moderate uncertainty

To simplify the stochastic Boussinesq model of Resseguier et al. (2017a), Quasi-Geostrophic (QG) models are developed for large horizontal length scales, L, such as:

$$\frac{1}{Bu} = \left(\frac{Fr}{Ro}\right)^2 = \left(\frac{L}{L_d}\right)^2 \sim 1 \text{ and } \frac{1}{Ro} = \frac{Lf_0}{U} \gg 1, \tag{3.3.1}$$

where U is the horizontal velocity scale, $L_d \triangleq \frac{Nh}{f}$ is the Rossby deformation radius, N is the stratification (Brunt-Väisälä frequency) and h is the characteristic vertical length scale. The Rossby deformation radius explicitly defines the mesoscale range, over which both kinetic and buoyancy effects are important, and strongly interact. In the following, both differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ, represent 2D operators.

3.3.1 Specific scaling assumptions

Hereafter, we explicit scaling assumptions to derive the non-dimensional version of the stochastic Boussinesq model.

Quadratic variation scaling

Besides traditional ones, another dimensionless number, Υ, is introduced to relate the large-scale kinetic energy to the energy dissipation due to the horizontal small-scale random component. In

the following, $\boldsymbol{\sigma}_{H\bullet}$ stands for the horizontal component of $\boldsymbol{\sigma}$, $\boldsymbol{a}_H$ for $\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T$ and A_u for its scaling. The new dimensionless number is defined by:

$$\Upsilon \triangleq \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{3.3.2}$$

This number compares horizontal advective and diffusive terms in the momentum and buoyancy equations. This number can also be related to the ratio between the Mean Kinetic Energy (MKE), U^2, and the Turbulent Kinetic Energy (TKE), A_u/T_σ, where T_σ is the small-scale correlation time. This reads:

$$\Upsilon = \frac{1}{\epsilon}\frac{MKE}{TKE}, \tag{3.3.3}$$

where $\epsilon = T_\sigma/T$ is the ratio of the small-scale to the large-scale correlation times. This parameter, ϵ, is central in homogenization and averaging methods (Majda et al., 1999; Givon et al., 2004; Gottwald and Melbourne, 2013). The number R_o/Υ can then be stated to measure the ratio between sub-grid terms and the Coriolis force. In the usual deterministic case and the limit of small Rossby number, the predominant terms of the horizontal momentum equation then correspond to the geostrophic balance. In the stochastic case, this balance also applies from weak ($\Upsilon \gg 1$) to moderate ($\Upsilon \sim 1$) uncertainty. However, if Υ/R_o is close enough to $O(1)$, this geostrophic balance is modified due to the diffusion effects introduced by the small-scale random velocity. Hereafter, developments focus on the moderate uncertainty case. Resseguier et al. (2017c) deals with the strong uncertainty case.

To evaluate Υ for a given flow at a given scale, eddy viscosity or diffusivity values help the determination of A_u. Boccaletti et al. (2007) give some examples of canonical values. Then, the typical resolved velocity and length scale lead to Υ. If no canonical values are known, absolute diffusivity or similar mixing diagnoses could be measured (Keating et al., 2011) as a proxy of the variance tensor.

Vertical unresolved velocity

The scaling to compare vertical to horizontal unresolved velocities is also considered:

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{R_o}{B_u}D, \tag{3.3.4}$$

where $D = \frac{h}{L}$ is the aspect ratio and the subscript H indicates horizontal coordinates. This scaling can be derived from the ω-equation (Giordani et al., 2006). For any velocity $\mathfrak{u} = (\mathfrak{u}_H, \mathfrak{w})^T$, which scales as $(\mathfrak{U}, \mathfrak{U}, \mathfrak{W})^T$, this equation reads

$$f_0^2\partial_z^2\mathfrak{w} + N^2\Delta\mathfrak{w} = \boldsymbol{\nabla}\cdot\boldsymbol{Q} \approx -\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\mathfrak{u}_H^T\boldsymbol{\nabla}\mathfrak{b}) \approx -f_0\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\mathfrak{u}_H^T\partial_z\mathfrak{u}_H^\perp\right), \tag{3.3.5}$$

where $\mathfrak{b}$ stands for the buoyancy variable and $\boldsymbol{Q}$ for the so-called $\boldsymbol{Q}$-vector. In its non-dimentional version, the ω-equation reads:

$$\frac{\mathfrak{W}}{\mathfrak{U}}\left(\partial_z^2\mathfrak{w} + B_u\Delta\mathfrak{w}\right) \approx DR_o\boldsymbol{\nabla}\cdot\boldsymbol{Q}. \tag{3.3.6}$$

At planetary scales, Burger number is small and the rotation dominates the stratification, $\frac{\mathfrak{W}}{\mathfrak{U}} \sim DR_o$. At smaller scales, with a larger Burger number, the stratification dominates the rotation, $\frac{\mathfrak{W}}{\mathfrak{U}} \sim DR_o/B_u$. For the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, the latter is thus more relevant.

Note that the angle between the small-scale component and the horizontal one can be assumed to be constrained by the angle between the isopycnical and the horizontal plane. Invoked to describe baroclinic instabilities theory, this statement helps to specify the anisotropy of the eddy diffusivity (Vallis, 2006). The argument of the orientation of the eddies activity with isentropic surfaces and the related mixing is also supported by several other authors (Gent and Mcwilliams, 1990; Pierrehumbert and Yang, 1993).

In the case of QG models, the large and small Burger scaling cases lead to the same result: the unresolved velocity is mainly horizontal.

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{R_o}{B_u} D \ll D. \tag{3.3.7}$$

This is consistent with the assumption of a large stratification, *i.e.* flat isopycnicals, if we admit that the eddies activity appears preferentially along the isentropic surfaces. As a consequence, the terms $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z$ scale as $\frac{R_o}{B_u}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \cdot \boldsymbol{\nabla}$. In the QG approximation, the scaling of the diffusion and effective advection terms including $\boldsymbol{\sigma}_{z\bullet}$ are one to two orders smaller (in power of R_o/B_u) than terms involving $\boldsymbol{\sigma}_{H\bullet}$. For any function ξ, the vertical diffusion $\partial_z(\frac{\boldsymbol{\sigma}_{z\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z\xi)$ is one order smaller than the horizontal-vertical diffusion term $\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z\xi\right)$ and two orders smaller than the horizontal diffusion term $\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T}{2}\boldsymbol{\nabla}\xi\right)$.

Beta effect

At mid-latitudes, the related term, given by $\beta \triangleq \partial_y f$, is much smaller than the constant part of the Coriolis frequency. Nevertheless, it can govern a large part of the relative vorticity at large scales. The following scaling is thus chosen (Vallis, 2006):

$$\beta y \sim \boldsymbol{\nabla}^\perp \cdot \boldsymbol{u} \sim \frac{U}{L} = R_o f_0. \tag{3.3.8}$$

3.3.2 Stratified Quasi-Geostrophic model under moderate uncertainty

The moderate uncertainty case corresponds to $\Upsilon \sim 1$. Horizontal advective terms and horizontal sub-grid terms are comparable.

Following similar principles as those used to derive the deterministic stratified QG model (Vallis, 2006), a stochastic QG model can be derived (see Appendix 3.B). This QG solution corresponds to the limit of the Boussinesq solution when the Rossby number goes to zero. The resulting potential vorticity (PV), Q, is then found to be conserved, along the horizontal random flow, up to three source terms:

$$\mathbb{D}_t^H Q \quad = \quad \frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\mathrm{d}t \;-\; \frac{1}{2}\boldsymbol{\nabla}\;\cdot\;(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_H f))^T\,\mathrm{d}t \;-\; \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right], \tag{3.3.9}$$

where the QG PV is:

$$Q \triangleq \Delta\psi + f + \left(\frac{f_0}{N}\right)^2\partial_z^2\psi, \tag{3.3.10}$$

ψ is the streamfunction, $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the $\frac{\pi}{2}$ rotation matrix,

$$\boldsymbol{S} = \frac{1}{2}\left[\boldsymbol{\nabla}\boldsymbol{u}^T + (\boldsymbol{\nabla}\boldsymbol{u}^T)^T\right] \text{ and } \boldsymbol{S}_{\sigma\mathrm{d}B_t} = \frac{1}{2}\left[\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T + (\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T)^T\right] \tag{3.3.11}$$

denotes the strain rate tensor of the horizontal resolved and unresolved velocities, $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, respectively. To interpret the source terms, we rather focus on the material derivative of the PV:

$$D_t^H Q \quad = \quad \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_H f))^T\,\mathrm{d}t - \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right], \tag{3.3.12}$$

with

$$\alpha^T \triangleq \sum_i \left(\boldsymbol{\nabla}\boldsymbol{\sigma}_{H_i}^T\right)^2, \tag{3.3.13}$$

which can be decomposed into a symmetric part, positive or negative diffusion of the stream function, and an anti-symmetric part, skew diffusion advection of the stream function. Compared to the traditional QG model, this system includes two smooth (continuous) source/sink terms that depend on the variance tensor, and a random forcing term. The first source term in (3.3.12) is correlated in time and may decrease or increase the PV energy. This term is due to the spatial variations of both the diffusion coefficient and the drift correction. The second term takes into account interactions between the Coriolis frequency, including beta effects, and inhomogeneous sub-grid eddies. The last source term in (3.3.12) is a noise term, encoding the interactions between the resolved and the unresolved strain rate tensors. Uncorrelated in time, this noise increases the potential enstrophy along time.

To further understand this source term, let us denote Ξ and Λ the eigenvalues associated with the stable directions (*i.e.* negative eigenvalue) of the strain rate tensors of the large-scale flow, $\boldsymbol{S}$, and of the small-scale flow, $\boldsymbol{S}_{\sigma \mathrm{d}B_t}$ respectively. We note θ, the angle between these two stable directions

$$-\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] \quad = \quad 2\underbrace{\Xi\Lambda}_{>0}\sin\left(2\theta\right). \tag{3.3.14}$$

The detailed derivation is provided in Appendix 3.B. This random source vanishes when the stable directions of $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$ are aligned or orthogonal. It is maximum and positive (respectively minimum and negative) when there is an angle of $\frac{\pi}{4}$ (respectively $-\frac{\pi}{4}$) between those directions. Around the local position $\boldsymbol{x}$, stable and unstable directions of the large-scale velocity define 2 axes and 4 quadrants. As understood, the strain rate tensor does not depend on the local vorticity. Yet, an hyperbolic deformation will almost resemble a positive vorticity in the upper-left and bottom-right quadrants, and a negative vorticity in the upper-right and bottom-left quadrants. For $\theta = \frac{\pi}{4}$, the stable direction of the small-scale velocity aligns along the upper-left to bottom-right direction. The small-scale velocity then compresses the flow in this direction and dilates the flow in the orthogonal direction (upper-right to bottom-left). The quadrants associated with a seemingly positive (resp. negative) vorticity are brought closer (resp. farther) to $\boldsymbol{x}$. Accordingly, the vorticity increases at $\boldsymbol{x}$. For $-\frac{\pi}{4}$, the vorticity would decrease.

Note the $\mathrm{d}\boldsymbol{B}_t$ factor has been omitted in the right-hand side of equation (3.3.14). This term remains a linear function of the uncorrelated noise $\boldsymbol{z} \mapsto \mathrm{d}\boldsymbol{B}_t(\boldsymbol{z})$. Whatever the angle between the stable directions, the source term always has a zero (ensemble) mean and increases the enstrophy since it is a term in $\mathrm{d}\boldsymbol{B}_t$. Equation (3.3.14) could then be used to define the horizontal inhomogeneous small-scale component of the velocity. If the conservation of PV is a strong constraint, this component can indeed be defined to ensure that its stable direction is always along or orthogonal to the stable direction of $\boldsymbol{u}$.

A two-layer model could also be deduced from equation (3.3.9) or (3.3.12). This would help identifying the stochastic parameterization effects on the barotropic and baroclinic modes. In particular, the particular forms of the operator $\boldsymbol{\sigma}$ able to trigger barotropization effects can bemore efficiently studied.

In the stochastic QG model, the stream function ψ is related to the buoyancy, b, the pressure, p', and the velocity, $\boldsymbol{u}$, by the usual relations:

$$b = f_0\partial_z\psi,\ p' = \rho_b f_0\psi \text{ and } \boldsymbol{u} = \boldsymbol{\nabla}^{\perp}\psi, \tag{3.3.15}$$

where ρ_b is the mean (background) density. The horizontal noise term, $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, appearing in both the horizontal stochastic material derivative and in the 2×2 horizontal variance tensor, $\boldsymbol{a}_H$, is in geostrophic balance with a pressure component uncorrelated in time. Due to their scaling, the vertical noise and its variance are neglected in the final equations.
For homogeneous turbulence conditions, the transport of PV (3.3.12) simplifies. The variance tensor becomes constant, the first two source terms disappear, to give

$$\mathbb{D}_t^H Q = D_t^H Q = -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right]. \tag{3.3.16}$$

The transport of the PV (equation (3.3.9) or (3.3.16)) determines the dynamics of the fluid interior. Boundary conditions are then necessary to specify completely the dynamics.

3.3.3 Surface Quasi-Geostrophic model under moderate uncertainty

A classical choice considers a vanishing solution in the deep ocean and a buoyancy transport at the surface (Vallis, 2006; Lapeyre and Klein, 2006):

$$\psi \underset{z\to-\infty}{\longrightarrow} 0 \text{ and } D_t^H b_{|z=0} = \mathbb{D}_t^H b_{|z=0} = 0. \tag{3.3.17}$$

Assuming zero PV in the interior but keeping these boundary conditions leads to the Surface Quasi-Geostrophic model (SQG) (Blumen, 1978; Held et al., 1995; Lapeyre and Klein, 2006; Constantin et al., 1994, 1999, 2012). Under the stochastic framework, the derivation is similar. The PV is indeed identical to the classical one (see equation (3.3.10)), assuming zero PV in the interior and vanishing solution as $z \to -\infty$ unsurprisingly yields the same SQG relationship:

$$\hat{b} = N\|\boldsymbol{k}\| \ \hat{\psi}. \tag{3.3.18}$$

The top boundary condition, equation (3.3.17), provides an evolution equation, namely the horizontal transport of surface buoyancy, in the stochastic sense:

$$\mathbb{D}_t^H b = 0. \tag{3.3.19}$$

The time-uncorrelated component of the velocity, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is divergence-free. Its inhomogeneous and anisotropic spatial covariance has then to be specified. The time-correlated component of the velocity is also divergence-free, with a stream function specified by the SQG relation (3.3.18). The buoyancy is randomly advected, and the resulting smooth velocity component is random as well.

3.3.4 Summary

For simplified models, stochastic versions are derived for scaling assumptions related to the subgrid terms. For moderate uncertainty, the PV is transported along the random flow up to three source terms. The first one, smooth in time, is due to spatial variations of the inhomogeneous diffusion and the drift correction. The second one, also smooth, encodes the interaction between inhomogeneous turbulence and Coriolis frequency. These terms disappear for an homogeneous turbulence. The last term, a time-uncorrelated multiplicative noise, involves the large-scale and the small-scale strain rate tensors. It is a source of potential enstrophy and its instantaneous value depends on the angle between the large-scale and small-scale stable directions. Assuming zero PV in the interior, a SQG model follows from this QG model.

3.4 Numerical results

We focus on this SQG_{MU} model (3.3.3). A high-resolution deterministic SQG simulation provides a reference. The MATLAB® codes are available online (`http://vressegu.github.io/sqgmu`). Numerical results are analyzed in terms of the resolution gains (when a single realization is simulated) and the potential for ensemble forecasting in estimating spatial and spectral reconstruction errors (for an ensemble of realizations).

3.4.1 Test flow

The initial conditions defining the test flow, Figure 3.1, consist of a spatially smooth buoyancy field with two warm elliptical anticyclones and two cold elliptical cyclones given by:

$$\begin{aligned} b(\boldsymbol{x}, t=0) \quad = \quad & F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) + F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) \\ & -F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right) - F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right), \end{aligned} \tag{3.4.1}$$

with

$$F(\boldsymbol{x}) \triangleq B_0 \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right) \text{ and } \left\{\begin{array}{lcl}\sigma_x & = & 67 \text{ km},\\ \sigma_y & = & 133 \text{ km}.\end{array}\right. \tag{3.4.2}$$

The size of the vortices is of order of the Rossby radius L_d. The buoyancy and the stratification have been set with: $B_0 = 10^{-3} m.s^{-2}$ and $N = 3f_0$. The Coriolis frequency is set to $1.028\times10^{-4}s^{-1}$ (45° N). Periodic boundaries conditions are considered.

The deterministic high-resolution SQG reference model is associated with a spatial mesh grid of 512^2 points, whereas the low-resolution (deterministic or stochastic) SQG models are run on 128^2 points. The simulations have been performed through a pseudo-spectral code in space. As for the temporal discrete scheme the deterministic simulation relies on a fourth-order Runge-Kutta scheme, whereas the stochastic ones are based on an Euler-Maruyama scheme (Kloeden and Platen, 1999). For our application, the weak precision of this scheme is balanced by the use of a small time step. In all the simulations (deterministic and random, high-resolution and low-resolution), a standard hyperviscosity model is used:

$$\mathbb{D}_t b = \alpha^{hv}\Delta^4 b \,\mathrm{d}t, \tag{3.4.3}$$

with a coefficient $\alpha^{hv} = (5\times10^{29} m^8.s^{-1})M_x^{-8}$ where M_x denotes the meshgrid size (*i.e.* 128 or 512).

Figure 3.1 displays the high-resolution buoyancy field at $t = 0, 5, 10, 13, 15, 16, 20$ and 30 days. During the first ten days, the vortices turn with slight deformation. Vortices of the same sign have their tails that draw closer. This creates high shears around four saddle points located at $(x, y) = (0, 250), (500, 250), (0, 750)$ and $(500, 750)$ (in km). A strong non-linearity in the neighborhood of a saddle point has been identified to become a major source of instability (Constantin et al., 1994, 1999, 2012). In our case, this effect is weak but yields an effective creation of turbulence 10 days later. Shears create long and fine filaments, wrapping around the vortices until the 15th day. At this time, the filaments become unstable, break and a so-called "pearl-necklace" appears, characteristic of the SQG model, days 17-18 in the simulation. These small vortices are then ejected from their orbits. Between days 17th and 25th, they interact with the large vortices, the filaments and other small vortices, to create a fully-developed SQG turbulence orbiting around the four large vortices.

3.4.2 Simulation of the random velocity

To simulate the $\mathrm{SQG_{MU}}$ model (3.3.18-3.3.19), the covariance of the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ must be specified. As this unresolved velocity field is assumed divergence-free, we introduce the following stream function linear operator, ψ_σ, and its kernel, $\breve{\psi}_\sigma$:

$$\boldsymbol{\sigma}_H(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\nabla}^\perp\psi_\sigma(\boldsymbol{x})\mathrm{d}B_t, = \int_\Omega \mathrm{d}\boldsymbol{z}\, \boldsymbol{\nabla}^\perp_{\boldsymbol{x}}\breve{\psi}_\sigma(\boldsymbol{x},\boldsymbol{z})\mathrm{d}B_t(\boldsymbol{z}). \tag{3.4.4}$$

As such, a single cylindrical Wiener process, B_t, is sufficient to sample our Gaussian process. This is specific to two-dimensional domains. In 3D, a vector of 3 independent $\mathbb{I}_d$-cylindrical Wiener processes, and a projection operator on the divergence-free vector space or a curl must be considered to simulate an isotropic small-scale velocity (Mémin, 2014). For a divergent unresolved velocity, equation (3.4.4) can additionally involve the gradient of a random potential, $\boldsymbol{\nabla}\tilde{\psi}_\sigma\mathrm{d}B_t$.

Then, similar to the Kraichnan's model, a solenoidal homogeneous field can be considered: (Kraichnan, 1968, 1994; Gawędzki and Kupiainen, 1995; Majda and Kramer, 1999):

$$\boldsymbol{\sigma}_H(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \int_\Omega \mathrm{d}\boldsymbol{z}\, \boldsymbol{\nabla}^\perp_{\boldsymbol{x}}\breve{\psi}_\sigma(\boldsymbol{x}-\boldsymbol{z})\mathrm{d}B_t(\boldsymbol{z}) = \left(\boldsymbol{\nabla}^\perp\breve{\psi}_\sigma \star \mathrm{d}B_t\right)(\boldsymbol{x}). \tag{3.4.5}$$

where $\star$ denotes a convolution. Although spatially inhomogeneous field would be more physically relevant, homogeneity greatly simplifies the random field simulation. Indeed, homogeneity in physical space implies independence between the Fourier modes

$$\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k}) = i\boldsymbol{k}^\perp\widehat{\breve{\psi}_\sigma}(\boldsymbol{k})\hat{\dot{B}}(\boldsymbol{k}), \tag{3.4.6}$$

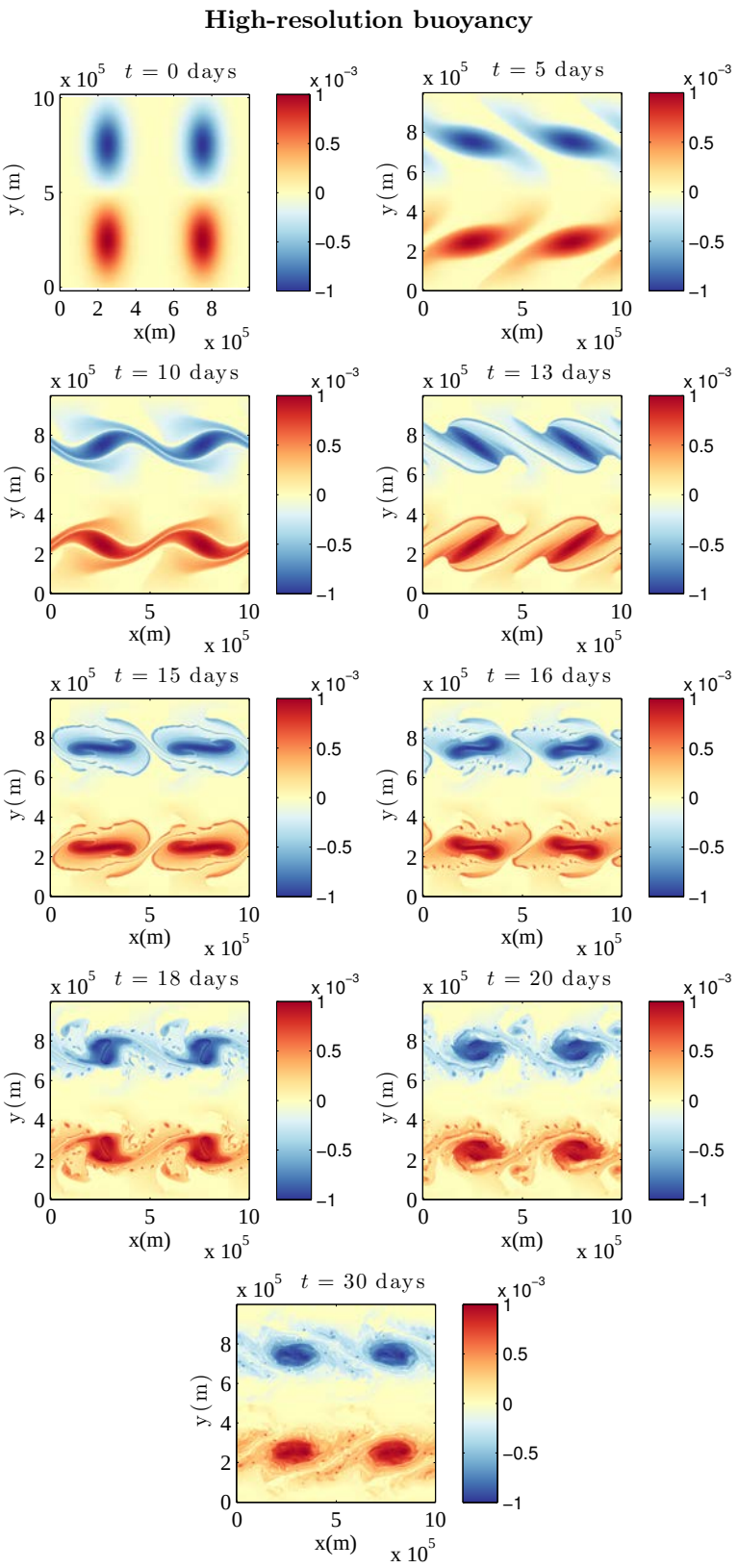


Figure 3.1: Buoyancy ($m.s^{-2}$) at $t = 0, 5, 10, 13, 15, 16, 20, 30$ days of advection for the usual SQG model at resolution 512^2.

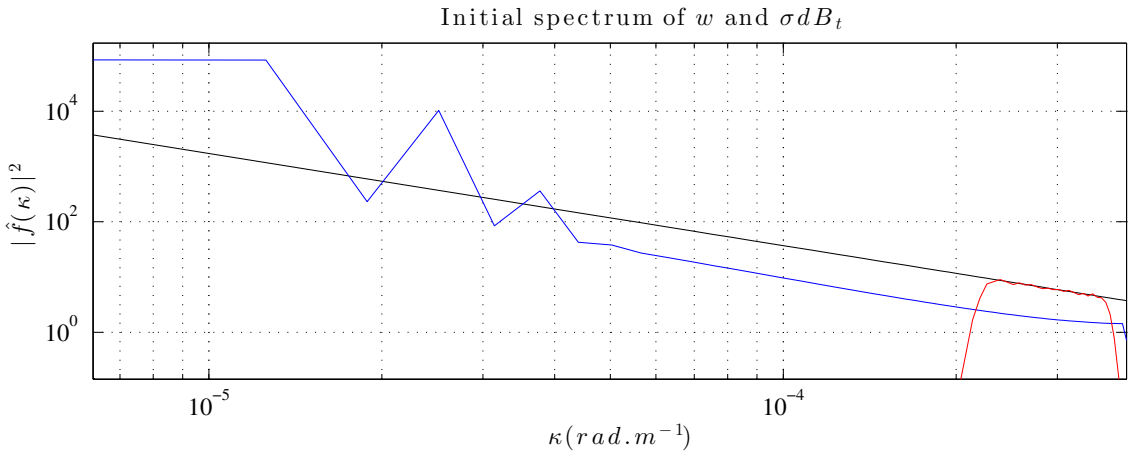


Figure 3.2: Spectrum ($m^2.s^{-2}/(\text{rad}.m^{-1})$) of $\boldsymbol{w}$, at the initial time, in blue, spectrum of $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$ (up to a multiplicative constant), in red, and slope $-\frac{5}{3}$ in black. In the simulation performed, $\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}$ is restricted to a narrow spectral band. Thus, this velocity component almost only acts near the resolution cutoff, where the large-scale component, $\boldsymbol{w}$, has a low energy.

in the half-space $\boldsymbol{k} \in (\mathbb{R} \times \mathbb{R}^{+*}) \cup (\mathbb{R}^+ \times \{0\})$. Thus, the small-scale velocity can be conveniently specified from its omnidirectional spectrum:

$$\boldsymbol{k} \mapsto \frac{1}{\mu(\Omega)}\mathbb{E}\oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}}\|\boldsymbol{k}\| \left\|\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})\right\|^2 = \frac{2\pi}{\Delta t}\|\boldsymbol{k}\|^3 \left|\widehat{\breve{\psi}_\sigma}\left(\|\boldsymbol{k}\|\right)\right|^2, \tag{3.4.7}$$

where $\mu(\Omega)$ is the surface of the spatial domain Ω, $\theta_{\boldsymbol{k}}$ is the angle of the wave-vector $\boldsymbol{k}$ and Δt the simulation time-step. Consistent with SQG turbulence, the omni-directional spectrum slope, denoted s, is fixed to $-\frac{5}{3}$. For 2D Euler equations, the slope would be set to -3. If the small scales spectrum slope is unknown, the spectrum slope of the resolve scales – estimated on line – may enable to specify s through a scale similarity assumption. The unresolved velocity should be energetic only where the dynamics cannot be properly resolved. Consequently, we apply to the spectrum a smooth band-pass filter, f_{BP}, which has non-zero values between two wavenumbers κ_{min} and κ_{max}. The parameter κ_{min} is inversely related to the spatial correlation length of the unresolved component. In practice, we set κ_{max} to the theoretical resolution, $\frac{\pi}{\Delta x}$, and κ_{min} to the effective resolution (hereafter $\kappa_{min} = \kappa_{max}/2$). Figure 3.2 illustrates this spectrum specification. The small scales' energy is specified by the diffusion coefficient a_H and the simulation time step:

$$\mathbb{E}\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)^T = \frac{1}{\Delta t}\boldsymbol{a}_H = \frac{1}{\Delta t}\begin{pmatrix} a_H & 0 \\ 0 & a_H \end{pmatrix}. \tag{3.4.8}$$

The diagonal structure of the variance tensor is due both to incompressiblity and isotropy. The scalar variance tensor, a_H, is similar to an eddy viscosity coefficient. So, a typical value of eddy viscosity used in practice is a good proxy to setup this parameter. Otherwise, this parameter can be tuned. For this paper, it is set to 9 $m^2.s^{-1}$. The time step depends itself, through the CFL conditions, on both the spatial resolution and the maximum magnitude of the resolved velocity. Finally, equation (3.4.6) writes:

$$\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}}\, i\boldsymbol{k}^\perp f_{BP}\left(\|\boldsymbol{k}\|\right)\|\boldsymbol{k}\|^{-\alpha}\,\frac{\widehat{\mathrm{d}B_t}}{\sqrt{\Delta t}}(\boldsymbol{k}) \text{ with } s = 3 - 2\alpha = -\frac{5}{3}, \tag{3.4.9}$$

where A is a constant to ensure $\mathbb{E}\left\|\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right\|^2 = 2\frac{a_H}{\Delta t}$ (see equation (3.4.8) above), $\widehat{\mathrm{d}B}_t$ is the spatial Fourier transform of $\mathrm{d}B_t$, with $\frac{\mathrm{d}B_t}{\sqrt{\Delta t}}$, a discrete scalar white noise process of unit variance in space and time. To sample the small-scale velocity, we first sample $\frac{\mathrm{d}B_t}{\sqrt{\Delta t}}$, to get $\frac{\widehat{\mathrm{d}B_t}}{\sqrt{\Delta t}}$, and finally $\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})$ with the above equation.

3.4.3 Resolution gain on a single simulation

In Figure 3.3, the buoyancy field and its spectrum for low resolution $\mathrm{SQG_{MU}}$ and deterministic SQG simulations are displayed for the day 17th. For the spectrum plots (right column), the slope $-\frac{5}{3}$ is superimposed. While the spectrum tail of the SQG model falls slightly before the stochastic one, the most significant gain is observed in the spatial domain, i.e. in the phase of the tracer. Indeed, the $\mathrm{SQG_{MU}}$ buoyancy field exhibits pearl-necklaces, only obtained at higher resolution. The low-resolved SQG simulation only generates smooth and stable filaments. Though small-scale energy distribution remains similar for both low-resolved models, the phase of the stochastic tracer is more accurate. This may seem surprising since the unresolved velocity, $\boldsymbol{\sigma}_H \dot{\boldsymbol{B}}$, is defined in a loose way, through its spectrum, without prescribing the nature of its phase. However, the noise is multiplicative, and the random forcing, $-(\boldsymbol{\sigma}_H \dot{\boldsymbol{B}}) \cdot \boldsymbol{\nabla} b$, does implicitly take into account the tracer phase.

Note, within the stochastic framework, the diffusion coefficient is explicitly related to the noise variance. If the small-scale velocity is set to a magnitude three times smaller than the one prescribed by the diffusion coefficient $\frac{a_H}{2}$, the tracer field becomes quickly too smooth (see Figure 3.4). Conversely, if the small-scale velocity is set to a magnitude three times larger than dictated by the stochastic transport model, the tracer field becomes rapidly too noisy. This is visible both in the spatial and Fourier spaces (Figure 3.4). The stochastic transport model thus imposes a correct balance between noise and diffusion.

3.4.4 Ensemble forecasts

While single realization of $\mathrm{SQG_{MU}}$ model carries more valuable information than a deterministic SQG formulation at the same resolution, our model further enables to perform ensemble forecasting and filtering. Straightforwardly, an ensemble of independently randomly forced realizations $\{b^{(i)} | i = 1, \cdots, N_e\}$ of tracer b can be simulated according to the SPDE (3.3.19). The probability density function and all the statistical moments of the simulated tracer can then be approximated. For instance, the (ensemble) mean of the buoyancy is a spatio-temporal field defined by:

$$\mathbb{E}(b)(\boldsymbol{x}, t) \approx \hat{\mathbb{E}}(b)(\boldsymbol{x}, t) \triangleq \frac{1}{N_e} \sum_{i=1}^{N_e} b^{(i)}(\boldsymbol{x}, t), \tag{3.4.10}$$

where N_e denotes the ensemble size. This is in essence a Monte-Carlo Markov Chain (MCMC) simulation. The ensemble size is deliberately kept small[2] in order to assess the proposed stochastic framework skills.

We compare the ensemble bias with the estimated error provided by the ensemble itself. The bias corresponds to the discrepancy between the tracer ensemble mean and the SQG simulation at high resolution[3] (512^2).

Our reference is deterministic since the initial condition is perfectly known and the target dynamics is deterministic, as the real ocean dynamics. The partial knowledge of initial conditions is a complementary issue not addressed in this paper. The reference being deterministic, the bias represents both the error of the mean and the mean of the error:

$$\widehat{\mathbb{E}}\{b\} - b^{ref} = \widehat{\mathbb{E}}\{\epsilon\}, \tag{3.4.11}$$

where $\epsilon = b - b^{ref}$ stands for the (random) error. We denote by e the absolute value of this bias. Another error metric could be the Root Mean Square Error (RMSE), $\sqrt{\widehat{\mathbb{E}}\{\epsilon^2\}}$. Yet slightly larger, it is found to have similar spatial and spectral distributions (not shown).

The estimated error, denoted ϵ_{est}, is set to 1.96 times the ensemble standard deviation. This specific value corresponds to the (Gaussian) 95% confidence interval. Although the tracer distribution is not Gaussian, this value provides an accurate conventional error estimate:

$$\epsilon_{est}^2(\boldsymbol{x}, t) = (1.96)^2 \widehat{Var}(b) \triangleq (1.96)^2 \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(b^{(i)} - \hat{\mathbb{E}}(b) \right)^2. \tag{3.4.12}$$

[2] All the random simulations are performed with 200 – 128^2 mesh-size – realizations.

[3] Note this simulation is afterward spatially filtered and subsampled to the same resolution as the ensemble

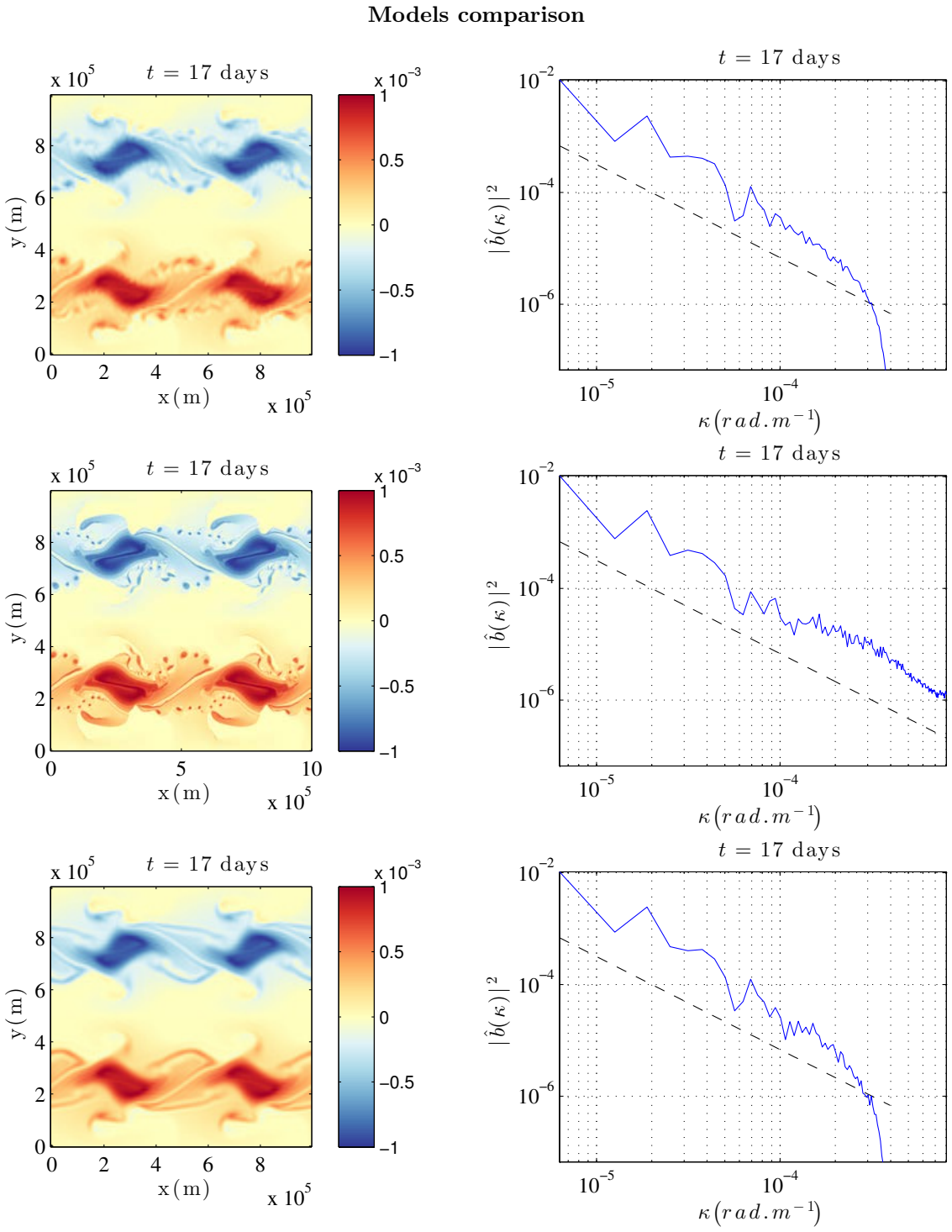


Figure 3.3: Buoyancy ($m.s^{-2}$) and its spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$) at the 17th day of advection for SQG_{MU} at resolution 128^2 (top), SQG at resolution 512^2 (middle) and at resolution 128^2 (bottom). Unlike SQG_{MU}, the low-resolved SQG simulation diffuses the "pearl necklaces", noticeable only at higher resolution.

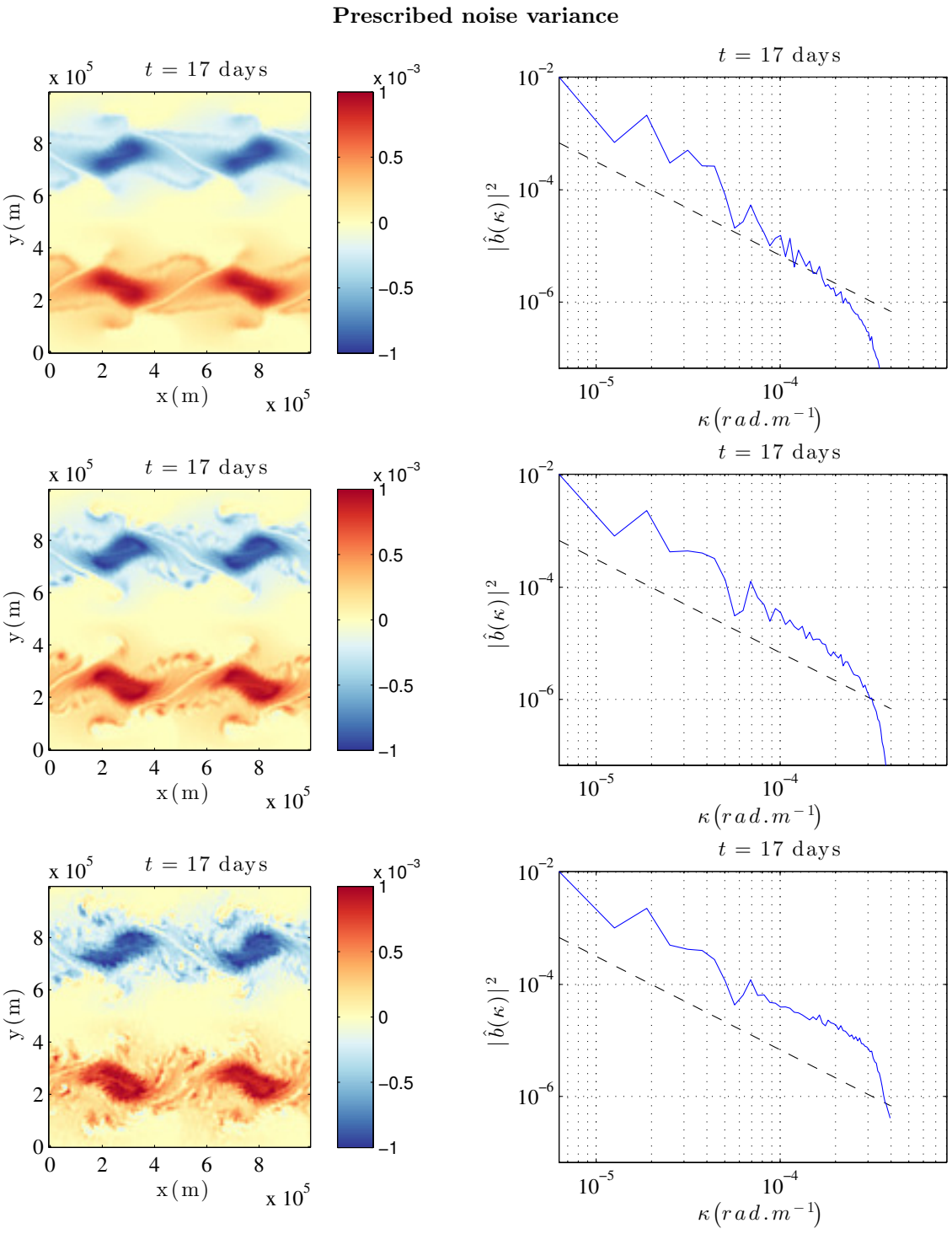


Figure 3.4: Buoyancy ($m.s^{-2}$) and its spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$) at the $t = 17^{th}$ day of advection for the SQG_{MU} model with a small-scale velocity component three times weaker than the one prescribed by the diffusion coefficient a_H (top), with the correct amount of small-scale energy (middle) and a small-scale velocity three times higher than the model diffusion. If the prescribed balance between noise and diffusion is not met the tracer field becomes quickly too smooth or too noisy.

As this error depends on time and space, several comparisons are performed at several distinct times in both the spatial and Fourier domains. In Figures 3.5 and 3.6, the absolute value of spatial fields (3.4.11) and (3.4.12) (*i.e.* e and ϵ_{est}) are compared at days 10, 13, 15, 17, 20 and 25. As obtained, the SQG_{MU} model enables the ensemble to predict the positions and the amplitudes of its own errors with a very good accuracy.

To compare the spread-error consistency of the proposed model, a more classical type of random simulation is considered. An ensemble of the same size is initialized with random perturbations of the initial conditions (3.4.1). The perturbations are assumed to be homogeneous, isotropic, Gaussian and are sampled from a $(-\frac{5}{3})$ spectrum restricted to the small spatial scales, as shown in Figure 3.7. Then, the ensemble is forecast with the deterministic SQG model.

Figures 3.8 and 3.9 represent the spectrum of the errors. The blue and red lines with crosses stand for the spectrum of the bias absolute value, e, of the SQG_{MU} with deterministic initial conditions and of the SQG model with random initial conditions, respectively.

Deterministic and stochastic models have close distribution of errors over the scales, although the SQG_{MU} ensemble mean generally leads to lower errors than the SQG ensemble mean.

The blue line with circles denotes the spectrum of the SQG_{MU} ensemble estimated error, ϵ_{est}. As a benchmark, we superimposed the spectrum of the same estimator, ϵ_{est}, but simulated with the usual model (red curve with circles). This estimation is dramatically underestimated. It is generally one order of magnitude smaller that the real error. To reduce this drawback, a solution would be to multiply by 10 the perturbations of the initial condition. However, this solution introduces strong errors on the realizations (not shown). Their small-scale errors are generally one order of magnitude larger than the ones of our model. These realizations of the deterministic model remain far from the reference for about ten days. On the contrary, the SQG_{MU} predicts the correct spectral distribution of errors at each time, except at very small-scales, and each of its realizations are accurate as shown in the previous subsection. Let us note however that most of the errors are concentrated at large scales.

SQG_{MU} thus appears to provide a relevant ensemble of realizations, as it enables us to estimate the amplitude of its own error with a good accuracy both in the spatial and spectral domains.

With such an ensemble of realizations, it is now possible to analyze the spatio-temporal evolution of the statistical moments. In Figure 3.10, we plotted the ensemble tracer mean and variance for $t = 17, 20$ and 30 days of advection. As expected, the mean field is more smooth than the realizations (see Figure 3.4 for comparison at $t = 17$ days). One realization provides a more realistic field than the mean from a topological point of view. Indeed, the realization exhibits physically relevant small-scale structures. Nevertheless, those structures have uncertain shapes and positions. Therefore, on average, the mean field is closer (in the sense of the norm $\|\bullet\|^2_{L^2(\Omega)}$) to the reference. Besides, those uncertain small-scale structures, forgotten by the mean field, are visible in the variance. The variance becomes significant after 10 days of advection, near the stretched saddle points. The strong tracer gradients create strong multiplicative noises. Indeed, strong large-scale gradients involve smaller scales, and thus interact with the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. Then, at $t = 17$ days, the filament instabilities are triggered by the unresolved velocity stretching effects. The appearance of "pearl necklaces" and the underlying motions of those small-scale eddies are mainly determined by the action of the unresolved velocity component. In consequence, these structures are associated with a high uncertainty in their shapes and locations. Hence, they appear naturally on the variance field. At $t = 20$, those sources of variance remain and mushroom-like structures also develop near $(x, y) = (0, 100), (500, 100), (0, 900)$ and $(500, 900)$ (in km). The evolution of these fronts are uncertain, and also show up in the variance field. On the day 30th, these random structures are transported by the zonal jets which are located at $y = 0$ and $y = 500$ km.

The empirical moments of order 3 and 4 can also be evaluated with the ensemble. A high 4th order moment directly relates to the occurrence of extreme events, which is very relevant for dynamical analysis. The point-wise 4-th order moment is centered and normalized to obtain the so-called kurtosis:

$$m_4 \triangleq \frac{\mathbb{E}\left(b - \mathbb{E}(b)\right)^4}{\left(\mathbb{E}\left(b - \mathbb{E}(b)\right)^2\right)^2}. \tag{3.4.13}$$

The excess kurtosis, $m_4 - 3$ highlights deviations from Gaussianity. In particular, positive values

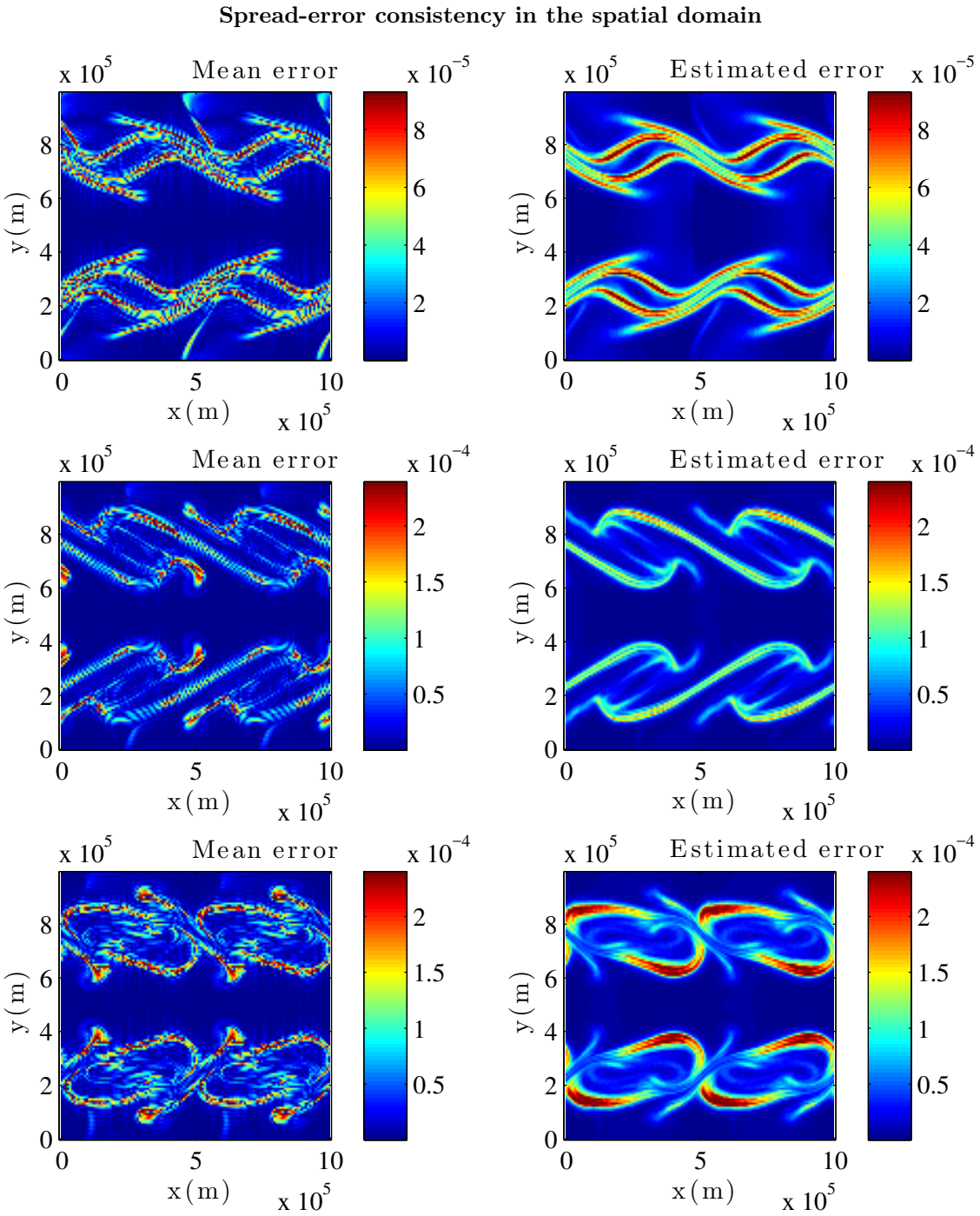


Figure 3.5: Buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, $(m.s^{-2})$ of the $\mathrm{SQG_{MU}}$ model (left) and its estimation, ϵ_{est}, ($1.96\times$ the standard deviation of the ensemble) (right) at resolution 128^2 at (from top to bottom) $t = 10, 13$ and 15 days of advection. The reference is the usual SQG model at resolution 512^2 – adequately filtered and subsampled.

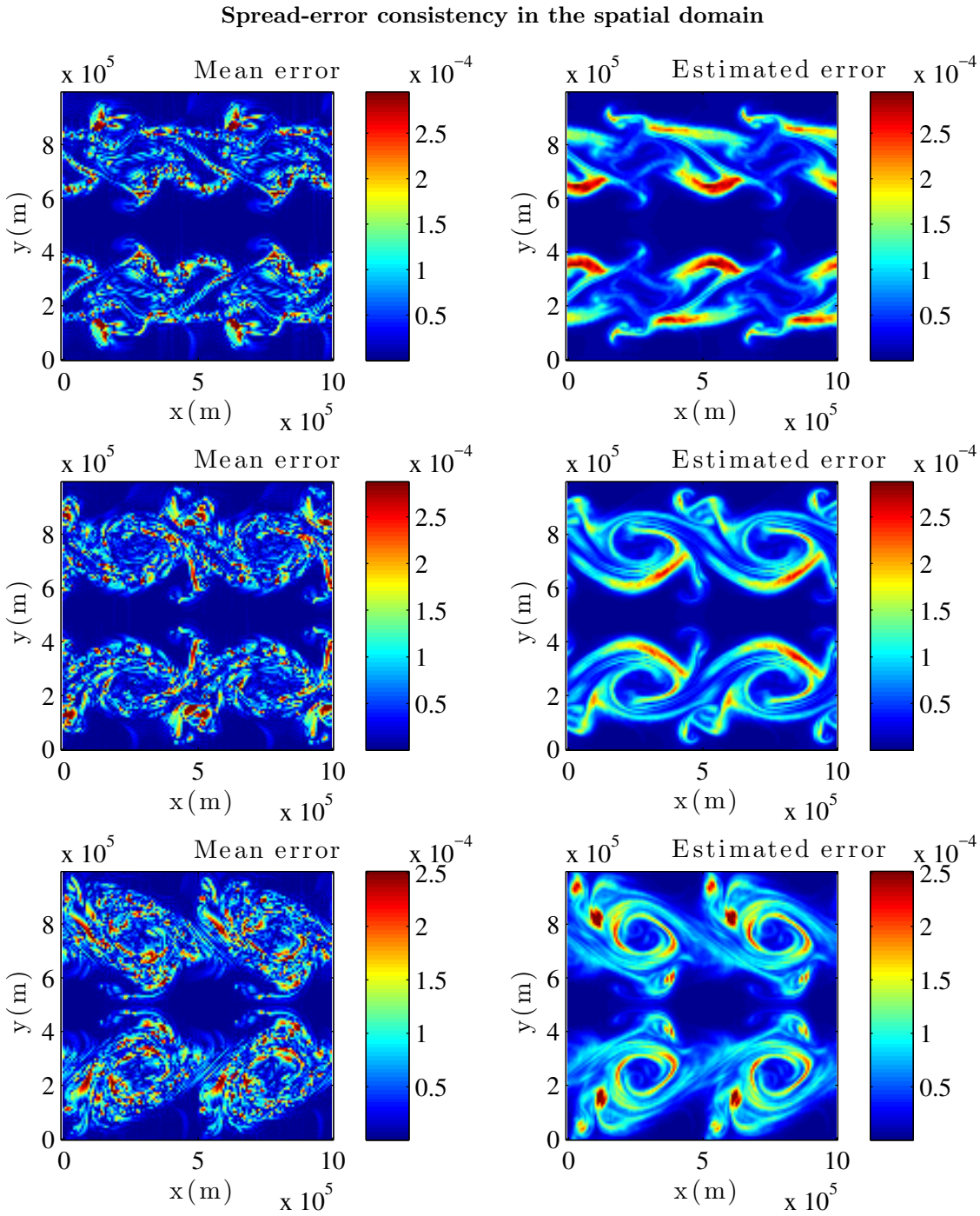


Figure 3.6: Buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, $(m.s^{-2})$ of the $\mathrm{SQG_{MU}}$ model (left) and its estimation, ϵ_{est}, ($1.96\times$ the standard deviation of the ensemble) (right) at resolution 128^2 at (from top to bottom) $t = 17, 20$ and 25 days of advection. The reference is the usual SQG model at resolution 512^2 – adequately filtered and subsampled.

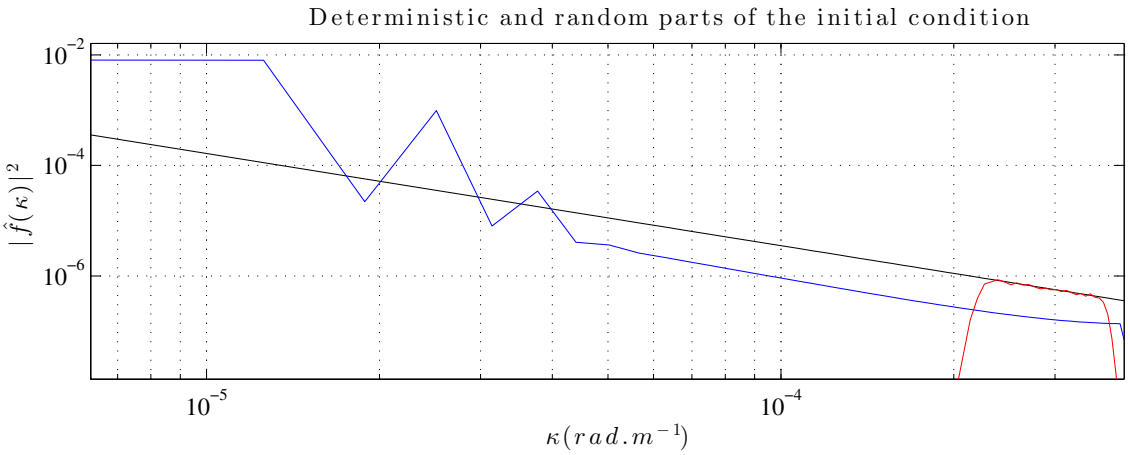


Figure 3.7: Spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$), at the initial time, of the mean buoyancy, in blue, spectrum of its random perturbation, in red, and slope $-\frac{5}{3}$ in black. The initial perturbation is restricted to a narrow spectral band. This random initial condition has been used to simulate an ensemble with the deterministic SQG model.

figure the existence of fat-tail distribution. On the right column of Figure 3.11, the logarithm of the excess kurtosis is displayed for several distinct times. Negative values of the excess kurtosis (which indicates a flatter peak around the mean) have been set to zero. The “pearl necklaces", identified in the variance plots, engender fat-tailed distribution at days $t = 17$ and 20. The small eddies of a “pearl necklace" have similar vorticity and are close to each other, creating high shears between them. A given eddy can be ejected from the necklace by its closest neighbors, and led up to the north or south down. In such a case, the eddy reaches a zone of the space, neither warm nor cold, with weak variability (e.g. with both local mean and variance being low compared to eddy's temperature). This brings extreme tracer values in statistical homogeneous areas. Finally, the random structures, associated with extreme events are trapped in the zonal jets.

The point-wise moment of order 3 marks the asymmetry of the point-wise tracer distribution. The skewness is the third-order moment of the centered and normalized tracer:

$$m_3 \triangleq \frac{\mathbb{E}\left(b - \mathbb{E}(b)\right)^3}{\left(\mathbb{E}\left(b - \mathbb{E}(b)\right)^2\right)^{\frac{3}{2}}}. \tag{3.4.14}$$

Considering the interpretation of excess-kurtosis, the skewness identifies the predominant occurrence of cold (resp. warm) extreme events, associated with the cold (resp. warm) “pearl-necklaces".

3.5 Conclusion

Models under location uncertainty involve a velocity partially time-uncorrelated. Accordingly, the material derivative, the interpretation of conservation laws, and the usual fluid dynamics models are modified. In this paper, the random Boussinesq model is approximated by the so-called QG equations. In our random framework, the approximation depends on sub-grid terms scaling. With moderate turbulent dissipation, the PV is randomly transported in the fluid interior up to three source/sink terms. Two of them are smooth in time and cancel out for homogeneous turbulence. The last forcing term – a random enstrophy source – is related to the angle between stable directions of resolved and unresolved velocities. Similarly to the deterministic case, a uniform PV yields a randomized SQG model, called SQG_{MU}, where the buoyancy is transported in the stochastic sense.

Simulation results are considered for the SQG_{MU} model which is a good representation of the transport under location uncertainty. As such, results are believed to hold for any fluid dynamics models under location uncertainty. As found, SQG_{MU} better resolves small-scale tracer structures

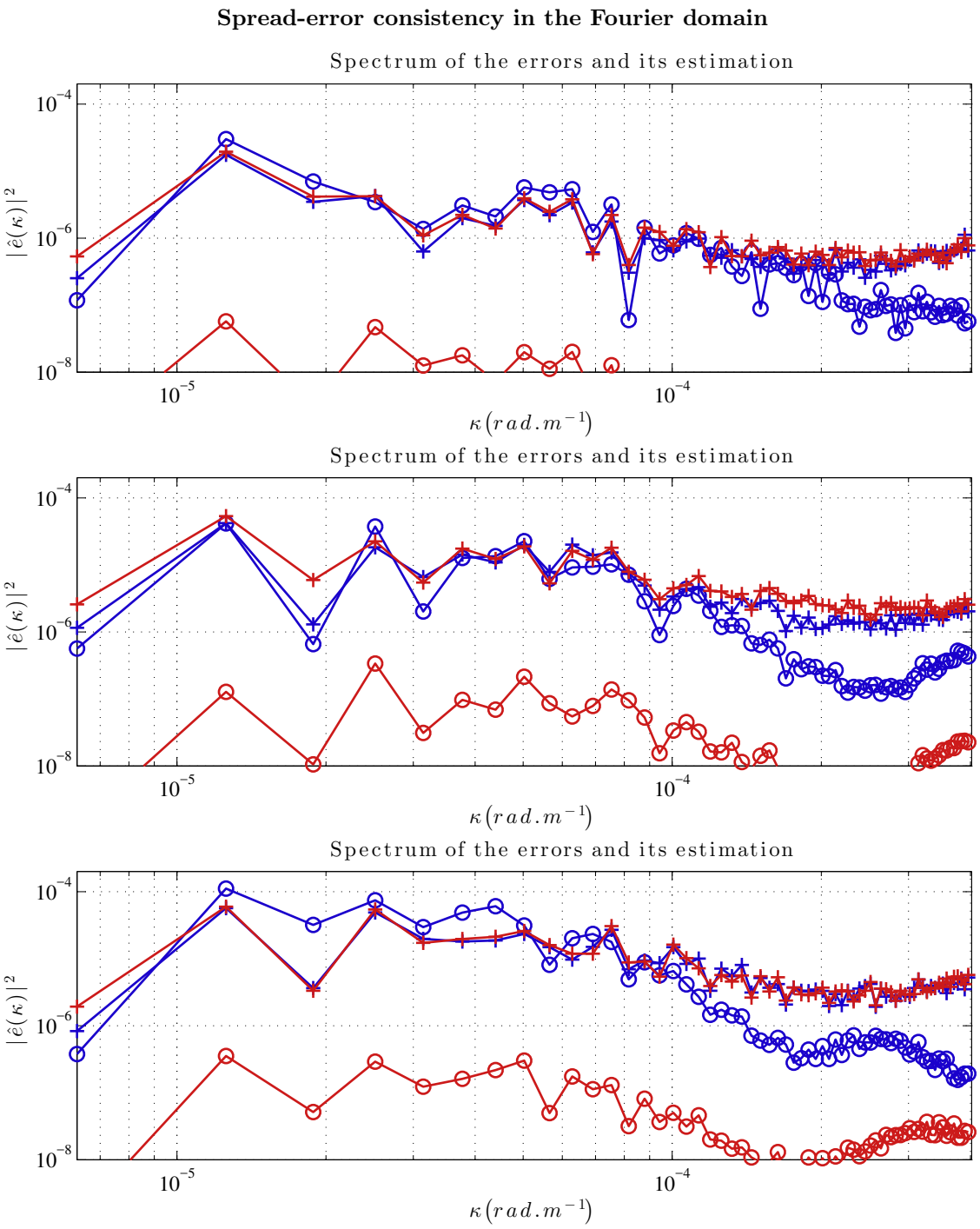


Figure 3.8: Spectrum of the buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, (lines with crosses) and spectrum of the estimated error, ϵ_{est}, (1.96× the standard deviation of the ensemble) (lines with circles) ($m^2.s^{-4}/(\text{rad}.m^{-1})$) of the low-resolution SQG model with random initial conditions (red) and of the SQG_{MU} model at the same resolution (blue), at (from top to bottom) $t = 10, 13$ and 15 days of advection. The reference is the usual SQG model at resolution 512^2– adequately filtered and subsampled. The low-resolution deterministic model with random initial conditions underestimates the error by at least one order of magnitude whereas our estimation is very precise except at small scales.

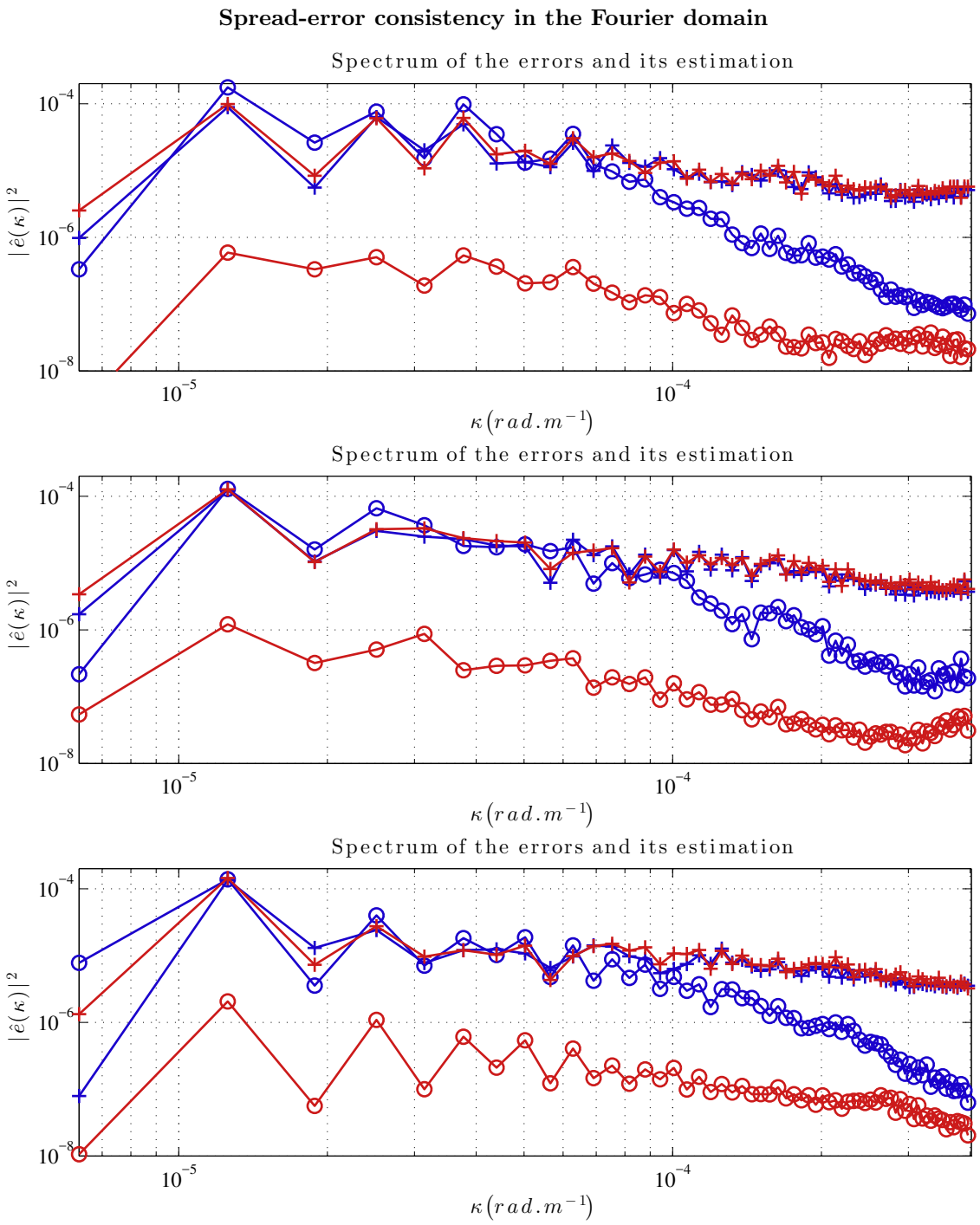


Figure 3.9: Spectrum of the buoyancy bias absolute value, $e = |\widehat{\mathbb{E}}\{b\} - b^{ref}|$, (lines with crosses) and spectrum of the estimated error, ϵ_{est}, (1.96× the standard deviation of the ensemble) (lines with circles) ($m^2.s^{-4}/(\text{rad}.m^{-1})$) of the low-resolution SQG model with random initial conditions (red) and of the SQG_{MU} model at the same resolution (blue), at (from top to bottom) $t = 17, 20$ and 25 days of advection. The reference is the usual SQG model at resolution 512^2– adequately filtered and subsampled. The low-resolution deterministic model with random initial conditions underestimates the error by at least one order of magnitude whereas our estimation is very precise except at small scales.

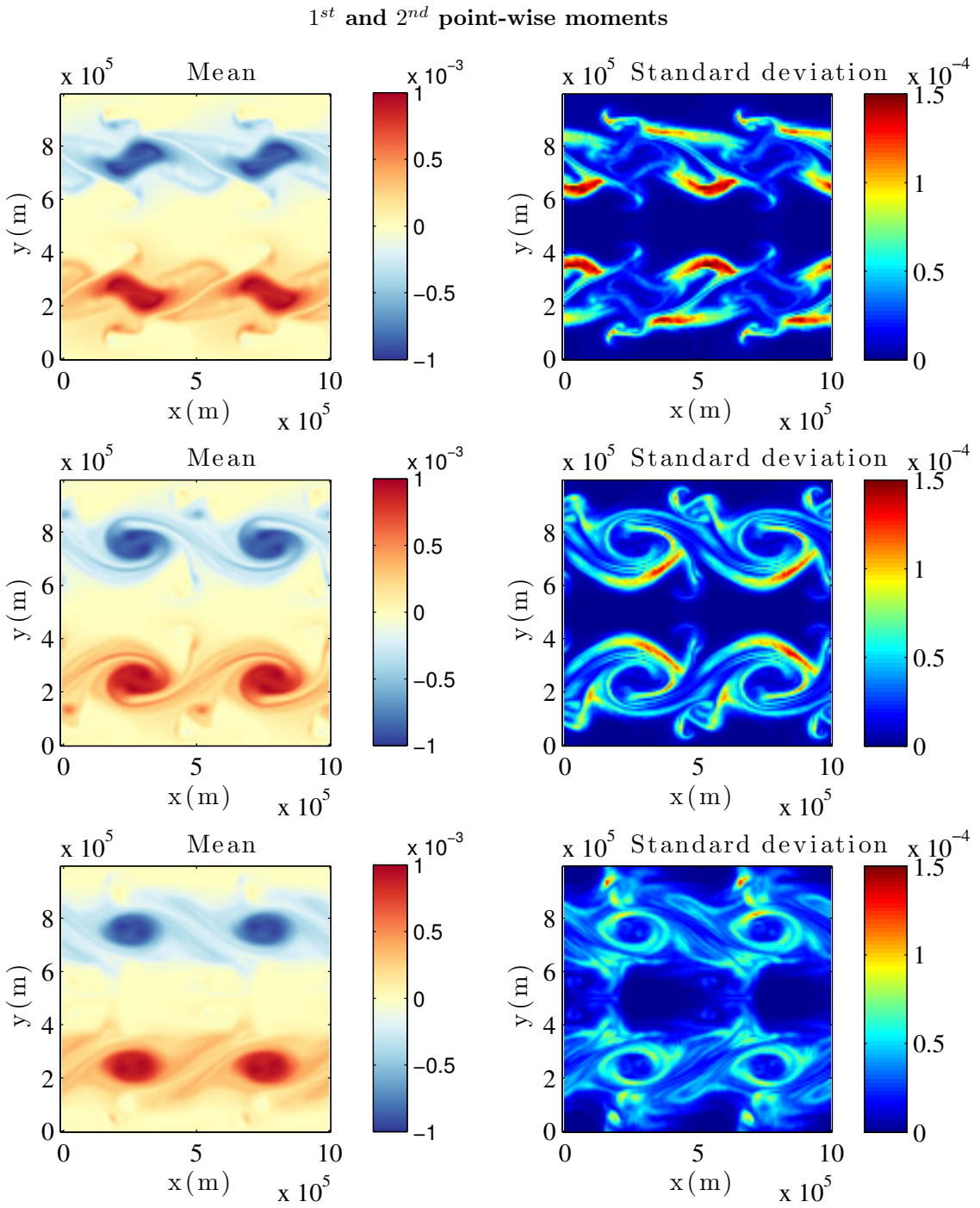


Figure 3.10: Point-wise mean (left) and standard deviation (right) of the buoyancy ($m.s^{-2}$) at $t = 17, 20$ and 30 days of advection for SQG_{MU} model at resolution 128^2. The moments are computed through MCMC simulations. The mean field is smoother than the individual realizations. Areas of higher variance appear first near the stretched saddle points. Then, at $t = 17$ days, the filament instabilities are triggered by the unresolved velocity component. The appearance of "pearl necklaces" can be observed. At $t = 20$, mushrooms-like structures also develop in the variance field near $(x, y) = (0, 100), (500, 100), (0, 900)$ and $(500, 900)$ (in km). At $t = 30$ days, these random structures are transported by the zonal jets.

3rd and 4th point-wise moments

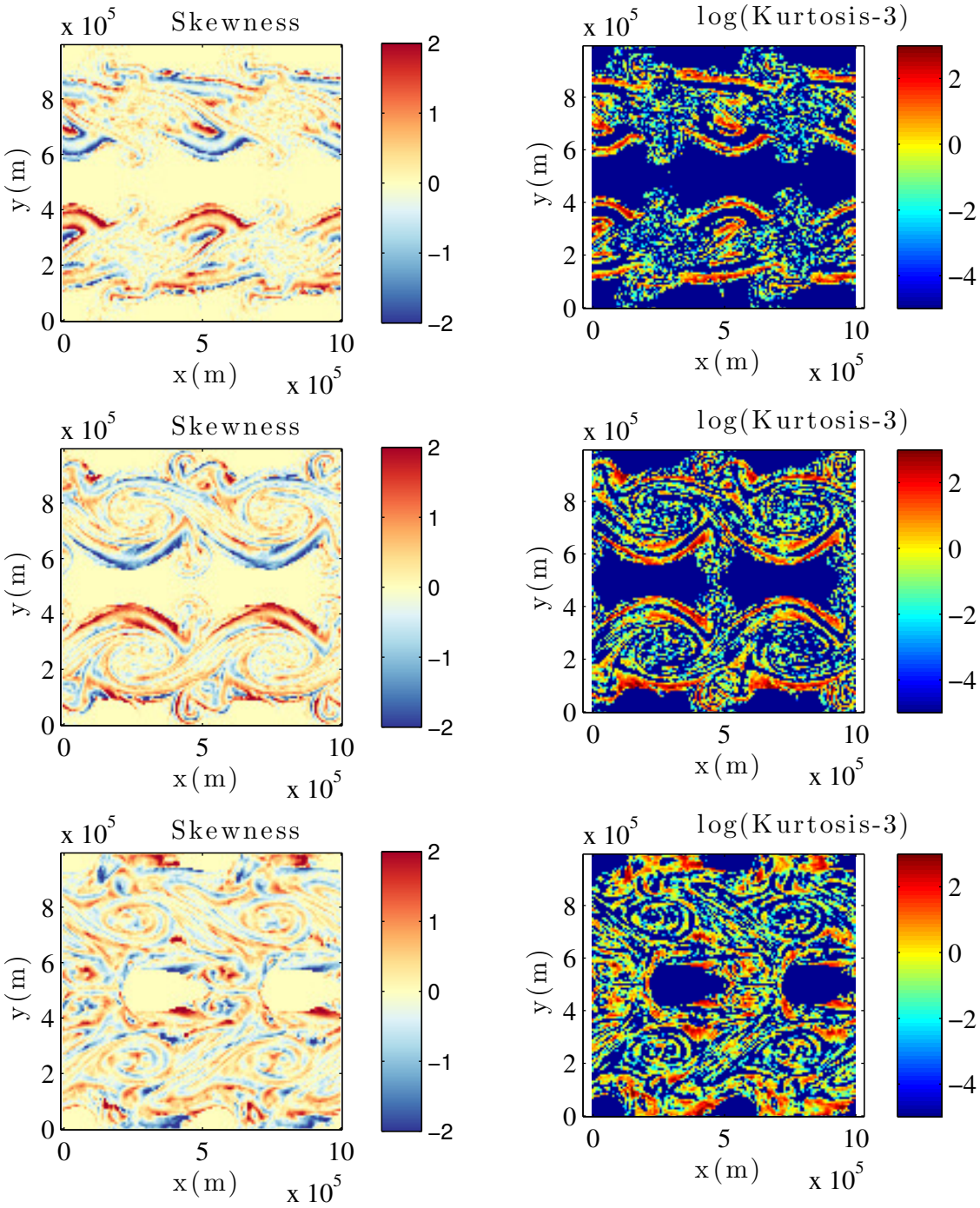


Figure 3.11: Point-wise skewness, and logarithm of the excess kurtosis of the buoyancy at $t = 17, 20$ and 30 days of advection for $\mathrm{SQG_{MU}}$ model at resolution 128^2. The moments are computed through MCMC simulations. Negative excess kurtosis is set to 0. The point-wise law of the tracer is clearly non-Gaussian. The "pearl necklace" events identified in the variance plots leads to fat-tailed distributions with skewness at $t = 17$ and 20 days. The random structures, associated with fat tails are then trapped in the zonal jets.

than a usual SQG model simulated at the same resolution. The prescribed balance between noise and diffusion has also been confirmed. As further highlighted, an ensemble of simulations was able to estimate the amplitude and the position of its own errors in both spatial and spectral domains. This result suggests that the proposed randomized dynamics should be well suited for filtering and other data assimilation methods. On the contrary, a deterministic model with randomized initial conditions, either creates strong errors in its realizations (one order of magnitude larger than the unperturbed deterministic dynamics), or underestimates its own errors (one order of magnitude too low). A MATLAB® code simulating the SQG_{MU} model is available online (`http://vressegu.github.io/sqgmu`).

As a discussion, we can address the problem of uncertainty quantification (UQ) of an unresolved dynamics from an opposite point of view as the usual setting. Instead of specifying a form for the sub-grid velocity, we can wonder what is the optimal form of SPDE for UQ in fluid dynamics. As demonstrated, randomization of initial conditions is far from being sufficient to quantify uncertainty. Therefore, a random forcing is needed to inject randomness at each time step. The simplest choice is a forcing uncorrelated in time. Otherwise, additional stochastic equations need to be simulated to sample a time-correlated process. This is not desirable in high dimension and the correlation time of the process is often small anyway (Berner et al., 2011). A forcing uncorrelated in time is a source of energy. So, to be physically acceptable, the SPDE should involve a dissipative term to exactly compensate this source, even in non-stationary regime. The simplest choices of dissipation are diffusion and linear drag. For small-scale processes, the first is more suitable. Now, what is the form of a noise which brings as much energy as a diffusion removes? The proposed approach constitutes a suitable solution toward this goal.

To further improve the accuracy of the UQ, spatial inhomogeneity of the variance tensor $\boldsymbol{a}$ can be introduced from data or from additional models, as discussed in Resseguier et al. (2017a). This inhomogeneity may reduce possible spurious oscillations of tracer stable isolines. Such oscillations are visible on Figure 3.3 on the sides of the largest vortices. The assumption of time decorrelation may also be a limitation. Nevertheless, as shown by the numerical simulations, the method already achieve very good outcomes with an homogeneous noise component and no real time-scale separation between the resolved and unresolved velocities. Note in particular that since the noise is multiplicative, the random forcing is inhomogeneous even for homogeneous small-scale velocity.

Resseguier et al. (2017c) focuses on a system with a clear time-scale separation between the meso and sub-meso scale dynamics to explore the consequences of the QG assumptions under a strong uncertainty assumption ($\Upsilon \sim R_o$). A zero PV directly appears in the fluid interior and the horizontal velocity becomes divergent. This divergence provides a simple diagnosis of the frontolysis on warm sides of fronts and frontogenesis on cold sides of fronts.

Future works shall also focus on the potential benefits of the stochastic transport for data assimilation issues. As foreseen, the proposed stochastic formalism opens new horizons for ensemble forecasting techniques and other UQ based dynamical approaches (e.g. Ubelmann et al., 2015). This stochastic setup has also been used to characterize chaotic transitions associated with breaking symmetries, also demonstrating interesting perspectives in that context.

Acknowledgments

The authors thank Aurélien Ponte, Jeroen Molemaker and Jonathan Gula for helpful discussions. We also acknowledge the support of the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project.

Appendix

3.A Non-dimensional Boussinesq equations

To derive a non-dimensional version of the Boussinesq equations under location uncertainty (Resseguier et al., 2017a), we scale the horizontal coordinates $\tilde{\boldsymbol{x}}_h = L\boldsymbol{x}_h$, the vertical coordinate $\tilde{z} = hz$, the aspect ratio $D = h/L$ between the vertical and horizontal length scales. A characteristic time $\tilde{t} = Tt$ corresponds to the horizontal advection time U/L with horizontal velocity $\tilde{\boldsymbol{u}} = U\boldsymbol{u}$. A vertical velocity $\tilde{w} = (h/L)Uw$ is deduced from the divergence-free condition. We further take a scaled buoyancy $\tilde{b} = Bb$, pressure $\tilde{\phi}' = \Phi\phi'$ (with the density scaled pressures $\phi' = p'/\rho_b$ and $\mathrm{d}_t\phi_\sigma = \mathrm{d}_t p_\sigma/\rho_b$), and the earth rotation $\boldsymbol{f}^* = f\boldsymbol{k}$. For the uncertainty variables, we consider a horizontal uncertainty $\tilde{\boldsymbol{a}}_H = A_u\, \boldsymbol{a}_H$ corresponding to the horizontal 2×2 variance tensor; a vertical uncertainty vector $\tilde{a}_{zz} = A_w a_{zz}$ and a horizontal-vertical uncertainty vector $\tilde{\boldsymbol{a}}_{Hz} = \sqrt{A_u A_w}\boldsymbol{a}_{Hz}$ related to the variance between the vertical and horizontal velocity components. The resulting non-dimensional Boussinesq system under location uncertainty becomes:

Nondimensional Boussinesq equations under location uncertainty

Momentum equations

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)\boldsymbol{u} + \left(\frac{Ro}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z\boldsymbol{u}\\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}\boldsymbol{u}\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right) + \frac{1}{Ro}\left(1+Ro\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)\\
&\quad = -E_u\,\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned} \tag{3.A.1a}$$

$$\begin{aligned}
&\mathrm{d}_t w + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,w\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)w + \left(\frac{Ro}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z w\\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}w\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right) = \frac{\Gamma}{D^2}b\mathrm{d}t - \frac{E_u}{D^2}\,\partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned} \tag{3.A.1b}$$

Buoyancy equation

$$\begin{aligned}
&\mathrm{d}_t b + \left(\boldsymbol{w}^*_\Upsilon\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)\right)\cdot\boldsymbol{\nabla}b - \frac{1}{2}\frac{1}{\Upsilon}\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{a}_H\boldsymbol{\nabla}b\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right)\\
&\qquad + \frac{1}{(Fr)^2}\frac{1}{\Gamma}\left(w^*_{\Upsilon/2}\mathrm{d}t + \left(\frac{Ro}{B_u}\right)\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right) = 0,
\end{aligned} \tag{3.A.1c}$$

Effective drift

$$\begin{aligned}
\boldsymbol{w}^*_\Upsilon &= \left(\boldsymbol{u}^*_\Upsilon, w^*_\Upsilon\right)^T,\\
&= \left(\left(\boldsymbol{w} - \frac{1}{2\Upsilon}\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right), \left(w - \left(\frac{Ro}{2\Upsilon B_u}\right)\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_{Hz} + O\left(\frac{Ro}{\Upsilon B_u}\right)^2\right)\right)^T,
\end{aligned} \tag{3.A.1d}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0, \tag{3.A.1e}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0, \tag{3.A.1f}$$

$$\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H\right)^T + 2\frac{Ro}{B_u}\boldsymbol{\nabla}_H\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{Ro}{B_u}\right)^2\right) = 0. \tag{3.A.1g}$$

Here, we do not separate the time-correlated components and the time-uncorrelated components in the momentum equations. The terms in $O\left(\frac{Ro}{Bu}\right)$ and $O\left(\frac{Ro}{Bu}\right)^2$ are related to the

time-uncorrelated vertical velocity. These terms are too small to appear in the final QG model ($B_u = O(1)$ in QG approximation) and not explicitly shown. We only make appear the big O approximations. Traditional non-dimensional numbers are introduced : the Rossby number $R_o = U/(f_0 L)$ with f_0 the average Coriolis frequency; the Froude number ($F_r = U/(Nh)$), ratio between the advective time to the buoyancy time; E_u, the Euler number, ratio between the pressure force and the inertial forces, $\Gamma = Bh/U^2 = D^2BT/W$ the ratio between the mean potential energy to the mean kinetic energy. To scale the buoyancy equation, the ratio between the buoyancy advection and the stratification term has also been introduced:

$$\frac{B/T}{N^2W} = \frac{B}{N^2h} = \frac{U^2}{N^2h^2}\frac{Bh}{U^2} = {Fr}^2\Gamma. \tag{3.A.2}$$

Besides those traditional dimensionless numbers, this system introduces Υ, relating the large-scale kinetic energy to the energy dissipated by the unresolved component:

$$\Upsilon = \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{3.A.3}$$

3.B QG model under moderate uncertainty

Hereafter, we consider the QG approximation ($R_o \ll 1$ and $B_u \sim 1$), for $\Upsilon \sim 1$. We focus on solutions of the Boussinesq model with Rossby number going to zero. To derive the evolution equations corresponding to this limit, the solution of the non-dimentional Boussinesq model (Appendix 3.A) is developed as a power series of the Rossby number:

$$\begin{pmatrix} \boldsymbol{w} \\ b \\ \phi \end{pmatrix} = \sum_{k=0}^{\infty} {R_o}^k \begin{pmatrix} \boldsymbol{w}_k \\ b_k \\ \phi_k \end{pmatrix}. \tag{3.B.1}$$

According to the horizontal momentum equation (3.A.1a), the scaling of the pressure still corresponds to the usual geostrophic balance. This sets the Euler number as:

$$E_u \sim \frac{1}{R_o}. \tag{3.B.2}$$

For the ocean, the aspect ratio, $D \triangleq H/L$, is small and $D^2 \ll 1$. As a consequence,

$$\frac{D^2}{E_u} \sim D^2 R_o \ll D^2 \ll 1 \text{ and } \frac{D^2}{E_u\Upsilon} \sim \frac{D^2 R_o}{\Upsilon} \leqslant D^2 \ll 1. \tag{3.B.3}$$

Therefore, the inertial and diffusion terms are negligible in the vertical momentum equation. The hydrostatic assumption is still valid. This leads to the classical QG scaling of the buoyancy equation:

$$\Gamma \sim E_u \sim \frac{1}{R_o} \text{ and } \frac{1}{{F_r}^2\Gamma} \sim \frac{R_o}{{F_r}^2} = \frac{B_u}{R_o}. \tag{3.B.4}$$

In the following, the subscript H is omitted for the differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ. They all represent 2D operators. Only keeping terms of order 0 and 1, we get the following system:

Momentum equations

$$R_o\left(\mathrm{d}_t\boldsymbol{u}+(\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{u}-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\big(a_{ij}\boldsymbol{u}\big)\mathrm{d}t+O\left(\frac{R_o}{\Upsilon B_u}\right)\right)$$
$$+\left(1+R_o\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)=-\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{3.B.5}$$

$$b\,\mathrm{d}t+O\left(R_oD^2\right)=\partial_z\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{3.B.6}$$

Buoyancy equation

$$\frac{R_o}{B_u}\left(\mathrm{d}_tb+\boldsymbol{\nabla}b\cdot\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)_H\right)+\partial_zb\;w\mathrm{d}t-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}b\right)\mathrm{d}t\right)$$
$$+w\mathrm{d}t-\frac{1}{\Upsilon}\frac{R_o}{B_u}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t+\frac{R_o}{B_u}\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z+O\left(\frac{{R_o}^2}{\Upsilon{B_u}^2}\right)=0, \tag{3.B.7}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}+\partial_zw=0, \tag{3.B.8}$$

$$\boldsymbol{\nabla}\cdot\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_H+\frac{R_o}{B_u}\partial_z\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_z=0, \tag{3.B.9}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T+2\frac{R_o}{B_u}\boldsymbol{\nabla}\cdot\partial_z\boldsymbol{a}_{Hz}+O\left(\left(\frac{R_o}{B_u}\right)^2\right)=0. \tag{3.B.10}$$

The thermodynamic equation (3.B.7) at 0 order leads to :

$$w_0=0, \tag{3.B.11}$$

and then, by the large-scale incompressibility equation (3.B.8), the 0-order horizontal velocity is divergence-free. Following the scaling assumption, the horizontal small-scale velocity is also divergence-free (3.B.9). The horizontal momentum equation (3.B.5) at the 0-th order leads to:

$$\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp\phi_0'\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\boldsymbol{\nabla}^\perp\mathrm{d}_t\phi_\sigma, \tag{3.B.12}$$

where time-correlated and time-uncorrelated components have been separated by the mean of uniqueness of the semi-martingale decomposition (Kunita, 1997). Being divergent-free, both components can be expressed with two stream functions ψ_0 and $\mathrm{d}_t\psi_\sigma$:

$$\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp\psi_0\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\boldsymbol{\nabla}^\perp\mathrm{d}_t\psi_\sigma, \tag{3.B.13}$$

exactly corresponding to the dimensionless pressure terms:

$$\psi_0=\phi_0'\text{ and }\mathrm{d}_t\psi_\sigma=\mathrm{d}_t\phi_\sigma. \tag{3.B.14}$$

Deriving these equations along z and introducing the hydrostatic equilibrium (3.B.6) – decomposed between correlated and uncorrelated components – yields the classical thermal wind balance at large-scale for the 0-th order terms. The buoyancy variable does not involve any white noise term, and the small-scale random velocity is thus almost constant along z, as

$$\partial_z\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp b_0\text{ and }\partial_z(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=O\left(R_oD^2\right). \tag{3.B.15}$$

Accordingly the variance tensor scales as:

$$\forall i,j\in H,\;\partial_za_{ij}=O\left({R_o}^2D^4\right), \tag{3.B.16}$$

which is negligible in all equations, and the uncertain random field solely depends on the horizontal coordinates. Since $R_o/B_u\sim R_o$, the 1-st order term of the buoyancy equation must be kept to describe the evolution of b_0:

$$\frac{1}{B_u}\mathbb{D}^H_{0t}b_0+w_1\mathrm{d}t-\frac{1}{\Upsilon}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z=0, \tag{3.B.17}$$

where, for all functions h,

$$\mathbb{D}^H_{0t} h \triangleq \mathrm{d}_t h + \boldsymbol{\nabla} h \cdot \left(\boldsymbol{u}_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial^2_{ij} \left(a_{ij} h \right) \mathrm{d}t. \tag{3.B.18}$$

Taking the derivative along z leads to:

$$\frac{1}{B_u} \left(\mathbb{D}^H_{0t} \partial_z b_0 + \boldsymbol{\nabla} b_0 \cdot \partial_z \left(\boldsymbol{u}_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} d\boldsymbol{B}_t)_H \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial^2_{ij} \left(\partial_z a_{ij} b_0 \right) \mathrm{d}t \right)$$
$$+ \partial_z w_1 \mathrm{d}t - \frac{1}{\Upsilon} \left(\boldsymbol{\nabla} \cdot \partial_z a_{Hz} \right)^T \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} \partial_z (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_z = 0. \tag{3.B.19}$$

The introduction of the thermal wind equations (3.B.15) and incompressibility conditions (3.B.8-3.B.10) helps simplifying this equation as:

$$\frac{1}{B_u} \mathbb{D}^H_{0t} \partial_z b_0 - \boldsymbol{\nabla} \cdot \boldsymbol{u}_1 \mathrm{d}t + \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{\Upsilon^{1/2}} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H = 0. \tag{3.B.20}$$

Note the factor $\left(\frac{R_o}{B_u} \right)^{-1}$ appears. It comes from the incompressible conditions (3.B.9) and (3.B.10), leading $\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$ to both scale as $\frac{R_o}{B_u}$. The hydrostatic balance at 0-order links the buoyancy to the pressure, and then to the stream function

$$\partial_z b_0 = \partial^2_z \phi_0 = \partial^2_z \psi_0. \tag{3.B.21}$$

The 1-st order term of the vertical velocity is not known. Yet, the system can be closed using the vorticity equation at order 1:

$$\boldsymbol{\nabla}^\perp \cdot \left(\mathbb{D}^H_{0t} \boldsymbol{u}_0 \right) + \left(\boldsymbol{\nabla} \cdot \boldsymbol{u}_1 + \left(\frac{R_o}{B_u} \right)^{-1} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) + \boldsymbol{\nabla}(\beta y) \cdot \left(\boldsymbol{u}_0 \mathrm{d}t + (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) = 0, \tag{3.B.22}$$

where the divergence terms come from the constant Coriolis term.

Again, factors $\left(\frac{R_o}{B_u} \right)^{-1}$ compensate the order of magnitude of $\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$. Then,

$$\mathbb{D}^H_{0t} \left(\Delta \psi_0 \right) + \boldsymbol{\nabla} \cdot u_1 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} \left(\frac{R_o}{B_u} \right)^{-1} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H + \beta \left(v_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_y \right)$$
$$+ \frac{1}{\Upsilon^{1/2}} \mathrm{tr} \left(\boldsymbol{\nabla}^\perp (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)^T_H \boldsymbol{\nabla} \boldsymbol{u}_0^T \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial^2_{ij} \left(\boldsymbol{\nabla}^\perp a_{ij} \cdot \boldsymbol{u}_0 \right) \mathrm{d}t = 0. \tag{3.B.23}$$

To make appear the transport of PV, we note that:

$$\begin{aligned} \mathbb{D}^H_{0t} \left(1 + \beta y \right) &= -\beta \left(v_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_y \right) + \boldsymbol{\nabla} \cdot a_{Hy} \beta \mathrm{d}t \\ &\quad - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{2\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t. \end{aligned} \tag{3.B.24}$$

Then, using (3.B.20), (3.B.21) and (3.B.23), we get:

$$\mathbb{D}^H_{0t} \left(\Delta \psi_0 + 1 + \beta y + \frac{1}{B_u} \partial^2_z \psi_0 \right) = -\boldsymbol{\nabla} \cdot a_{Hy} \beta \mathrm{d}t - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{2\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$$
$$- \mathrm{tr} \left(\boldsymbol{\nabla}^\perp (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)^T_H \boldsymbol{\nabla} \boldsymbol{u}_0^T \right) + \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial^2_{ij} \left(\boldsymbol{\nabla}^\perp a_{ij} \cdot \boldsymbol{u}_0 \right) \mathrm{d}t. \tag{3.B.25}$$

We recall that coefficients $\left(\frac{R_o}{B_u}\right)^{-1}$ are still present since

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \sim \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H)^T\,\mathrm{d}t \sim \left(\frac{R_o}{B_u}\right). \tag{3.B.26}$$

If we rewrite the equation with dimensional quantities, the evolution equation for $u_0 = \lim_{Ro\to 0} u$ is obtained (dropping the index 0 for clarity):

$$\mathbb{D}_t^H Q = -\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right) + \frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_H f))^T\,\mathrm{d}t, \tag{3.B.27}$$

where Q is the QG potential vorticity:

$$Q \triangleq \Delta\psi + f + \left(\frac{1}{N}\right)^2\partial_z^2\psi. \tag{3.B.28}$$

Note, (3.B.12) provides the geostrophic balance for the small-scale velocity component. To express the material derivative of Q, the noise term is expanded:

$$-\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right) = -\sum_{k,j\in H}\partial_{kj}^2\psi\partial_k\boldsymbol{\sigma}_{j\bullet}\mathrm{d}\boldsymbol{B}_t. \tag{3.B.29}$$

According to Resseguier et al. (2017a), the difference between the material derivative, D_tQ, and the stochastic transport operator $\mathbb{D}_tQ$, is a function of the time-uncorrelated forcing:

$$\left\{\begin{array}{lcl}\mathbb{D}_tQ &=& f_1\mathrm{d}t + \boldsymbol{h}_1^T\mathrm{d}\boldsymbol{B}_t,\\ D_tQ &=& f_2\mathrm{d}t + \boldsymbol{h}_2^T\mathrm{d}\boldsymbol{B}_t,\end{array}\right. \iff \left\{\begin{array}{lcl} f_2 &=& f_1 + \mathrm{tr}\big((\boldsymbol{\sigma}^T\boldsymbol{\nabla})\,\boldsymbol{h}_1^T\big),\\ \boldsymbol{h}_1 &=& \boldsymbol{h}_2.\end{array}\right. \tag{3.B.30}$$

The expression of h_1 is given by equation (3.B.27) and the above formulas give:

$$D_tQ - \mathbb{D}_tQ = \sum_{i\in H}\boldsymbol{\sigma}_{i\bullet}\partial_i\left(-\sum_{j,k\in H}\partial_k\boldsymbol{\sigma}_{j\bullet}\partial_{kj}^2\psi\right)^T, \tag{3.B.31}$$

$$= -\sum_{i,j,k\in H}\left(\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_{kj}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ijk}^3\psi\right). \tag{3.B.32}$$

With the use of the small-scale incompressibility, we obtain:

$$\frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ik}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_{jk}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ijk}^3\psi\right). \tag{3.B.33}$$

From (3.B.32) and (3.B.33), it yields:

$$D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ik}^2\psi\right). \tag{3.B.34}$$

Denoting, α, the following matrix

$$\alpha_{ij} \triangleq \sum_{k\in H}\partial_k\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}_{k\bullet}^T = \sum_{k\in H}\partial_k(\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}_{k\bullet}^T), \tag{3.B.35}$$

we have

$$\boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{ik}\psi\right), \tag{3.B.36}$$

$$= D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H}\partial^2_{ij}\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right), \tag{3.B.37}$$

and the material derivative of the PV finally reads:

$$D^H_tQ = \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_Hf))^T\,\mathrm{d}t - \mathrm{tr}\left[\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H\boldsymbol{\nabla}\boldsymbol{u}^T\right]. \tag{3.B.38}$$

To note, the transpose of the matrix α has a compact expression:

$$\alpha^T = \sum_p\left(\boldsymbol{\nabla}\boldsymbol{\sigma}^T_{Hp}\right)^2. \tag{3.B.39}$$

To better assess the role of the random source term (the last term of (3.B.38)), it is decomposed in terms of symmetric and anti-symmetric parts of the small-scale/large-scale deformation tensors. Let us denote $\boldsymbol{S}$ and $\boldsymbol{S}_{\sigma\mathrm{d}B_t}$ the symmetric parts of $\boldsymbol{\nabla}\boldsymbol{u}^T$ and $\boldsymbol{\nabla}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H$, respectively. Associated with divergence-free velocities, these symmetric parts, so-called strain rate tensors, have zero trace. Terms $-\frac{1}{2}\omega\boldsymbol{J}$ and $-\frac{1}{2}\omega_{\sigma\mathrm{d}B_t}\boldsymbol{J}$ will stand for the anti-symmetric parts, where $\boldsymbol{J} = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ is the $\frac{\pi}{2}$ rotation. The factors ω and $\omega_{\sigma\mathrm{d}B_t}$ are the large-scale and the small-scale components of the vorticity, respectively. Using $\boldsymbol{J}\boldsymbol{J} = -\mathbb{I}_d$ and $\mathrm{tr}[\boldsymbol{M}\boldsymbol{N}] = tr[\boldsymbol{N}\boldsymbol{M}]$ yields:

$$-\mathrm{tr}\left[\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H\boldsymbol{\nabla}\boldsymbol{u}^T\right] = -\mathrm{tr}\left[\boldsymbol{J}\left(\boldsymbol{S}_{\sigma\mathrm{d}B_t} - \frac{1}{2}\omega_{\sigma\mathrm{d}B_t}\boldsymbol{J}\right)\left(\boldsymbol{S} - \frac{1}{2}\omega\boldsymbol{J}\right)\right], \tag{3.B.40}$$

$$= -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right] - \frac{1}{2}\omega_{\sigma\mathrm{d}B_t}\underbrace{\mathrm{tr}\left[\boldsymbol{S}\right]}_{=0}$$

$$-\frac{1}{2}\,\omega\underbrace{\mathrm{tr}\left[\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right]}_{=0} + \frac{1}{4}\,\omega\,\omega_{\sigma\mathrm{d}B_t}\underbrace{\mathrm{tr}\left[\boldsymbol{J}\right]}_{=0}, \tag{3.B.41}$$

$$= -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right]. \tag{3.B.42}$$

This term thus only depends on the strain rate tensors of $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$. The PV transport can thus be rewritten as:

$$D^H_tQ = \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_Hf))^T\,\mathrm{d}t - \mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right]. \tag{3.B.43}$$

The noise term can be further expressed using the stable directions of the flows defined by $\boldsymbol{u}$ and $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H$, respectively. In the following, we will omit writing the $\mathrm{d}\boldsymbol{B}_t$ factor. The two strain rate tensors are decomposed in orthogonal basis:

$$\boldsymbol{S} = \boldsymbol{V}\boldsymbol{\Xi}\boldsymbol{V}^T = \sum_{p=1}^{2}\Xi_{pp}\boldsymbol{V}_{\bullet p}\boldsymbol{V}^T_{\bullet p} \text{ and } \boldsymbol{S}_{\sigma\mathrm{d}B_t} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^T, \tag{3.B.44}$$

where $\boldsymbol{V}^T_{\bullet p}\boldsymbol{V}_{\bullet q} = \boldsymbol{W}^T_{\bullet p}\boldsymbol{W}_{\bullet q} = \delta_{pq}$, $\Xi_{11} = -\Xi_{22} < 0$ and $\Lambda_{11} = -\Lambda_{22} < 0$.

$$-\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma\mathrm{d}B_t}\right] = -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\mathrm{tr}\left[\boldsymbol{V}_{\bullet p}\boldsymbol{V}^T_{\bullet p}\boldsymbol{J}\boldsymbol{W}_{\bullet q}\boldsymbol{W}^T_{\bullet q}\right], \tag{3.B.45}$$

$$= -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\left(\boldsymbol{V}^T_{\bullet p}\boldsymbol{W}_{\bullet q}\right)\left(\boldsymbol{V}^T_{\bullet p}\boldsymbol{J}\boldsymbol{W}_{\bullet q}\right), \tag{3.B.46}$$

$$= -\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\,\cos(\theta_{pq})\cos\left(\theta_{pq} + \frac{\pi}{2}\right), \tag{3.B.47}$$

$$= \frac{1}{2}\sum_{p,q=1}^{2}\Xi_{pp}\Lambda_{qq}\,\sin(2\theta_{pq}), \tag{3.B.48}$$

where $\theta_{pq} \triangleq (\widehat{\boldsymbol{V}_{\bullet p}, \boldsymbol{W}_{\bullet q}})$ is the angle between $\boldsymbol{V}_{\bullet p}$ and $\boldsymbol{W}_{\bullet q}$. Using the relations between the eigenvalues and the orthogonality of the eigenvectors, it finally comes:

$$\begin{aligned} -\mathrm{tr}\left[\boldsymbol{S}\boldsymbol{J}\boldsymbol{S}_{\sigma \mathrm{d}B_t}\right] &= \frac{1}{2}\Xi_{11}\Lambda_{11}\left(\sin\left(2\theta_{11}\right) - \sin\left(2\left(\theta_{11} - \frac{\pi}{2}\right)\right) - \sin\left(2\left(\theta_{11} + \frac{\pi}{2}\right)\right) + \sin\left(2\theta_{11}\right)\right), \\ &= 2\underbrace{\Xi_{11}\Lambda_{11}}_{>0}\sin\left(2\theta_{11}\right). \end{aligned} \tag{3.B.49}$$

Chapter 4

Geophysical flows under location uncertainty, Part III SQG and frontal dynamics under strong turbulence conditions

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017c

Abstract

Models under location uncertainty are derived assuming that a component of the velocity is uncorrelated in time. The material derivative is accordingly modified to include an advection correction, inhomogeneous and anisotropic diffusion terms and a multiplicative noise contribution. This change can be consistently applied to all fluid dynamics evolution laws. This paper continues to explore benefits of this framework and consequences of specific scaling assumptions. Starting from a Boussinesq model under location uncertainty, a model is developed to describe a mesoscale flow subject to a strong underlying submesoscale activity. Specifically, turbulent diffusion and rotation effects have similar orders of magnitude. As obtained, the geostrophic balance is modified and the Quasi-Geostrophic (QG) assumptions remarkably lead to a zero Potential Vorticity (PV). The ensuing Surface Quasi-Geostrophic (SQG) model provides a simple diagnosis of warm frontolysis and cold frontogenesis.
Keywords: stochastic subgrid tensor, uncertainty quantification, upper ocean dynamics.

4.1 Introduction

Quasi-Geostrophic (QG) models are standard models to study mesoscale barotropic and baroclinic dynamics. Assuming uniform Potential Vorticity (PV) in the fluid interior, the Surface Quasi-Geostrophic (SQG) model helps describe the surface dynamics (Blumen, 1978; Held et al., 1995; Lapeyre and Klein, 2006; Constantin et al., 1994, 1999, 2012). Despite its simplicity, the SQG relation provides a good diagnosis to relate mesoscale surface buoyancy fields to surface and interior velocity fields. Nevertheless, QG and SQG paradigms assume strong rotation and strong stratification ($Fr \sim Ro \ll 1$) and thus neglect the submesoscale ageostrophic dynamics. In particular, the QG velocity is horizontal and solenoidal. This structure prevents the emergence and development of realistic submesoscale features such as frontogenesis, restratification, and asymmetry between cyclones and anticyclones (Lapeyre et al., 2006; Klein et al., 2008). In contrast, the QG^{+1} (Muraki et al., 1999) and SQG^{+1} (Hakim et al., 2002) models capture such phenomenon with a

(one degree) higher order power series expansions in the Rossby number. This comes with an additional complexity. In particular, the SQG^{+1} model involves a nonlinear PV. Semi-Geostrophic (SG) (Eliassen, 1949; Hoskins, 1975) and Surface Semi-Geostrophic (SSG) models (Hoskins, 1976; Hoskins and West, 1979; Badin, 2013; Ragone and Badin, 2016) also offer simple alternatives to the QG framework. Within a weaker stratification context ($Fr^2 \sim Ro \ll 1$), ageostrophic terms emerges to better represent fronts and filaments than QG dynamics. The SSG model is formally similar to SQG as it is in the same way associated with a zero PV. Yet, SSG involves a space remapping (from geostrophic coordinates to physical coordinates) together with a nonlinear term in the PV that is often neglected (Ragone and Badin, 2016). These terms – both of order 1 in Rossby – bring relevant horizontal velocity divergence as in SQG^{+1} model. Nevertheless, these terms require a more involved numerical inversion.

In this paper, we derive a linear SQG model enabling to cope with frontal dynamics without explicitly resolving higher Rossby order. PV is not arbitrarily set to zero, it rigorously results from a strong submesoscale activity. Indeed, this underlying turbulence makes the turbulent diffusion comparable to the Coriolis force, and consequently cancels the PV. Such a derivation is a direct consequence of the dynamics under location uncertainty (Mémin, 2014; Resseguier et al., 2017a,b), for which the velocity is decomposed between a large-scale resolved component and a time-uncorrelated unresolved component. Derived models then rigorously handle sub-grid tensors. In particular, they link together small-scale velocity statistics, turbulent diffusion, small-scale induced velocity and backscattering effects.

After briefly recalling the main features of models under location uncertainty (section 2), a modified SQG model is derived (section 3). Finally, the ensuing diagnostic relation is tested on realistic very-high resolution model outputs (section 4).

4.2 Models under location uncertainty

Hereafter, we briefly outline the main ideas for the derivation of these stochastic models (for a more complete description, see Resseguier et al. (2017a)). This relies on a decomposition of the flow velocity in terms of a large-scale component, $\boldsymbol{w}$, and a random field uncorrelated in time, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$:

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} = \boldsymbol{w} + \boldsymbol{\sigma}\dot{\boldsymbol{B}}. \tag{4.2.1}$$

The latter represents the small-scale velocity component. This solenoidal, possibly anisotropic and non-homogeneous random field corresponds to the aliasing effect of the unresolved velocity component. To parametrize its spatial correlations, an infinite-dimensional linear operator, $\boldsymbol{\sigma}$, is applied to a space-time white noise, $\dot{\boldsymbol{B}}$. The decomposition (4.2.1) leads to a stochastic representation of the Reynolds transport theorem (RTT) and of the material derivative, D_t (derivative along the flow (4.2.1)). In most cases, this derivative coincides with the stochastic transport operator, $\mathbb{D}_t$, defined for every field, Θ, as follow:

$$\mathbb{D}_t\Theta \triangleq \underbrace{\mathrm{d}_t\Theta}_{\substack{\triangleq\ \Theta(\boldsymbol{x},t+\mathrm{d}t)-\Theta(\boldsymbol{x},t)\\ \text{Time increment}}} + \underbrace{(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\Theta}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}}\mathrm{d}t, \tag{4.2.2}$$

where the time increment term $\mathrm{d}_t\Theta$ stands instead of the partial time derivative as Θ is non differentiable. The diffusion coefficient matrix, $\boldsymbol{a}$, is solely defined by the one-point one-time covariance of the unresolved displacement per unit of time:

$$\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^T = \frac{\mathbb{E}\left\{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}}{\mathrm{d}t}, \tag{4.2.3}$$

and the modified drift is given by

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T. \tag{4.2.4}$$

For a divergent small-scale velocity, this drift would involve an additional component (Resseguier et al., 2017a). The stochastic RTT and material derivative involve a diffusive subgrid term, a

multiplicative noise and a modified advection drift induced by the small-scale inhomogeneity. This material derivative has a remarkable conservative property. Indeed, for any field, Θ, randomly transported, *i.e.*

$$\Theta(\boldsymbol{X}(t+\Delta t), t+\Delta t) \quad = \quad \Theta(\boldsymbol{X}(t), t), \tag{4.2.5}$$

Resseguier et al. (2017a) showed that the energy of each realization is conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega}\Theta^2 = 0. \tag{4.2.6}$$

The RTT enables us to express the conservation law of mechanics (linear momentum, energy, mass) with a partially known velocity. Deterministic and random subgrid parametrizations for various geophysical flow dynamics can then directly be obtained. Stochastic Navier-Stokes and Boussinesq models can be derived as discussed by Mémin (2014) and Resseguier et al. (2017a). The latter model involves random transports of buoyancy and velocity, together with incompressibility constraints.

4.3 Mesoscale flows under strong uncertainty

From the Boussinesq model, the QG assumptions state a strong rotation and a strong stratification. This is of particular interest to study flows at mesoscale, where both kinetic and buoyant dynamics are important. More specifically, we focus on horizontal length scales, L, such as:

$$\frac{1}{Bu} = \left(\frac{Fr}{Ro}\right)^2 = \left(\frac{L}{L_d}\right)^2 \sim 1 \text{ and } \frac{1}{Ro} = \frac{Lf_0}{U} \gg 1, \tag{4.3.1}$$

where U is the horizontal velocity scale, $L_d \triangleq \frac{Nh}{f}$ is the Rossby deformation radius, N is the stratification (Brunt-Väisälä frequency) and h is the characteristic vertical length scale. In the following, both differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ, represent 2D operators. Moreover, $\boldsymbol{\sigma}_{H\bullet}$ stands for the horizontal component of $\boldsymbol{\sigma}$, $\boldsymbol{a}_H$ for $\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T$ and A_u for its scaling.

4.3.1 Specific scaling assumptions

Similarly to Resseguier et al. (2017b), scalings within the QG framework (4.3.1) can authorize the set up of a non-dimensional stochastic Boussinesq model amenable to further simplifications.

Quadratic variation scaling

Models under location uncertainty involve subgrid terms which have also to be scaled. A new dimentionless number, Υ, quantifying the ratio of horizontal advection and horizontal turbulent diffusion is therefore introduced:

$$\Upsilon \triangleq \frac{U/L}{A_u/L^2} = \frac{U^2}{A_u/T}. \tag{4.3.2}$$

We can also relate it to the ratio of Mean Kinetic Energy (MKE), U^2, to the Turbulent Kinetic Energy (TKE), A_u/T_σ, where T_σ is the small-scale correlation time. This reads:

$$\Upsilon = \frac{1}{\epsilon}\frac{MKE}{TKE}, \tag{4.3.3}$$

where $\epsilon = T_\sigma/T$ is the ratio of the small-scale to the large-scale correlation times. This parameter, ϵ, is central in homogenization and averaging methods (Majda et al., 1999; Givon et al., 2004; Gottwald and Melbourne, 2013). The number Υ/Ro measures the ratio between rotation and horizontal diffusion. For a parameter Υ close or larger than unity, the geostrophic balance still holds (Resseguier et al., 2017b), whereas for $\Upsilon \sim Ro$, this balance is modified. Throughout this paper, we focus on this specific scaling.

The parameter Υ depends through A_u on the flow and on the resolution scale. In order to specify the scaling and the resulting associated model, knowledge of the characteristic horizontal eddy diffusivity or eddy viscosity is needed. Tuning experiences of usual subgrid parametrizations may provide such information, and Boccaletti et al. (2007) give some examples of canonical values.

If absence of characteristic values, absolute diffusivity or similar mixing diagnoses could be measured (Keating et al., 2011) as a proxy of the variance tensor. Small values of Υ are generally relevant for the ocean where the TKE is often one order of magnitude larger than the MKE (Wyrtki et al., 1976; Richardson, 1983; Stammer, 1997; Vallis, 2006). Note that here the TKE may encompass all the unresolved dynamics down to the Kolmogorov scale.

Vertical unresolved velocity

To scale the vertical unresolved velocity, we consider

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{Ro}{Bu} D, \tag{4.3.4}$$

where $D = \frac{h}{L}$ is the aspect ratio and the subscript H indicates horizontal coordinates. The ω-equation (Giordani et al., 2006) justifies such a scaling. For any velocity $\mathbf{u} = (\mathbf{u}_H, \mathfrak{w})^T$, which scales as $(\mathfrak{U}, \mathfrak{U}, \mathfrak{W})^T$, this equation reads

$$f_0^2 \partial_z^2 \mathfrak{w} + N^2 \Delta \mathfrak{w} = \boldsymbol{\nabla} \cdot \boldsymbol{Q} \approx -\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \mathbf{u}_H^T \boldsymbol{\nabla} \mathfrak{b}) \approx -f_0 \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \mathbf{u}_H^T \partial_z \mathbf{u}_H^\perp\right), \tag{4.3.5}$$

where $\mathfrak{b}$ stands for the buoyancy variable and $\boldsymbol{Q}$ for the so-called $\boldsymbol{Q}$-vector. In its non-dimentional version, the ω-equation reads:

$$\frac{\mathfrak{W}}{\mathfrak{U}} \left(\partial_z^2 \mathfrak{w} + Bu \Delta \mathfrak{w}\right) \approx DRo \boldsymbol{\nabla} \cdot \boldsymbol{Q}. \tag{4.3.6}$$

The Burger number is small at planetary scales where the rotation dominates ($\frac{\mathfrak{W}}{\mathfrak{U}} \sim DRo$) and is large at submesoscales where the stratification dominates ($\frac{\mathfrak{W}}{\mathfrak{U}} \sim DRo/Bu$). For the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, the latter is thus more relevant.

Relations between the isopicnical tilt and mixing give another justification of the scaling (4.3.4). Based on baroclinic instabilities theory, anisotropy specifications of eddy diffusivity sometimes rely on this tilt (Vallis, 2006). Moreover, several other authors suggest that the eddy activity and the associated mixing mainly occur along isentropic surfaces (Gent and Mcwilliams, 1990; Pierrehumbert and Yang, 1993).

For QG dynamics, the Burger number is of order one and the scaling in DRo and in DRo/Bu coincides. In particular, they encode a mainly horizontal unresolved velocity:

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{Ro}{Bu} D \ll D. \tag{4.3.7}$$

This is consistent with the assumption of a large stratification, *i.e.* flat isopycnicals, if we admit that the activity of eddies preferentially appears along the isentropic surfaces. As a consequence, the terms $(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z \partial_z$ scale as $\frac{Ro}{Bu}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \cdot \boldsymbol{\nabla}$. In the QG approximation, the scaling of the diffusion and effective advection terms including $\boldsymbol{\sigma}_{z\bullet}$ are one to two orders smaller (in power of Ro/Bu) than terms involving $\boldsymbol{\sigma}_{H\bullet}$. For any function ξ, the vertical diffusion $\partial_z(\frac{\boldsymbol{\sigma}_{z\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z \xi)$ is one order smaller than the horizontal-vertical diffusion term $\boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{z\bullet}^T}{2}\partial_z \xi\right)$ and two orders smaller than the horizontal diffusion term $\boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T}{2}\boldsymbol{\nabla}\xi\right)$.

Beta effect

The beta effect is weak at mid-latitude mesoscales. Yet, at the first order, it influences the absolute vorticity. So, we choose the same scaling as Vallis (2006):

$$\beta y \sim \boldsymbol{\nabla}^\perp \cdot \boldsymbol{u} \sim \frac{U}{L} = Ro f_0. \tag{4.3.8}$$

4.3.2 Stratified Quasi-Geostrophic model under strong uncertainty

Strong uncertainty condition corresponds to Υ having an order of magnitude close to the Rossby number. More specifically, we assume $Ro \leqslant \Upsilon \ll 1$. In this situation, the random eddies have larger energy than the large-scale mean kinetic energy. Accordingly, the diffusion and drift terms are one order of magnitude larger than the advection terms.

In the case of strong ratio Υ, the diffusion is very large and the system is not approximately in geostrophic balance anymore. The large-scale horizontal velocity becomes divergent, and decoupling the system is more tedious. For sake of simplicity, in the following we consider the case of homogeneous and horizontally isotropic turbulence. As a consequence, the variance tensor, $\boldsymbol{a}$, is constant in space and diagonal:

$$\boldsymbol{a} = \begin{pmatrix} a_H & 0 & 0 \\ 0 & a_H & 0 \\ 0 & 0 & a_z \end{pmatrix}. \tag{4.3.9}$$

Modified geostrophic balance under strong uncertainty

For horizontal homogeneous turbulence, the large-scale geostrophic balance is modified by the horizontal diffusion, whereas the unresolved velocity is in geostrophic balance:

$$\begin{cases} \boldsymbol{f} \times \boldsymbol{u} - \dfrac{a_H}{2}\Delta \boldsymbol{u} = -\dfrac{1}{\rho_b}\boldsymbol{\nabla} p', & (4.3.10a) \\ \boldsymbol{f} \times \boldsymbol{\sigma}_H \mathrm{d}\boldsymbol{B}_t = -\dfrac{1}{\rho_b}\boldsymbol{\nabla} \mathrm{d}_t p_\sigma, & (4.3.10b) \end{cases}$$

where $\boldsymbol{u}$ is the large-scale horizontal velocity, p' the time-correlated component of the pressure, $\dot{p}_\sigma = \frac{\mathrm{d}p_\sigma}{\mathrm{d}t}$ the time-uncorrelated component, and ρ_b is the mean density. For a constant Coriolis frequency, the first equation can be solved in Fourier space. The Helmholtz decomposition of the velocity reads:

$$\begin{cases} \boldsymbol{u} = \boldsymbol{\nabla}^\perp \psi + \boldsymbol{\nabla}\tilde{\psi}, & (4.3.11a) \\ \hat{\psi} = \left(1 + \left\| \dfrac{\boldsymbol{k}}{k_c} \right\|_2^4 \right)^{-1} \dfrac{\hat{p}'}{\rho_b f}, & (4.3.11b) \\ \tilde{\psi} = \dfrac{1}{k_c^2}\Delta\psi, & (4.3.11c) \end{cases}$$

where $k_c = \sqrt{\frac{2f_0}{a_H}}$ and the hat accent indicates a horizontal Fourier transform. This solution is derived in Appendix 4.A using geometric power series of matrices. The obtained formula is valid for any right-hand side in equation (4.3.10a). For instance, additional forcing such as an Ekman stress could be taken into account. In equation (4.3.11), the solenoidal component of the velocity, $\boldsymbol{\nabla}^\perp\psi$, corresponds to the usual geostrophic velocity multiplied by a low-pass filter (4.3.11b). The irrotational (ageostrophic) component of the velocity, $\boldsymbol{\nabla}\tilde{\psi}$, dilates the anticyclones (maximum of pressure and negative vorticity) and shrinks the cyclones (minimum of pressure and positive vorticity) at small scales. Indeed, according to equation (4.3.11c), the divergence of the velocity corresponds to the vorticity Laplacian divided by k_c^2. Naturally, this structure is reminiscent of the Ekman model where divergence and vorticity would be related by a double vertical derivative:

$$\delta = \ell_{E_k}^2 \partial_z^2 \zeta \quad \text{where} \quad \begin{cases} \delta &= \boldsymbol{\nabla}\cdot\boldsymbol{u}, \\ \zeta &= \boldsymbol{\nabla}^\perp \cdot \boldsymbol{u}, \end{cases} \tag{4.3.12}$$

and ℓ_{E_k} is the thickness of the Ekman layer. The turbulent diffusion involved in equation (4.3.11c) is rather horizontal due to the strong stratification assumption (see (4.3.4)). In the proposed stochastic model, the divergent component and the low-pass filter of the system (4.3.11) are parameterized by the spatial cutoff frequency k_c, which moves toward larger scales when the diffusion coefficient a_H increases. If both the vorticity and the divergence can be measured at large scales, the previous relation should enable to estimate the cutoff frequency k_c by fitting terms of equation (4.3.11c). Then, the horizontal diffusive coefficient, a_H, or the variance of the horizontal small-scale velocity (at the time scale Δt), $a_H/\Delta t$, can be deduced.

Modified SQG relation under strong uncertainty

To derive a QG model, we use the other equations of the stochastic Boussinesq model at the 0-order. After some algebra (see Appendix 4.C), we obtain directly a zero PV in the fluid interior:

$$PV = \left(\Delta + \left(1 + \frac{\Delta^2}{k_c^4}\right)\partial_z\left(\left(\frac{f_0}{N}\right)^2\partial_z\right)\right)\psi = 0, \tag{4.3.13}$$

where $k_c = \sqrt{\frac{2f_0}{a_H}}$. The assumptions used here correspond to the same used for a classical QG model (Vallis, 2006), except that the dissipation, due to the noise, is strong. It is a striking result. Instead of finding a model in the form of a classical QG model, developments, through a strong uncertainty, directly leads to the description of surface dynamics, a SQG model. It means that the subgrid dissipation prevents the development of the interior dynamics. Without this dynamics, no baroclinic instabilities can grow (Lapeyre and Klein, 2006). If the stratification is vertically invariant, this static linear equation can be solved by imposing a vanishing condition in the deep ocean ($z \to -\infty$) and a specified boundary value at a given depth ($z = \eta$). The horizontal Fourier transform of the solution then reads:

$$\hat{\psi}(\boldsymbol{k}, z) = \hat{\psi}(\boldsymbol{k}, \eta)\ \exp\left(\frac{N\|\boldsymbol{k}\|_2}{f_0\sqrt{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4}}\ (z - \eta)\right). \tag{4.3.14}$$

At $z = \eta$, the modified SQG relation is:

$$\hat{b}(\boldsymbol{k}, \eta) = N\|\boldsymbol{k}\|_2\sqrt{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4}\ \hat{\psi}(\boldsymbol{k}, \eta), \tag{4.3.15}$$

where b stands for the buoyancy. In the following, we will refer to (4.3.15) as the SQG relation under Strong Uncertainty (SQG_{SU}). For low wave number or moderate uncertainty ($\|\boldsymbol{k}/k_c\|_2^2 \sim Ro/\Upsilon \ll 1$), we retrieve the standard SQG relation. The expression of the stream function as a function of the buoyancy is expressed as the convolution with a Green function, $G_{SQG} = \frac{1}{2\pi N}\|\boldsymbol{x}\|^{-1}$, and the velocity decays rapidly as the inverse of the square distance to the point vortex center. On the other hand, for very high wave numbers or very large uncertainty ($\|\boldsymbol{k}/k_c\|_2^2 \sim Ro/\Upsilon \gg 1$), the velocity tends very quickly to zero. For strong uncertainty or small scales ($\|\boldsymbol{k}/k_c\|_2^2 \sim Ro/\Upsilon \sim 1$),

$$\hat{b} = \frac{\sqrt{2}N}{k_c}\left(\|\boldsymbol{k}\|_2^2 + \underset{\|\boldsymbol{k}\|\to k_c}{O}\left(\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2 - 1\right)^2\right)\hat{\psi}. \tag{4.3.16}$$

Accordingly, we may see the SQG_{SU} relation as an intermediary between two relevant models in geophysics: the SQG dynamics where the tracer (the buoyancy) is proportional to $\|\boldsymbol{k}\|_2\hat{\psi}$ and a two-dimensional flow dynamics where the tracer (the vorticity) is proportional to $\|\boldsymbol{k}\|_2^2\hat{\psi}$. In the latter case, the streamfunction can be expressed as the convolution of the buoyancy with the Green function, $G_{2D} = \frac{k_c}{2\sqrt{2}\pi N}\ln\|\boldsymbol{x}\|$, and the velocity decays slowly as the inverse of the distance to the point vortex center. Nevertheless, contrary to the two-dimensional flow and the SQG models, the 2D velocity $\boldsymbol{u}$ is divergent (see equation (4.3.11c)). The total horizontal velocity can be computed from the buoyancy, through the Helmholtz decomposition (4.3.11a), the modified SQG relation (4.3.15) and the equation (4.3.11c). As derived, the vertical velocity is finite and given by the main balance of the buoyancy equation:

$$w = \frac{f_0}{N^2}\frac{1}{k_c^2}\Delta b. \tag{4.3.17}$$

Note that this equation is not derived from a non-hydrostatic vertical momentum equation. Equation (4.3.17) is directly obtained from the thermodynamic equation. It expresses the fact that, under strong stratification and strong horizontal diffusion, the buoyancy anomalies are mainly created by vertical advection. This relation is similar to the result of Garrett and Loder (1981), except the proportionality coefficient. Indeed, Garrett and Loder (1981) consider vertical diffusion

and neglect the horizontal one. Invoking the thermal wind relation and the stratification structure, vertical variations are then associated with horizontal buoyancy variations. In the present development, the vertical velocity scales as $\frac{R_o}{\Upsilon B_u} DU \sim \|\boldsymbol{k}/k_c\|_2^2 \frac{DU}{B_u}$. It is prominent at small scales and proportional to the variance tensor, such as the divergent component of the horizontal velocity.

Figures 4.1 and 4.2 show the static link between the 3D velocity and buoyancy for two isolated vortices and a front, respectively. As obtained, the solenoidal component is similar to the classic SQG velocity. In Figure 4.1, the non-rotational component forces the anticyclone (warm spot) to spread, and the cyclone (cold spot) to shrink. Note that our study focuses on the ocean dynamics. For atmospheric applications, the vertical axis should be inverted and the sign of the temperature anomaly changed (Ragone and Badin, 2016). In Figure 4.2, the irrotational component is weak on the warm side of the front, but strongly strengthens the cold side. As modeled in the SQG^{+1} (Hakim et al., 2002) and Surface Semi-Geostrophic (SSG) (Badin, 2013; Ragone and Badin, 2016) models, a frontolysis (resp. frontogenesis) develops on the warm (resp. cold) side of the front. In Figure 4.1, a downwelling of warm water and a upwelling of cold water appear. As the vertical velocity comes from the thermodynamic equation and not from the vertical momentum equation, it is the cause of the buoyancy anomaly not its consequence. Whereas the irrotational horizontal component is stronger close to a front than within an eddy, the vertical velocity associated with a front is found much weaker than the one associated with an isolated eddy.

4.4 Diagnostic under strong uncertainty

As derived, under strong uncertainty, the eddy diffusion is substantial and modifies the geostrophic balance (4.3.11). The velocity becomes divergent and equation (4.3.11c) offers a diagnostic of this divergence. This diagnostic states that the divergence should be proportional to the Laplacian of the vorticity:

$$\delta = \frac{1}{k_c^2} \Delta \zeta. \tag{4.4.1}$$

To evaluate the relevance of this diagnostic, outputs of a realistic 3D high-resolution oceanic simulation are used. During winter, the eddy activities are usually stronger, especially close to energetic currents. For this reason, the Gulf-Stream during winter season is a test-bed region for high-resolution simulation (Gula et al., 2015).

Figure 4.3 shows the temperature of the first and of the 58th day. Simulations are three-dimensional and involve a fine spatial and temporal resolutions. Equation (4.3.11c) is a surface mesoscale diagnostic valid far from the coasts. Consequently, the surface fields are filtered temporally and spatially. The final time step is one day and the final resulting spatial resolution is 3 km. Figure 4.3 displays the original surface field and the filtered cropped fields.

Figure 4.4 compares the reference divergence field to our estimate, the Laplacian of the vorticity. An overall agreement clearly emerges. Nonetheless, the small scales of our estimate are more energetic than the small scales of the real divergence field. For this reason, the spatial fields are further filtered at a resolution of 30 km. Except for some small spots, estimation and reference are similar. In particular, fronts – associated with two length scales: one at sub-mesoscales and one at mesoscales – are highlighted.

Figure 4.5 specifies the relevance and the limitations of the proposed diagnostic. The spectra of the two fields unveil a very good match at mesoscale range ($L > 60$km *i.e.* $\kappa < 10^{-4}$), whereas they differ at sub-mesoscales. This difference is certainly not surprising, the estimation being derived for large scale components. Note, the velocity divergent component is far from being zero in the mesoscale range. Compared to the solenoidal component, its spectrum is certainly much flatter and smaller in this range. Nevertheless, the mesoscale divergence is stronger than the sub-mesoscales divergence. The ratio of Fourier transform modulus further confirms the accuracy of our diagnostic at mesoscales and makes clear the difference at sub-mesoscales. The -1 slope may suggest that a fractional diffusion would be preferable to a Laplacian diffusion at those scales.

The complementary analysis is the coherence, which is a measure of the phase relationship

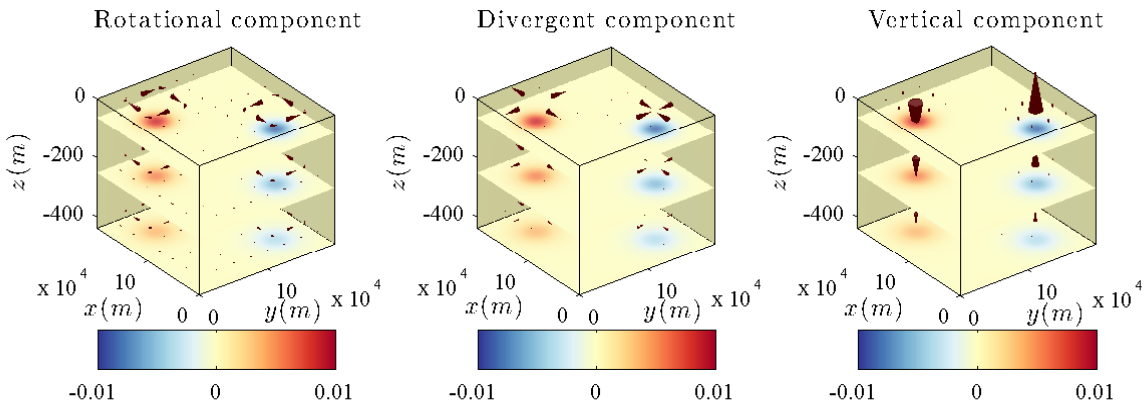


Figure 4.1: Value of the interior buoyancy created by a warm spot and a cold spot at the surface. The two components of the velocity are also shown. The upwelling and shrinking of the cyclone (cold spot) and the downwelling and spreading of the anticyclone (warm spot) are clearly visible.

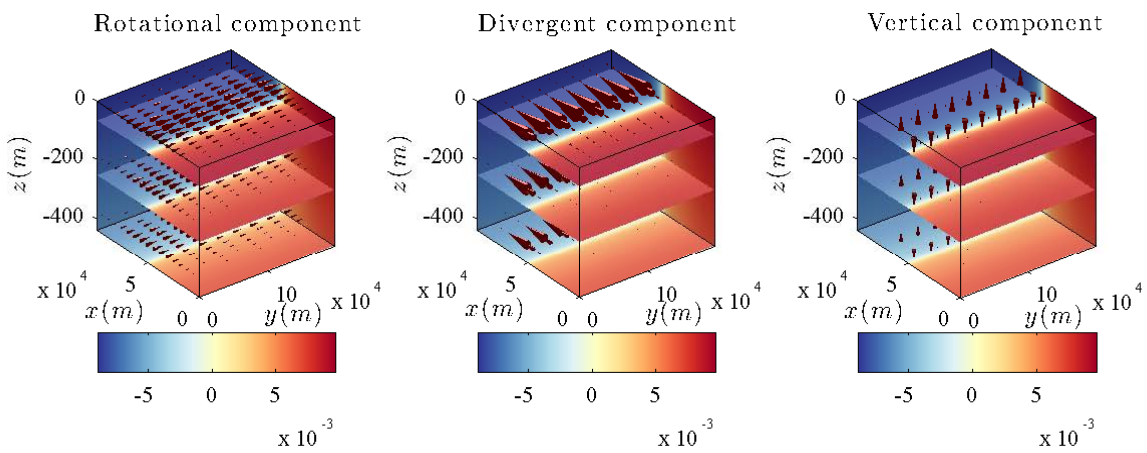


Figure 4.2: Value of the interior buoyancy created by a front at the surface. The two components of the velocity are also shown. The divergence effects will strengthen the front on the cold side (frontogenesis) and smooth the front on the warm side (frontolysis). The vertical velocity is here much weaker than in the case of isolated spots.

between two fields. Specifically, the coherence is the Fourier modes correlation coefficient:

$$\Re\left(\frac{\widehat{\delta}(\boldsymbol{k})\ \overline{\widehat{\Delta\zeta}(\boldsymbol{k})}}{\left|\widehat{\delta}(\boldsymbol{k})\ \overline{\widehat{\Delta\zeta}(\boldsymbol{k})}\right|}\right), \tag{4.4.2}$$

where $\Re$ denotes the real part. The coherence is the cosinus of the phase shift, θ, between the two fields. Here, we directly show the phase shift averaged on angular spatial frequencies.

For our estimate the phase-shift is about $0.8 \approx \frac{\pi}{4}$. It means that a linear transformation of the large-scale vorticity can explain more than half of the divergence. As a comparison, the same analysis was done with the SQG relation, using temperature anomaly instead of buoyancy (not shown). The phase shift was similar.

From Figure 4.4, one further get a rough estimation for the multiplicative constant of the proposed diagnostic: $k_c^2 \approx 10^{-7}$. It suggests a spatial cutoff $k_c^{-1} \approx 3$ km and a diffusion coefficient $a_H/2 \approx 1000\ m^2.s^{-1}$. This value is canonical, according to Boccaletti et al. (2007), which upholds the proposed approach. To confirm the validity of our strong uncertain assumption, it can be evaluated:

$$\frac{Ro}{\Upsilon} \sim \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \sim \frac{a_H}{2f_0}\kappa^2 \sim 0.1, \tag{4.4.3}$$

with $\frac{2\pi}{\kappa} = 60$ km.

The unresolved energy can also be estimated. From a mesoscale point of view, motions induced with diurnal cycles can be approximated as delta-correlated processes. Hence, an estimation of the unresolved horizontal velocity amplitude shall follow from $\sqrt{a_H/\Delta t} \approx 10^{-1} m.s^{-1}$, with $\Delta t = 1$ day. Considering the present simulation, this is consistent with the sub-mesoscale velocity field.

4.5 Conclusion

To develop models under location uncertainty, the highly-oscillating unresolved velocity component is assumed to be uncorrelated in time. Consequently, the expression of the material derivative and hence most fluid dynamics models are modified, taking into account an inhomogeneous and anisotropic diffusion, an advection correction and a multiplicative noise. In this work, we simplify a Boussinesq model under location uncertainty assuming strong rotation, stratification, and subgrid turbulence. From this last assumption, the geostrophic balance is modified, and an horizontal divergent velocity explicitly appears. Furthermore, the QG approximation implies a zero PV. In other words, the strong uncertainty prevents interior dynamics at mesoscales. This provides a new derivation of the SQG model from the Boussinesq equations. The ensuing SQG model with divergent velocity is denoted SQG_{SU}. It exhibits physically relevant asymmetry between cold and warm areas, and suggests a diagnostic of the mesoscale divergence from the vorticity, as successfully tested on very high-resolution simulated data.

A more complete model could encompass white noise components for temperature, salinity and density. At mesoscales, a thermal wind relation should relate these time uncorrelated components to the unresolved velocity. Therefore, these additional terms should provide the vertical structure of the unresolved velocity, without increasing the complexity of the parametrization.

Finally, besides solar forcing, the restratification is certainly a complicated process related to frontal dynamics. In the Mixed Layer (ML), the ML instabilities are often triggered by non-hydrostatic motions. They generate very-small-scale baroclinic instabilities and slumpings of the fronts (Boccaletti et al., 2007). For such phenomena, subgrid parameterizations are necessary. They must act to horizontally homogenize and restratify the ML. In such a context, the SQG_{SU} model may constitute a simple solution or, at least a first step to develop models under location uncertainty in this direction. To encode the weak stratification of the ML, stochastic Semi-Geostrophic (SG) and Surface Semi-Geostrophic (SSG) models could also be derived. According to our scaling of the vertical unresolved velocity (4.3.4), a weaker stratification should then enhance the vertical mixing compared to the SQG_{SU} model. The modified geostrophic balance (4.3.11) would involve both

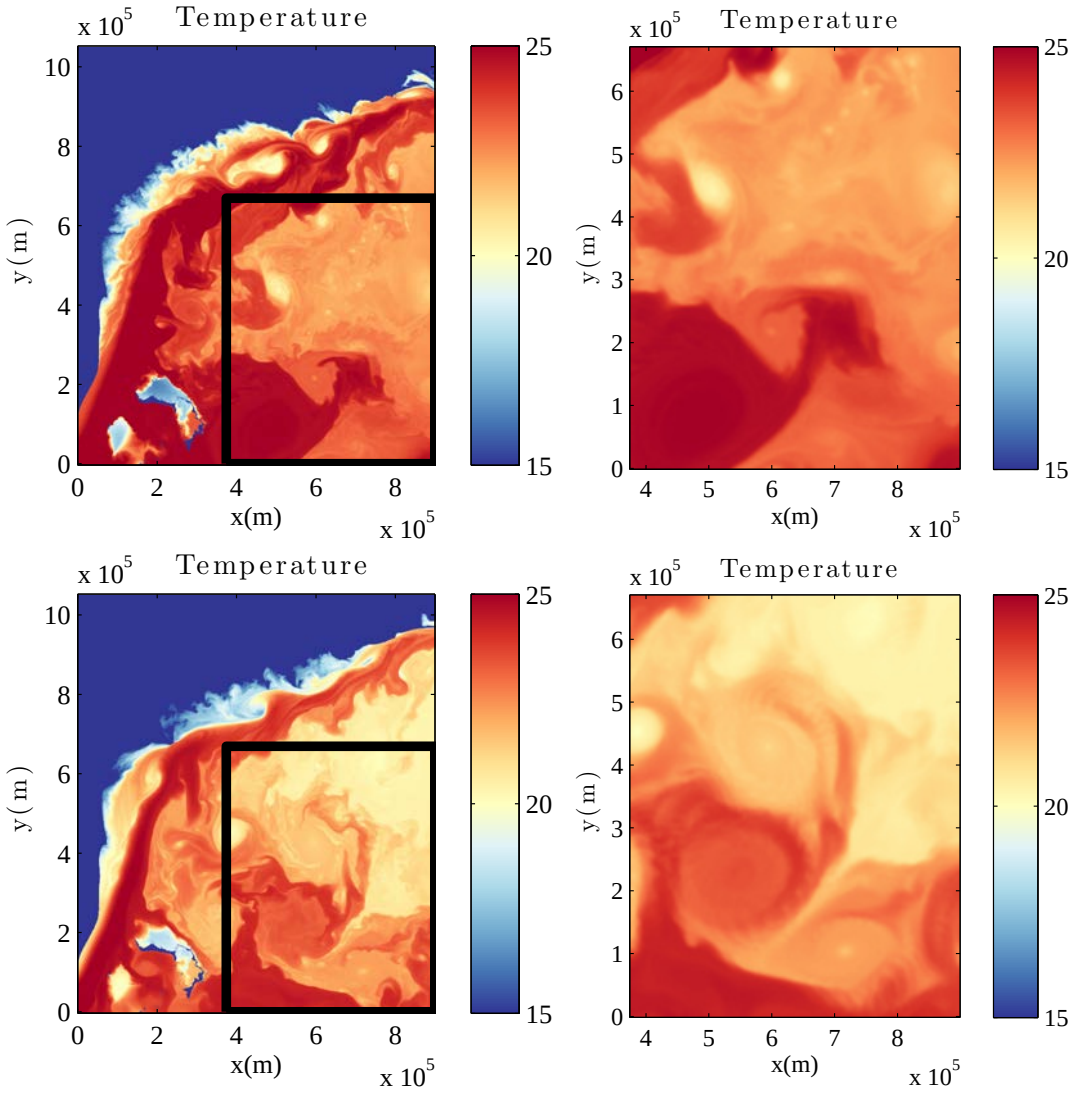


Figure 4.3: Temperature (in Celsius degree) for the first (top) and 58th day (bottom) at high temporal and spatial resolution ($\Delta t = 12$h and $\Delta x = 750$m) (left) and after filtering ($\Delta t = 1$ day and $\Delta x = 3$km) (right). The black line on the top pictures highlight the region selected for the diagnostic.

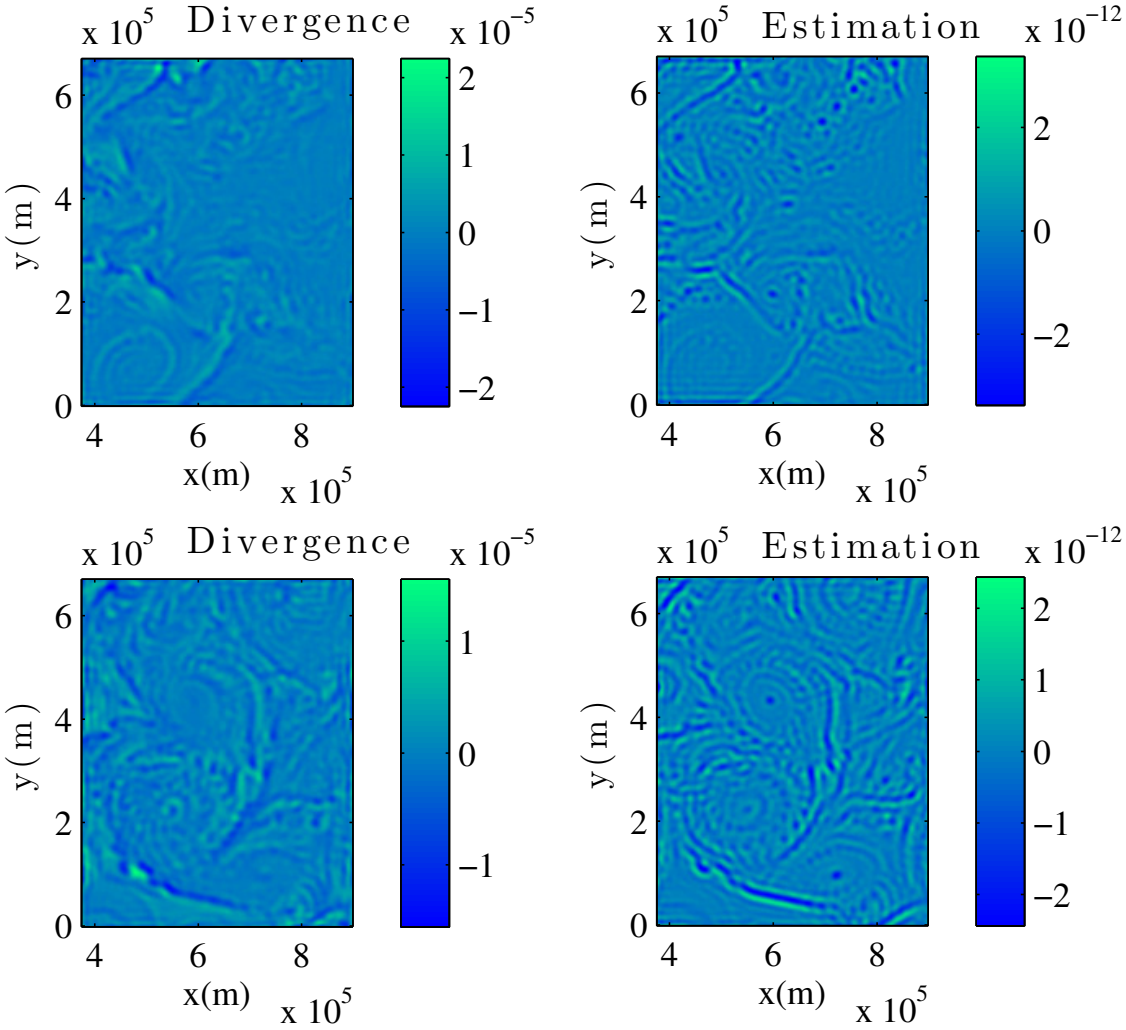


Figure 4.4: Divergence (s^{-1}) and Laplacian of the vorticity ($m^{-2}.s^{-1}$) for the first and the 58th day at a 30-km resolution. According to our modified geostrophic balance under strong uncertainty, the latter is an estimation of the mesoscale divergence up to a multiplicative constant.

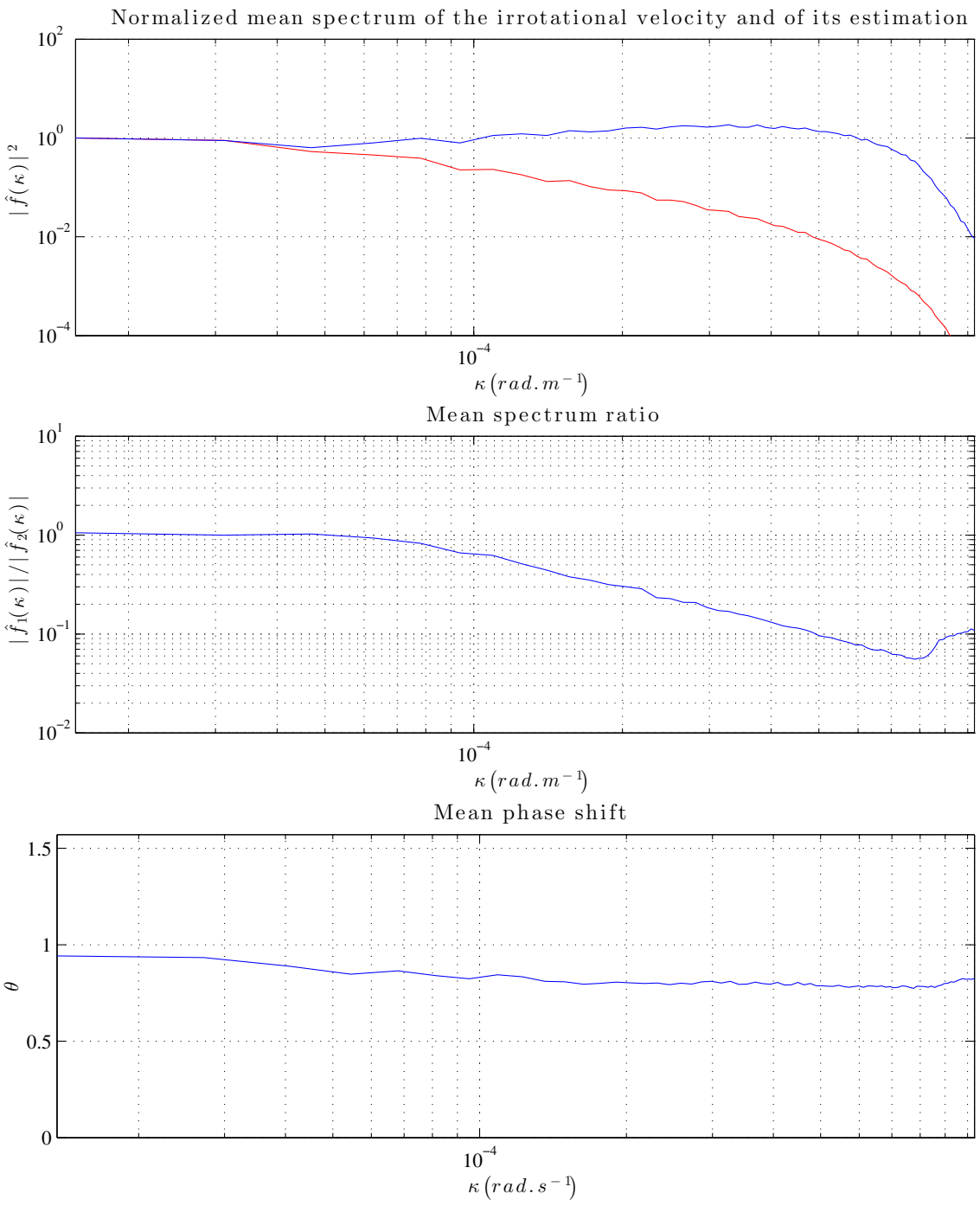


Figure 4.5: From top to bottom: Normalized spectrum of the irrotational velocity component (red) and of our estimate of this component (blue), ratio of the Fourier transform modulus of the divergence to the one of our estimate and phase shift (rad) between the divergence and its estimate. Each of these spectral quantities is averaged on angular spatial frequencies and on the 58 winter days.

horizontal and vertical diffusions. Moreover, since the stratification is weaker in the SG scalings, each term of the buoyancy equation (3D transport, 3D turbulent dissipation and stratification) would have the same scaling.

Acknowledgments

The authors thank Aurélien Ponte, Jeroen Molemaker and Jonathan Gula for helpful discussions. We also acknowledge the support of the ESA DUE GlobCurrent project, the "Laboratoires d'Excellence" CominLabs, Lebesgue and Mer through the SEACS project.

Appendix

4.A Modified geostrophic balance

Under strong horizontal homogeneous turbulence, the large-scale geostrophic balance is modified by the horizontal diffusion:

$$\boldsymbol{f} \times \boldsymbol{u} - \frac{a_H}{2}\Delta_H \boldsymbol{u} = \boldsymbol{\xi}, \tag{4.A.1}$$

where $\boldsymbol{u}$ is the resolved horizontal velocity and $\Delta_H \overset{\triangle}{=} \partial_x^2 + \partial_y^2$ the horizontal Laplacian. On the right-hand side, $\boldsymbol{\xi}$ is the pressure gradient. Let us note that $\boldsymbol{f} \times \boldsymbol{u} = f\boldsymbol{J}\boldsymbol{u}$ with $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and that $\boldsymbol{J}^T = \boldsymbol{J}^{-1} = -\boldsymbol{J}$. For a constant Coriolis frequency, the previous equation can be solved in the horizontal Fourier space :

$$\hat{\boldsymbol{u}} = \left(f\boldsymbol{J} + \frac{a_H}{2}\|\boldsymbol{k}\|_2^2 \mathbb{I}_d\right)^{-1} \hat{\boldsymbol{\xi}} = \left(\mathbb{I}_d - \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right)^{-1} \left(-\frac{1}{f}\hat{\boldsymbol{\xi}}^{\perp}\right), \tag{4.A.2}$$

with $k_c = \sqrt{\frac{2f}{a_H}}$. $-\frac{1}{f}\boldsymbol{\xi}^{\perp} = -\frac{1}{f}\boldsymbol{J}\boldsymbol{\xi}$ is the solution without diffusion. Expanding the right-hand side operator in Taylor series and using the properties $\boldsymbol{J}^{2p} = (-1)^p \mathbb{I}_d$ and $\boldsymbol{J}^{2p+1} = (-1)^p \boldsymbol{J}$,

$$\begin{aligned}
\left(\mathbb{I}_d - \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right)^{-1} &= \sum_{p=0}^{+\infty} \left(\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right)^p, && (4.A.3)\\
&= \sum_{p=0}^{+\infty} (-1)^p \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^{4p} \mathbb{I}_d + \sum_{p=0}^{+\infty} (-1)^p \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^{4p+2} \boldsymbol{J}, && (4.A.4)\\
&= \sum_{p=0}^{+\infty} \left(-\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4\right)^p \left(\mathbb{I}_d + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right), && (4.A.5)\\
&= \frac{1}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(\mathbb{I}_d + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right). && (4.A.6)
\end{aligned}$$

This leads to the following solution for the modified geostrophic balance:

$$\hat{\boldsymbol{u}} = \frac{1}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(-\frac{1}{f}\hat{\boldsymbol{\xi}}^{\perp}\right) + \frac{\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(\frac{1}{f}\hat{\boldsymbol{\xi}}\right). \tag{4.A.7}$$

4.B Non-dimensional Boussinesq equations

To derive a non-dimensional version of the Boussinesq equations under location uncertainty (Resseguier et al., 2017a), each term of the evolution laws is scaled (Resseguier et al., 2017b): the horizontal

coordinates $\tilde{\boldsymbol{x}}_h = L\boldsymbol{x}_h$, the vertical coordinate $\tilde{z} = hz$, the aspect ratio $D = h/L$ between the vertical and horizontal length scales. A characteristic time $\tilde{t} = Tt$ corresponds to the horizontal advection time U/L with horizontal velocity $\tilde{\boldsymbol{u}} = U\boldsymbol{u}$. A vertical velocity $\tilde{w} = (h/L)Uw$ is deduced from the divergence-free condition. We further take a scaled buoyancy $\tilde{b} = Bb$, pressure $\tilde{\phi}' = \Phi\phi'$ (with the density scaled pressures $\phi' = p'/\rho_b$ and $\mathrm{d}_t\phi_\sigma = \mathrm{d}_t p_\sigma/\rho_b$), and the earth rotation $\boldsymbol{f}^* = f\boldsymbol{k}$. For the uncertainty variables, we consider a horizontal uncertainty $\tilde{\boldsymbol{a}}_H = A_u\,\boldsymbol{a}_H$ corresponding to the horizontal 2×2 variance tensor; a vertical uncertainty vector $\tilde{a}_{zz} = A_w a_{zz}$ and a horizontal-vertical uncertainty vector $\tilde{\boldsymbol{a}}_{Hz} = \sqrt{A_u A_w}\boldsymbol{a}_{Hz}$ related to the variance between the vertical and horizontal velocity components. The resulting non-dimensional Boussinesq system under location uncertainty becomes:

Nondimensional Boussinesq equations under location uncertainty

Momentum equations

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)\boldsymbol{u} + \left(\frac{Ro}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z\boldsymbol{u}\\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}\boldsymbol{u}\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right) + \frac{1}{Ro}\left(1+Ro\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)\\
&\quad = -E_u\,\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned}\tag{4.B.1a}$$

$$\begin{aligned}
&\mathrm{d}_t w + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,w\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)w + \left(\frac{Ro}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z w\\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}w\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right) = \frac{\Gamma}{D^2}b\mathrm{d}t - \frac{E_u}{D^2}\,\partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned}\tag{4.B.1b}$$

Buoyancy equation

$$\begin{aligned}
&\mathrm{d}_t b + \left(\boldsymbol{w}^*_\Upsilon\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)\right)\cdot\boldsymbol{\nabla}b - \frac{1}{2}\frac{1}{\Upsilon}\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{a}_H\boldsymbol{\nabla}b\right)\mathrm{d}t + O\left(\frac{Ro}{\Upsilon B_u}\right)\\
&\qquad + \frac{1}{(F_r)^2}\frac{1}{\Gamma}\left(w^*_{\Upsilon/2}\mathrm{d}t + \left(\frac{Ro}{B_u}\right)\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right) = 0,
\end{aligned}\tag{4.B.1c}$$

Effective drift

$$\begin{aligned}
\boldsymbol{w}^*_\Upsilon &= \left(\boldsymbol{u}^*_\Upsilon, w^*_\Upsilon\right)^T,\\
&= \left(\left(\boldsymbol{w} - \frac{1}{2\Upsilon}\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right), \left(w - \left(\frac{Ro}{2\Upsilon B_u}\right)\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_{Hz} + O\left(\frac{Ro}{\Upsilon B_u}\right)^2\right)\right)^T,
\end{aligned}\tag{4.B.1d}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0,\tag{4.B.1e}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0,\tag{4.B.1f}$$

$$\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H\right)^T + 2\frac{Ro}{B_u}\boldsymbol{\nabla}_H\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{Ro}{B_u}\right)^2\right) = 0.\tag{4.B.1g}$$

Here, the time-correlated components and the time-uncorrelated components in the momentum equations have not been separated. The terms in $O\left(\frac{Ro}{B_u}\right)$ and $O\left(\frac{Ro}{B_u}\right)^2$ are related to the time-uncorrelated vertical velocity. These terms are too small to appear in the final QG model ($B_u = O(1)$ in QG approximation) and not explicitly shown. We only make appear the big O approximations. Traditional non-dimensional numbers are introduced : the Rossby number $R_o = U/(f_0L)$ with f_0 the average Coriolis frequency; the Froude number ($F_r = U/(Nh)$), ratio between the advective time to the buoyancy time; E_u, the Euler number, ratio between the pres-

sure force and the inertial forces, $\Gamma = Bh/U^2 = D^2BT/W$ the ratio between the mean potential energy to the mean kinetic energy. To scale the buoyancy equation, the ratio between the buoyancy advection and the stratification term has also been introduced:

$$\frac{B/T}{N^2W} = \frac{B}{N^2h} = \frac{U^2}{N^2h^2}\frac{Bh}{U^2} = Fr^2\Gamma. \tag{4.B.2}$$

Besides those traditional dimensionless numbers, this system introduces Υ, relating the large-scale kinetic energy to the energy dissipated by the unresolved component:

$$\Upsilon = \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{4.B.3}$$

4.C QG model under strong uncertainty

For the case Υ close to the Rossby number, the diffusion term is not negligible anymore and the geostrophic balance is modified. As the terms of the geostrophic balance remain large ($Ro \leqslant \Upsilon \ll 1$), the scaling of the pressure can still be done with the Coriolis force. This leads to an Euler number scaling as

$$Eu \sim \frac{1}{Ro}. \tag{4.C.1}$$

Keeping a small aspect ratio $D^2 \ll 1$, we get

$$\frac{Eu}{D^2} \sim \frac{1}{RoD^2} \gg \frac{1}{Ro} \geqslant \frac{1}{\Upsilon}. \tag{4.C.2}$$

As the Rossby number and the ratio Υ are both small in the vertical momentum equation, the inertial terms are dominated by the diffusion term which is itself negligible in front of the pressure term. The hydrostatic balance is hence conserved. The buoyancy scaling still correspond to the thermal winds relation:

$$\Gamma \sim Eu \sim \frac{1}{Ro}. \tag{4.C.3}$$

Considering the scaling $\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim D\frac{Ro}{Bu}$ for the vertical small-scale velocity, the non-dimensional evolution equations are now given by:

Momentum equations

$$Ro\left(\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{u} + O\left(\frac{Ro}{\Upsilon Bu}\right)\right) - \frac{Ro}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}(a_{ij}\boldsymbol{u})\mathrm{d}t$$
$$+ (1 + Ro\beta y)\,\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right) = -\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{4.C.4}$$

$$b\,\mathrm{d}t + O\left(\frac{RoD^2}{\Upsilon^{1/2}}\right) = \partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{4.C.5}$$

Buoyancy equation

$$\frac{Ro}{Bu}\left(\mathrm{d}_t b + \boldsymbol{\nabla}b\cdot\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)_H\right) + \partial_z b\; w\mathrm{d}t\right) - \frac{Ro}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}b\right)\mathrm{d}t$$
$$+ w\mathrm{d}t - \frac{1}{\Upsilon}\frac{Ro}{Bu}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t + \frac{Ro}{Bu}\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z + O\left(\frac{{Ro}^2}{\Upsilon {Bu}^2}\right) = 0, \tag{4.C.6}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} + \partial_z w = 0, \tag{4.C.7}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)_H + \frac{Ro}{Bu}\partial_z\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)_z = 0, \tag{4.C.8}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T + 2\frac{Ro}{Bu}\boldsymbol{\nabla}\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{Ro}{Bu}\right)^2\right) = 0. \tag{4.C.9}$$

The operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ represent 2D operators. If $R_o \sim \Upsilon$, the system is not anymore approximately in geostrophic balance. The large-scale velocity becomes divergent and decoupling the system is more involved. For sake of simplicity, we thus focus on the case of homogeneous and horizontally isotropic turbulence. As a consequence, the variance tensor a is constant in space and diagonal:

$$a = \begin{pmatrix} a_h & 0 & 0 \\ 0 & a_h & 0 \\ 0 & 0 & a_z \end{pmatrix}. \tag{4.C.10}$$

The time-correlated components of the horizontal momentum at the 0-th order can be written as:

$$-\frac{a_H}{2}\Delta \boldsymbol{u}_0 + \boldsymbol{k} \times \boldsymbol{u}_0 = -\boldsymbol{\nabla}\phi_0', \tag{4.C.11}$$

Then, equation (4.A.7) of Appendix 4.A expresses the result in Fourier space. In the physical space, the solution reads:

$$\boldsymbol{u}_0 = \boldsymbol{\nabla}^{\perp} \underbrace{\left(1 + \frac{\Delta^2}{k_c^4}\right)^{-1} \phi_0'}_{=\psi_0} + \boldsymbol{\nabla} \underbrace{\left(1 + \frac{\Delta^2}{k_c^4}\right)^{-1} \frac{\Delta}{k_c^2}\phi_0'}_{=\tilde{\psi}_0} \text{ with } k_c = \sqrt{\frac{2}{a_H}} \tag{4.C.12}$$

which is the Helmholtz decomposition of the horizontal velocity u_0 into its rotational and divergent component with a stream function ψ_0 and a velocity potential $\tilde{\psi}_0$. Differentiating the buoyancy equation at the order 0 along z, we obtain

$$\frac{a_H}{2}\Delta\partial_z\left(\frac{b_0}{B_u}\right) = \partial_z w_0 = -\boldsymbol{\nabla}\cdot\boldsymbol{u}_0 = -\Delta\tilde{\psi}_0 = -\frac{\Delta^2}{k_c^4}\psi_0. \tag{4.C.13}$$

The time-correlated part of the 0-th order hydrostatic equation relates the buoyancy to the pressure ϕ_0':

$$\frac{a_H}{2}\Delta\partial_z\left(\frac{b_0}{B_u}\right) = \frac{a_H}{2}\Delta\partial_z^2\phi_0' = \frac{a_H}{2}\Delta\partial_z^2\left(1 + \frac{\Delta^2}{k_c^4}\right)\psi_0. \tag{4.C.14}$$

Gathering these two equations leads to:

$$\left(\Delta + \left(1 + \frac{\Delta^2}{k_c^4}\right)\partial_z\left(\left(\frac{f_0}{N}\right)^2\partial_z\right)\right)\psi = 0. \tag{4.C.15}$$

Using the horizontal Fourier transform, it writes:

$$\left(-\|\boldsymbol{k}\|_2^2 + \left(1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4\right)\partial_z\left(\left(\frac{f_0}{N}\right)^2\partial_z\right)\right)\hat{\psi} = 0. \tag{4.C.16}$$

Under an uniform stratification, with a fixed value at a specific depth ($z = \eta$), and a vanishing condition in the deep ocean ($z \to -\infty$), a solution is:

$$\hat{\psi}(\boldsymbol{k}, z) = \hat{\psi}(\boldsymbol{k}, \eta)\ exp\left(\frac{N\|\boldsymbol{k}\|_2}{f_0\sqrt{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4}}\ (z - \eta)\right). \tag{4.C.17}$$

Accordingly, the buoyancy is:

$$\hat{b} = \partial_z\hat{\phi}' = f_0\left(1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4\right)\partial_z\hat{\psi} = N\|\boldsymbol{k}\|_2\sqrt{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4}\ \hat{\psi}. \tag{4.C.18}$$

Chapter 5

Bifurcations and location uncertainty in geophysical fluid flows

Abstract

A bifurcation associated with a symmetry breaking is identified in a high-resolution simulation of a simplified geophysical model, the so-called Surface Quasi-Geostrophic (SQG) model. Depending on the initial condition, two subsequent scenarios are possible. Yet, for the same initial field, low-resolution and high-resolution SQG simulations follow different scenarios. To retrieve both scenarios, two ensembles of stochastic simulations are performed on the coarser grid. The ensemble based on random initial conditions does not appear to converge and requires a large number of realizations. The other ensemble is obtained according to a randomized transport, referred to as dynamics under location uncertainty. Within this stochastic framework, many attractive theoretical properties are kept, and ensemble results robustly identify both scenarios and their respective probability density functions. MATLAB® codes are available online.

Keywords: Bifurcation, stochastic sub-grid tensor, geophysical fluid dynamics, ensemble forecasts.

For geophysical flows, chaos is ubiquitous and imposes finite limits on the predictability of fully developed turbulent flows. Long-term evolutions can then encompass several scenarios. The latter are customarily unveiled using an ensemble of deterministic simulations with random initial conditions. The large dimension of geophysical systems can prevent the efficiency of this method. Here, we illustrate on a simple example the potential of a new stochastic approach: the transport under location uncertainty. As a testbed, a bifurcation is studied for a known geophysical fluid dynamics model. As derived and numerically proved using very few simulations, the randomized dynamics largely outperforms the classical method, to robustly anticipate and describe several likely transitions and evolutions.

5.1 Introduction

For geophysical flows, a loss of predictability is generally attributed to the rapid upscale cascade of small-scale initial error (Lorenz, 1969; Leith, 1971; Leith and Kraichnan, 1972; Métais and Lesieur, 1986). This strongly hampers the applicability of deterministic large-scale forecasts.

For that reason, ensemble methods are routinely used in operational weather or climate centers for short to seasonal time scales forecasting (Berner et al., 2015, and references therein). Proposed schemes mostly consist in tracking several likely-scenarios and to increase *in fine* the models predictability skill. This necessitates a good estimation of uncertainties, both in terms of their magnitude and location. Underdispersive ensembles shall lead to overconfident dynamical models with poor forecasting skills. A good representation of the ensemble dispersion, but accompanied with a bad spatial localization of the errors, will also lead to a poor state-space representation.

The predictability skills thus critically depends on the scheme capability to model and maintain unresolved small-scale events.

In this paper, we explore the performances of a new stochastic framework with respect to this issue. Introduced by Mémin (2014), this approach is referred to as dynamics under location uncertainty. The formalism has been successfully applied to deterministic Large Eddy Simulations, more specifically to POD-Galerkin reduced order models (Resseguier et al., 2015). It can further provide randomized versions of fluid dynamics models.

Hereafter, this stochastic framework is applied to the surface quasi-geostrophy model corresponding to the evolution of a scalar quantity with 2D velocities determined by the scalar properties. As demonstrated, a bifurcation associated with a symmetry breaking is efficiently tracked using the proposed stochastic framework. Using few realizations, the probability density functions of each subsequent scenarios is well characterized, whereas methods based on random initial conditions do not converge. The identification of several scenarios, done at each time step in a reduced subspace, is obtained by Principal Component Analysis (PCA), also termed Empirical Orthogonal Functions method (EOF).

Notations and properties of the transport under uncertainty are first recalled. The test flow is then presented. Finally, a diagnosis of bifurcation is performed and discussed from both the stochastic model and a method based on randomized initial conditions.

5.2 Transport under location uncertainty

This section briefly recall notation and developments of the stochastic framework.

5.2.1 Informal description

Numerical models describing the ocean/atmosphere dynamics introduce *de facto* a coarse scale truncation of the system. Those dynamical models emanate from physical deterministic representations whose solutions are assumed to be smooth (*i.e.* differentiable) in time. Although small-scale fluid flow velocities can be characterized by local and intermittent energy bursts, possibly associated with infinitesimal characteristic time-scales, it is generally assumed that these unresolved flow components remain smooth in time. At the model resolution, the resolved (large-scale) flow can thus be considered as a coarse-grained representation of the actual Eulerian flow, with the unresolved flow component rapidly varying in time. From an observer point of view, such sub-grid dynamics can be conveniently modeled by a delta-correlated process. The smooth velocity field, denoted $\boldsymbol{w}$, represents a large-scale, possibly random, component continuous in time. The unresolved contribution, expressed as $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is then assumed Gaussian, volume preserving (divergence-free) and uncorrelated in time. This contribution can be non-homogeneous and anisotropic in space. Due to the irregularity of the resulting flow, the transport of a conserved quantity, b, by the whole velocity, defined as

$$b(\boldsymbol{X}_{t+\Delta t}, t+\Delta t) = b(\boldsymbol{X}_t, t) \tag{5.2.1}$$

reads in an informal stochastic way as

$$\partial_t b + \underbrace{\boldsymbol{w}^\star \cdot \boldsymbol{\nabla} b}_{\substack{\text{Corrected}\\ \text{advection}}} = \underbrace{\boldsymbol{\nabla} \cdot \left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla} b\right)}_{\text{Diffusion}} - \underbrace{\boldsymbol{\sigma}\dot{\boldsymbol{B}} \cdot \boldsymbol{\nabla} b}_{\substack{\text{Random}\\ \text{forcing}}}, \tag{5.2.2}$$

with a drift velocity corrected as

$$\boldsymbol{w}^\star = \boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla} \cdot \boldsymbol{a})^T. \tag{5.2.3}$$

Hence, the deterministic (for a fixed realization of velocity field $\boldsymbol{w}$) evolution equation is replaced by a stochastic equation with respect to $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. As a result, the conserved quantity is now advected by an “effective" velocity, $\boldsymbol{w}^\star$, taking into account the possible spatial variation of the small-scale velocity variance. The random forcing term in (5.2.2) relates to the advection by the unresolved

velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$. This term continuously backscatters random energy to the system. The diffusion term then accounts for the mixing effect of the small-scale random velocity. This term plays a role similar to the eddy diffusivity models introduced in classical large-scale representations (Gent and Mcwilliams, 1990; Germano et al., 1991; Lilly, 1966; Smagorinsky, 1963), in analogy with the molecular diffusion mechanism (Boussinesq, 1877).

The inhomogeneous and anisotropic diffusion coefficient matrix, $\boldsymbol{a}$, is then defined by the one-point one-time covariance of the unresolved displacement per unit of time:

$$\boldsymbol{a} = \frac{\mathbb{E}\left\{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)^T\right\}}{\mathrm{d}t}. \tag{5.2.4}$$

This ensures an exact energy balance between the amount of diffusion and the random forcing.

5.2.2 Uncertainty formalism

To derive more formally the evolution law of a scalar quantity transported by a stochastic flow, the stochastic Lagrangian description of the infinitesimal displacement associated with a particle trajectory $\boldsymbol{X}_t$ writes:

$$\mathrm{d}\boldsymbol{X}_t \;=\; \boldsymbol{w}(\boldsymbol{X}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t,t)\mathrm{d}\boldsymbol{B}_t. \tag{5.2.5}$$

In this equation, the second term explicitly figures the flow location uncertainty. Formally, this random field is defined over the fluid domain, $\Omega \subset \mathbb{R}^d$, from a d-dimensional Brownian function $\boldsymbol{B}_t$. Such a function can be interpreted as a white noise process in space and a Brownian process in time. Formally it is a cylindrical I_d-Wiener process (see Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007) for more information on infinite dimensional Wiener process and cylindrical I_d-Wiener process). The time derivative of the Brownian function, in a distribution sense, is informally denoted $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$, and is a white noise distribution. The spatial correlations of the flow uncertainty are specified through the diffusion operator $\boldsymbol{\sigma}(.,t)$, defined for any vectorial function, $\boldsymbol{f}$, through the matrix kernel $\breve{\boldsymbol{\sigma}}(.,.,t)$:

$$\boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{f} \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z},t)\boldsymbol{f}(\boldsymbol{z},t)\mathrm{d}\boldsymbol{z}. \tag{5.2.6}$$

This quantity is assumed to have a finite norm. More precisely, the operator $\boldsymbol{\sigma}$ is assumed to be Hilbert-Schmidt. We also assume that the above expression have periodic or null boundary conditions on the domain frontier. The resulting d-dimensional random field, $\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t$, is a centered vectorial Gaussian function, correlated in space and uncorrelated in time with covariance tensor:

$$\boldsymbol{Cov}(\boldsymbol{x},\boldsymbol{y},t,t') \;\triangleq\; \mathbb{E}\left\{\left(\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t\right)\left(\boldsymbol{\sigma}(\boldsymbol{y},t')\mathrm{d}\boldsymbol{B}_{t'}\right)^T\right\}, \tag{5.2.7}$$

$$= \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z},t)\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y},\boldsymbol{z},t)\mathrm{d}\boldsymbol{z}\;\delta(t-t')\mathrm{d}t. \tag{5.2.8}$$

We note that the uncertainty field has a (mean) bounded norm for any finite time. Indeed, since $\boldsymbol{\sigma}$ is Hilbert-Schmidt, the trace of operator Q – defined by the kernel $(\boldsymbol{x},\boldsymbol{y}) \mapsto \boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{\sigma}^T(\boldsymbol{y},t)$ – is finite, and for all $t \leqslant T < \infty$,

$$\mathbb{E}\left\|\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\right\|^2_{L^2(\Omega)} \;=\; \int_0^t\int_\Omega \|\breve{\boldsymbol{\sigma}}(\bullet,\boldsymbol{z})\|^2_{L^2(\Omega)}\,\mathrm{d}\boldsymbol{z}\mathrm{d}t', \tag{5.2.9}$$

$$= \int_0^t \|\boldsymbol{\sigma}\|^2_{HS,L^2(\Omega)}\mathrm{d}t' \;=\; \int_0^t tr(\boldsymbol{Q})\mathrm{d}t' \;<\; \infty, \tag{5.2.10}$$

where the index HS refers to the Hilbert-Schmidt norm. Hereafter, the diagonal of the covariance tensor, $\boldsymbol{a}$, will be referred to as the variance tensor:

$$\boldsymbol{a}(\boldsymbol{x},t)\delta(t-t')\mathrm{d}t = \boldsymbol{Cov}(\boldsymbol{x},\boldsymbol{x},t,t').$$

By definition, it is a symmetric positive definite matrix at all spacial points, $\boldsymbol{x}$. This quantity, also denoted in short as $\boldsymbol{\sigma}\boldsymbol{\sigma}^T$, corresponds to the time derivative of the so-called quadratic variation process:

$$\boldsymbol{\sigma}\boldsymbol{\sigma}^T \triangleq \boldsymbol{a} = \partial_t \left\langle \int_0^t \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_s, \left(\int_0^t \boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_r \right)^T \right\rangle.$$

with $\langle f, g \rangle$ to stand for the quadratic cross-variation process of f and g.

All along this study, we further assume that the small-scale random flow component is associated with a divergence-free diffusion tensor:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = 0. \tag{5.2.11}$$

5.2.3 Scalar advection

For a fluid flow defined by equation (5.2.5), the material derivative of a tracer writes:

$$0 = D_t b \;=\; \underbrace{\mathrm{d}_t b}_{\substack{\triangleq\, b(\boldsymbol{x},t+\mathrm{d}t)-b(\boldsymbol{x},t) \\ \text{Time increment}}} + \underbrace{(\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla} b}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla} \cdot \left(\frac{1}{2} \boldsymbol{a} \boldsymbol{\nabla} b \right)}_{\text{Diffusion}} \mathrm{d}t. \tag{5.2.12}$$

The time increment $\mathrm{d}_t b$ can be interpreted as the analog of the partial time derivative $\partial_t b$ in deterministic partial differential equations. The above expression is derived in Appendix 5.A. To ensure a stochastic isochoric flow, an incompressibility constraint on the modified drift is considered, $\boldsymbol{\nabla} \cdot \boldsymbol{w}^\star = 0$. This adds up to the divergence-free condition of the unresolved velocity (5.2.11). One can show that those two constraints enable to establish a strong energy conservation property for any realizations:

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega b^2 \;=\; 0. \tag{5.2.13}$$

The noise energy intake is exactly compensated by the diffusion term. Note that when the unresolved velocity is homogeneous, the variance tensor, $\boldsymbol{a}$, is uniform in space, as $\boldsymbol{a} = \frac{tr(\boldsymbol{a})}{d} \mathbb{I}_d$.

5.2.4 SQG model under moderate uncertainty

The buoyancy, b, is proportional to the density anomaly ρ'

$$b \triangleq -g \frac{\rho'}{\rho_0} \text{ with } \rho(x,y,z,t) = \rho_0 \left(1 - \frac{N^2}{g} z \right) + \rho'(x,y,z,t), \tag{5.2.14}$$

where ρ is the density, and N the stratification. In the ocean, the density anomaly is small compared to the total density (Boussinesq approximation) and the flow is approximately isochoric. The conservation of salinity and temperature, with a linearized equation of state, provides the transport of buoyancy (Vallis, 2006). Then, considering rapid rotation, strong stratification and uniform potential vorticity leads to the so-called SQG model (Blumen, 1978; Held et al., 1995). A random version of this model (denoted SQG_{MU}) can be derived from the location uncertainty principle. It will keep the same structure except that the buoyancy is now transported in the stochastic sense (5.2.12). The horizontal velocity $\boldsymbol{u} = \boldsymbol{w}$ is related to the buoyancy b in Fourier space through the usual SQG relation:

$$\hat{\boldsymbol{u}} = i\boldsymbol{k}^\perp \frac{\hat{b}}{N \|\boldsymbol{k}\|_2}, \tag{5.2.15}$$

where $\boldsymbol{k}$ is the horizontal wave-vector. The unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ is also horizontal and denoted $\boldsymbol{\sigma}_H \dot{\boldsymbol{B}}$. Consequently, the variance tensor is a 2×2 matrix denoted $\boldsymbol{a}_H$ and the SQG_{MU} model is two-dimensional.

5.3 Chaotic test flow and resolution issues

High-resolution deterministic SQG simulations of test flows will provide references to which we will compare random simulations performed at a lower resolution.

The initial conditions defining our test flows are shown in Figures 5.1 and 5.2, consisting of a spatially smooth buoyancy field with two warm elliptical anticyclones (positive buoyancy) and two cold elliptical cyclones (negative buoyancy) given by:

$$\begin{aligned} b_0(\boldsymbol{x}) &\triangleq F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) + F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right) \\ &\quad -F\left(\boldsymbol{x} - \begin{pmatrix} 250 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right) - F\left(\boldsymbol{x} - \begin{pmatrix} 750 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right), \end{aligned} \tag{5.3.1}$$

with

$$F(\boldsymbol{x}) \triangleq B_0 \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right) \text{ and } \left\{ \begin{array}{lcl} \sigma_x &=& 67 \text{ km}, \\ \sigma_y &=& 133 \text{ km}. \end{array} \right. \tag{5.3.2}$$

The domain size is a square box $L_x \times L_y = 1000$ km$\times$1000 km. The boundaries conditions of the simulation are doubly periodic and, for the above initial condition, there is a meridional line of symmetry at $x = 500$ km. Therefore, the zonal period of the initial condition is $L_x/2 = 500$ km. This periodicity is relatively stable and holds during the first month. Nevertheless, the SQG dynamics, is subject to an inverse cascade of energy (Capet et al., 2008), and vortices of the same sign tend to merge. When this merging occurs, this affects the global shape of the flow. In particular, the periodicity that remains in the first month eventually disappears. This symmetry is hence metastable rather than stable. The symmetry breaking corresponds to a transition from one "state" to an another. By "state", we mean a relatively "compact" and connected subspace of the state space. Warm vortices can merge at $x = 0$ or at $x = 500$km. In the following, we will refer to the first case as "scenario 1", and to the second case as "scenario 2". Because of the periodic boundary conditions, these two possible transitions are likely to occur. In a deterministic numerical simulation, the appearance of one transition or the other is determined by an infinitesimal asymmetry in the initial condition or possibly by a numerical error. This is a bifurcation. The bifurcation related to the merging of cold vortices is analog. With those two simultaneous bifurcations, there are thus $2 \times 2 = 4$ likely transitions.

To trigger a particular transition, we introduce two infinitesimal modifications in the initial condition. For the sake of simplicity, we focus on the bifurcation associated with the warm vortices. To do so, the merging of cold vortices in $x = 500$ km will be forced, by adding an infinitesimal small-scale cold eddy in $(x, y) = (480, 750)$ (in km). The barycentre of northern structures becomes slightly closer to $x = 500$ km. This gives rise to the desired transition, as shown in Figures 5.1 and 5.2. To trigger the bifurcation associated with the southern warm vortices, an infinitesimal small-scale eddy has been added in $(x, y) = (20, 250)$ (in km). If the eddy is warm, the southern barycentre is moved closer to $x = 0$, and the two warm vortices merge near $x = 0$ (scenario 1), as shown in Figure 5.1. If this eddy is cold, the southern barycentre is moved closer to $x = 500$ km, and the two warm vortices merge near $x = 500$ km (scenario 2), as shown in Figure 5.2. The exact expression of the initial condition is the following:

$$\begin{aligned} b(\boldsymbol{x}, t = 0) &= b_0(\boldsymbol{x}) + 0.3 s_w F\left(\frac{1}{40}\left(\boldsymbol{x} - \begin{pmatrix} 20 \text{ km} \\ 250 \text{ km} \end{pmatrix}\right)\right) \\ &\quad -0.3F\left(\frac{1}{40}\left(\boldsymbol{x} - \begin{pmatrix} 480 \text{ km} \\ 750 \text{ km} \end{pmatrix}\right)\right), \end{aligned} \tag{5.3.3}$$

where the large-scale field, b_0, and the two-dimensional Gaussian function, F, remain defined by (5.3.1) and (5.3.2). The factor s_w is set to 1 (respectively -1) if one wants to force the scenario 1 (respectively the scenario 2). The size of the large vortices of b_0 is of the order of the Rossby radius L_d, whereas the small-scale eddies spread only over few kilometers.

The amplitude of the buoyancy and the stratification are set to: $B_0 = 10^{-3} m.s^{-2}$ and $N = 3f_0$. The Coriolis frequency is fixed to $1.028 \times 10^{-4} s^{-1}$, which corresponds to a latitude of 45°. The

reference deterministic high-resolution SQG model is associated with a mesh-grid of 512^2 points. The low-resolution (deterministic or stochastic) SQG models are run with 128^2 points. Simulations are performed through a pseudo-spectral code. For all simulations (deterministic and random, high-resolution and low-resolution), a standard hyperviscosity model has been introduced (Held et al., 1995):

$$D_t b = \alpha^{hv} \Delta^4 b \, \mathrm{d}t, \tag{5.3.4}$$

with a coefficient $\alpha^{hv} = (5 \times 10^{29} m^8.s^{-1}) M_x^{-8}$ where M_x denotes the grid size (*i.e.* 128 or 512).

Scenario 1

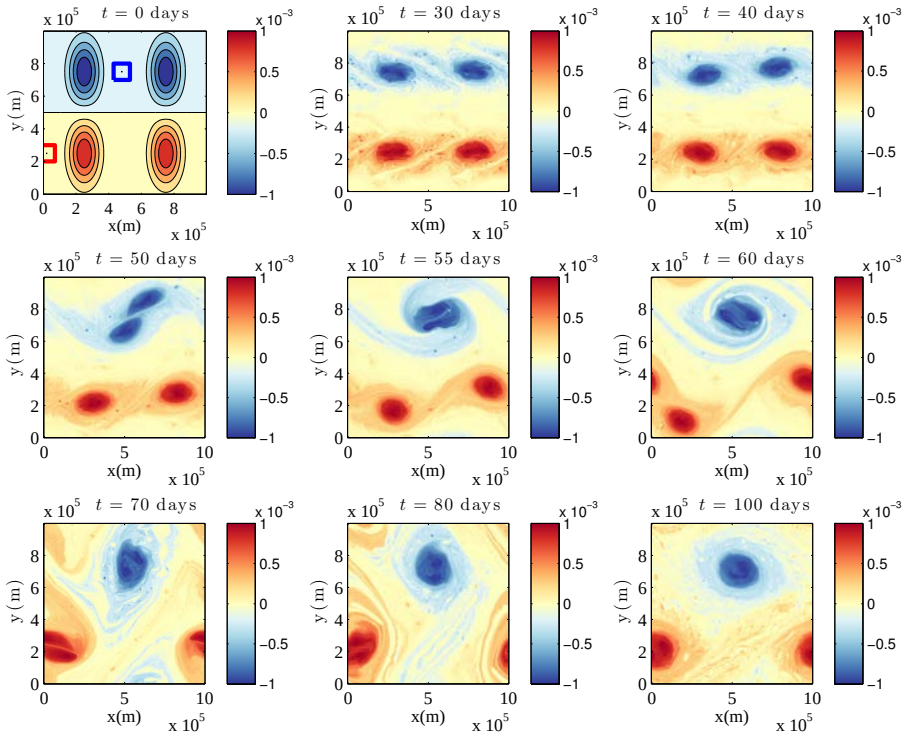


Figure 5.1: Buoyancy ($m.s^{-2}$) at $t = 0, 30, 40, 50, 55, 60, 70$ and 80 days of advection for the SQG model at resolution 512^2. Here, a cold and a warm very small eddies are added respectively in the top and the bottom of the initial condition. These eddies are highlighted by respectively a blue square and a red square. This small difference in the initial condition does not modify the flow until the onset of the symmetry breaking, the 40th day. Since the flow is chaotic, the small perturbation at $t = 0$ determines how the symmetry breaking occurs a month and a half later.

The high-resolution simulations corresponding to scenario 1 and 2 are displayed in Figures 5.1 and 5.2, respectively. At this resolution, the evolution toward scenario 1 or scenario 2 is determined by the value of the parameter s_w (*i.e.*$+/-1$). The associated variations are hardly taken into account by the low-resolved SQG model, as shown in Figure 5.3. The two SQG simulations correspond to 70 days of advection at high and low resolution, respectively. Both simulations have been initialized in the same way ($s_w = 1$). The low-resolution field differs from the high-resolution field, as it apparently followed the wrong transition. As understood, whatever the random or deterministic nature of the tracer evolution law, a dissipation or a filtering at small scales is necessary to remove aliasing effects. In present simulations, the dissipation is created by an hyperviscosity scheme. At low resolution, the initial perturbation is rapidly diffused (few days). When the symmetry breaking occurs, after 40 days of advection, this initial perturbation has been completely forgotten. Moreover, another infinitesimal asymmetry triggers the other likely transition. According to the expression of b_0 (equations (5.3.1) and (5.3.2)), the large scale of the initial condition is not exactly zonally periodic with period 500 km. The southern part is

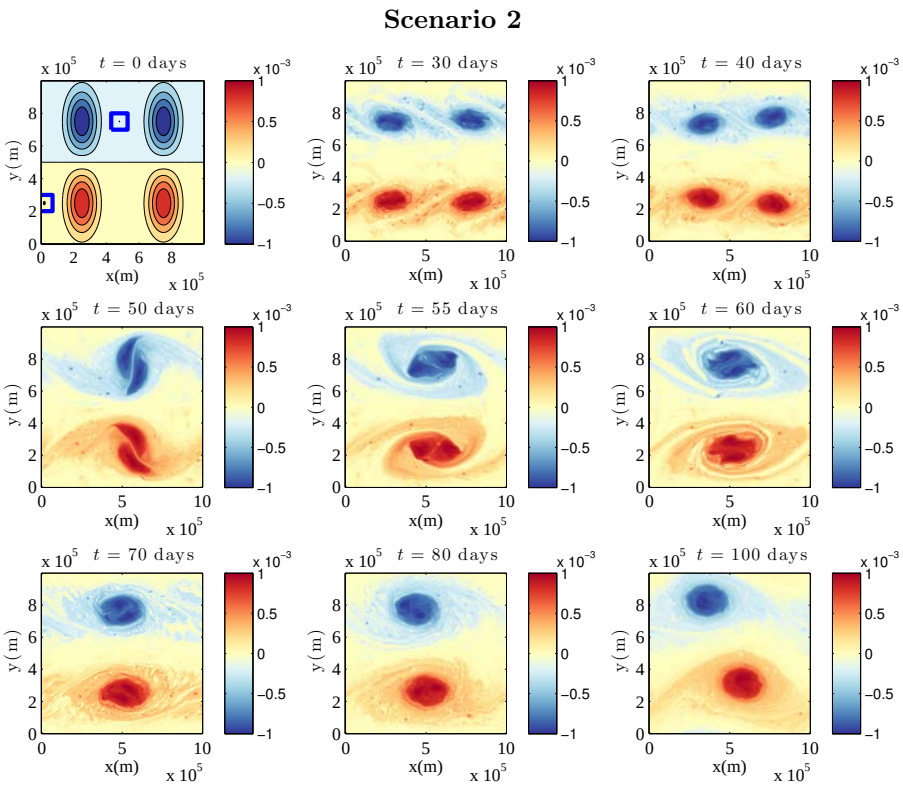


Figure 5.2: Buoyancy $(m.s^{-2})$ at $t = 0, 30, 40, 50, 55, 60, 70$ and 80 days of advection for the usual SQG model at resolution 512^2. Here, two very small cold eddies are added in the top and the bottom of the initial condition. They are highlighted by two blue squares. This small difference in the initial condition does not modify the flow until the onset of the symmetry breaking, the 40th day. Since the flow is chaotic, the small perturbation at $t = 0$ determines how the symmetry breaking occurs a month and a half later.

slightly warmer in the middle. Indeed, the value of b_0 on $(x, y) = (500, 250)$ (in km) is about $1.8 \times 10^{-5} m.s^{-2}$, and on $(x, y) = (0, 250)$ (in km) is about $8.8 \times 10^{-6} m.s^{-2}$. The initial barycentre of the southern structures is thus closer to $x = 500$ km. This asymmetry has a very weak amplitude but a large spatial length scale which prevents its diffusion. This explains the merging in the wrong location.

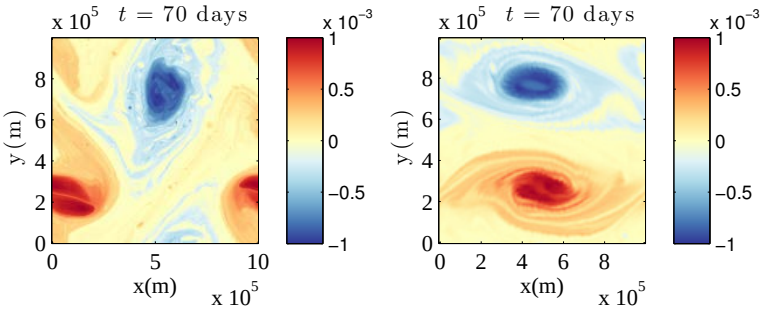


Figure 5.3: Buoyancy ($m.s^{-2}$) after 70 days of advection for the SQG model at resolution 512^2 (left) and at resolution 128^2 (right). The small-scale perturbation in the initial condition activating the symmetry breaking in the reference simulation (512^2), is dissipated in few days in the low-resolution simulation. This makes this single low-resolution simulation erroneous.

In the next section, we will show that low-resolution simulations of appropriate randomized dynamics can retrieve the right scenario.

5.4 Stochastic analysis

With low-resolution simulations of the $\mathrm{SQG_{MU}}$ model described in 5.2.4, ensemble analysis can be performed. The unresolved velocity component is assumed to be isotropic, with a $(-\frac{5}{3})$ spectrum restricted to the small spatial scales, as illustrated by Figure 5.4. The energy is then specified by the diffusion coefficient $\frac{a_H}{2} = 9\ m^2.s^{-1}$. The numerical simulation of this component is detailed in Appendix 5.B. A MATLAB code simulating the SQG_{MU} model is available online (`http://vressegu.github.io/sqgmu`).

Unlike the SQG model, the random simulations of the SQG_{MU} model, with the exact same initialization $s_w = 1$, yields several likely transitions. In Figure 5.5, we show two realizations. One of those realizations corresponds to the reference scenario (scenario 1), the other does not. The model encodes several likely transitions, and thus several potential scenarios. Indeed, the random forcing provides various small-scale perturbations that may trigger these transitions. As this triggering is random, the large-scale changes are also random. In other words, there is a backscattering of uncertainty toward the large scales, as illustrated in Figure 5.6. We decomposed the mean omni-directional spectrum, *i.e.* the mean energy at a given scale, $\hat{\mathbb{E}}\{\overline{\Gamma}_b\}$, into the spectrum of the mean tracer, $\overline{\Gamma}_{\hat{\mathbb{E}}\{b\}}$, (blue line) and the mean spectrum of the tracer random component, $\hat{\mathbb{E}}\left\{\overline{\Gamma}_{b-\hat{\mathbb{E}}\{b\}}\right\}$, (shaded grey):

$$\hat{\mathbb{E}}\{\overline{\Gamma}_b\} = \hat{\mathbb{E}}\left\{\frac{1}{A}\oint \|\boldsymbol{k}\|_2 |\hat{b}|^2 \mathrm{d}\theta_{\boldsymbol{k}}\right\}, \tag{5.4.1}$$

$$= \frac{1}{A}\oint \|\boldsymbol{k}\|_2 \Bigg(\underbrace{|\hat{\mathbb{E}}\{\hat{b}\}|^2}_{\substack{\text{Energy of}\\ \text{the mean}}} + \underbrace{\hat{\mathbb{E}}\left\{|\hat{b} - \hat{\mathbb{E}}\{\hat{b}\}|^2\right\}}_{\text{Variance}}\Bigg) \mathrm{d}\theta_{\boldsymbol{k}}, \tag{5.4.2}$$

$$= \overline{\Gamma}_{\hat{\mathbb{E}}\{b\}} + \hat{\mathbb{E}}\left\{\overline{\Gamma}_{b-\hat{\mathbb{E}}\{b\}}\right\}, \tag{5.4.3}$$

where A denotes the area of the domain Ω, $\hat{\mathbb{E}}\{f\}$ the empirical mean of f, computed from the ensemble, and $\hat{f}$ the Fourier transform of f.

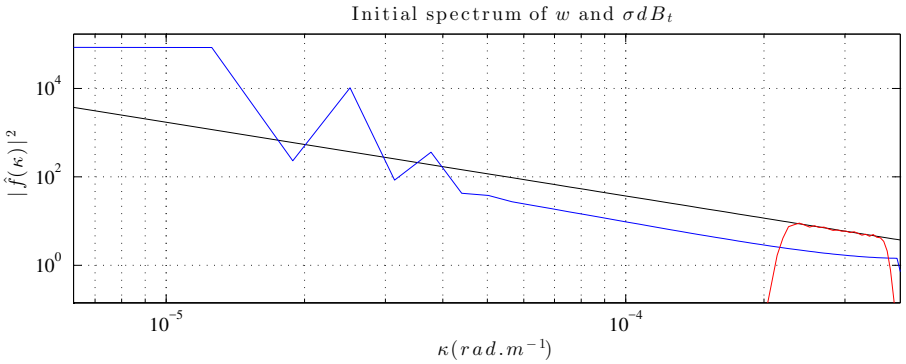


Figure 5.4: Spectrum ($m^2.s^{-2}/(\text{rad}.m^{-1})$) of $\boldsymbol{w} = \boldsymbol{u}$, at the initial time, in blue, spectrum of $\boldsymbol{\sigma}_H \dot{\boldsymbol{B}}$ (up to a multiplicative constant) for the SQG_{MU} simulation, in red, and slope $-\frac{5}{3}$ in black. In the simulation performed, $\boldsymbol{\sigma}_H \dot{\boldsymbol{B}}$ is restricted to a narrow spectral band. Thus, this velocity component acts almost only near the resolution cutoff, where the large-scale component, $\boldsymbol{w}$, has a low energy.

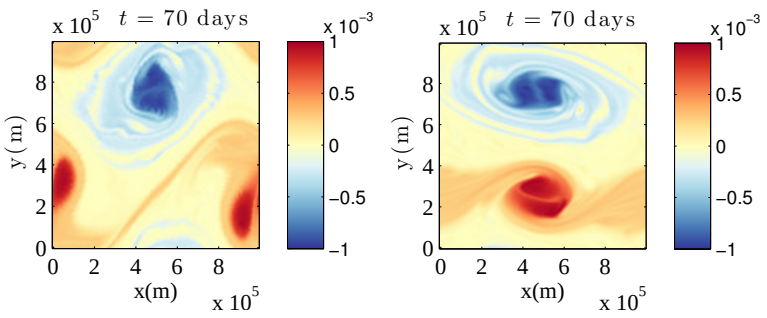


Figure 5.5: Two realizations of buoyancy ($m.s^{-2}$) after 70 days of advection for the SQG_{MU} at resolution 128^2. Event though the small-scale perturbation in the initial condition is dissipated in few days, the small-scale component of SQG_{MU} triggers randomly the symmetry breaking. Therefore, some realizations follow the right transitions (left) and some do not (right).

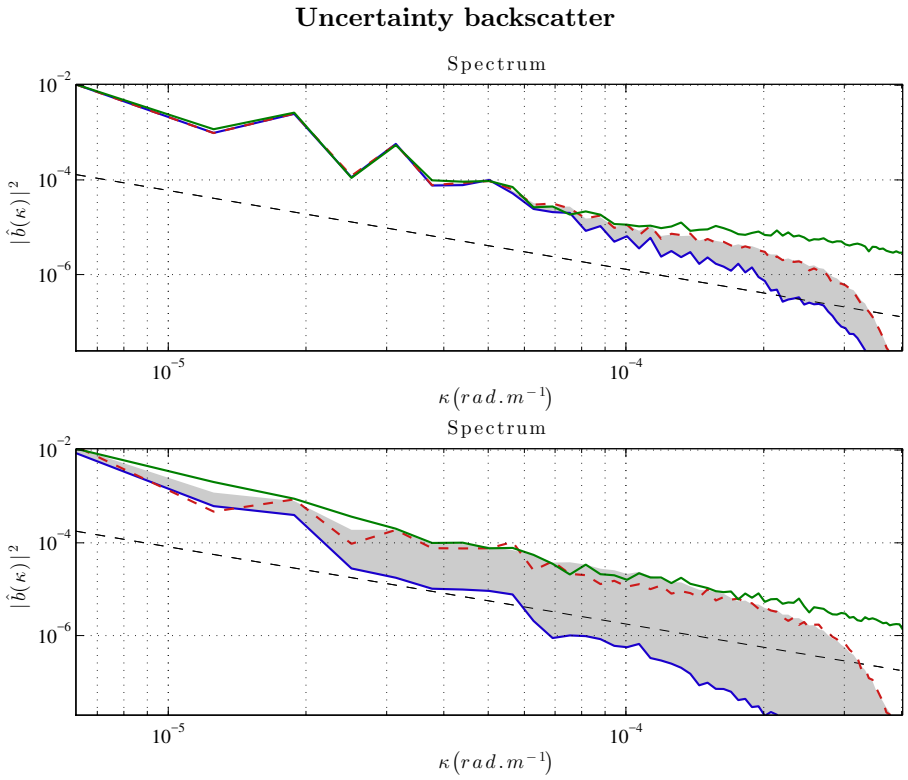


Figure 5.6: Tracer spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$) after 30 days (top) and 70 days (bottom) of advection for SQG model at resolution 512^2 (green), one realization of SQG_{MU} model, $\overline{\Gamma}_{b^{(1)}}$, at resolution 128^2 (red dashed line), the spectrum of the mean, $\overline{\Gamma}_{\hat{\mathbb{E}}\{b\}}$, (blue line) and the mean spectrum of the tracer random component, $\hat{\mathbb{E}}\{\overline{\Gamma}_{b-\hat{\mathbb{E}}\{b\}}\}$, (shaded grey). After being stacked, the two last plots represent the mean spectrum: $\hat{\mathbb{E}}\{\overline{\Gamma}_b\} = \overline{\Gamma}_{\hat{\mathbb{E}}\{b\}} + \hat{\mathbb{E}}\{\overline{\Gamma}_{b-\hat{\mathbb{E}}\{b\}}\}$. The more thick the shaded grey area is, the more variance is contained at this scale. At $t = 30$ days, the variance remains at small scales but this small-scale uncertainty activates the symmetry breaking. This results in a variance backscattering with a thickening of the spectrum of the random component at large scales (visualized at $t = 70$). Since large scales influence strongly the small scales the small-scale variance is also enforced.

One can then wonder about the number of realizations following the right scenario. This necessitates the analysis of the 4-dimensional spatio-temporal random field $\left(b^{(i)}(x_j, y_k, t_l)\right)_{ijkl}$. The superscript (i) designates the i-th realization of the ensemble. To reduce the associated dimension, a Principal Component Analysis (PCA), also termed Empirical Orthogonal Function (EOF) representation, is performed over the realizations, at a fixed time t. Within this analysis, unlike usual EOF representation, the time axis is replaced by the realization index. At a given time, it helps to represent the whole ensemble (200 realizations of the random field) by the ensemble mean field and few other EOF spatial modes. In the present case, the buoyancy is approximated as:

$$b^{(i)}(\boldsymbol{x}, t) = \hat{\mathbb{E}}(b)(\boldsymbol{x}, t) + \sum_{n=1}^{N_e} c_n^{(i)}(t)\Psi(\boldsymbol{x}, t), \tag{5.4.4}$$

$$\approx \hat{\mathbb{E}}(b)(\boldsymbol{x}, t) + \sum_{n=1}^{N_{EOF}} c_n^{(i)}(t)\Psi_n(\boldsymbol{x}, t), \tag{5.4.5}$$

where N_e is the size of the ensemble, $N_{EOF} \ll N_e$ is the number of EOF modes chosen to described the whole ensemble and the $(\Psi_n)_{1 \leqslant n \leqslant N_{EOF}}$ denote the EOF spatial modes. Those spatial fields are orthogonal:

$$\frac{1}{M}\sum_{j=1}^{M} \Psi_n(\boldsymbol{x}_j, t)\Psi_m(\boldsymbol{x}_j, t) = \delta_{nm}, \tag{5.4.6}$$

with M the number of grid points. The mean energies – or variances – of EOF coefficients correspond to the eigenvalues of the two-points correlation matrix; they are ordered in decreasing order and represent the energies associated with each spatial mode. To describe the ensemble with respect to a maximal variance point of view, only the EOF coefficients c_n concentrating the largest part of the buoyancy mean energy are kept. This energy, which differs from the energy of the mean, reads:

$$\underbrace{\hat{\mathbb{E}}\left\{\frac{1}{M}\sum_{j=1}^{M} b^2(\boldsymbol{x}_j, t)\right\}}_{\text{Mean energy}} = \frac{1}{M}\sum_{j=1}^{M}\left(\hat{\mathbb{E}}(b)\right)^2(\boldsymbol{x}_j, t) + \frac{1}{M}\sum_{j=1}^{M}\widehat{Var}\,(b)\,(\boldsymbol{x}_j, t), \tag{5.4.7}$$

$$= \underbrace{\frac{1}{M}\sum_{j=1}^{M}\left(\hat{\mathbb{E}}(b)\right)^2(\boldsymbol{x}_j, t)}_{\text{Energy of the mean}} + \sum_{n=1}^{N_e} \underbrace{\hat{\mathbb{E}}\left\{c_n^2(t)\right\}}_{\substack{\text{Mean energy of}\\ \text{the } n\text{-th EOF}}}. \tag{5.4.8}$$

In Figure 5.7, left part, the energy associated with the different spatial modes is displayed. At $t = 30$ days, the energy of the mean field, denoted as a 0-th order EOF (index+1 = 1), is much larger than the variance field, described by the other EOFs. On the contrary, at $t = 70$ days, after the breaking symmetry, the energy of the mean and the variance have the same order of magnitude. The variance is mainly explained by the fist EOF. Thus, at the first order, the randomness of the tracer is approximately encoded by this first EOF coefficient, c_1. Its probability density function and the joint probability density function for the two first EOF coefficients presented in Figure 5.7 are unimodal at $t = 30$ days and bimodal at $t = 70$ days. Note that the tracer is clearly non-Gaussian. The symmetry breaking has created two likely scenarios in the ensemble. The scenario corresponding to negative values of the first EOF coefficient (probability of 47%) is called scenario A and the scenario corresponding to positive values (probability of 53%) is called scenario B. The red line at zero separates the probability density function between the two scenarios. In Figure 5.8, the same probability density function along time is plotted. The bifurcation is clearly visible. Also shown, the mean buoyancy ($m.s^{-2}$) of the two likely scenarios are represented after 70 days

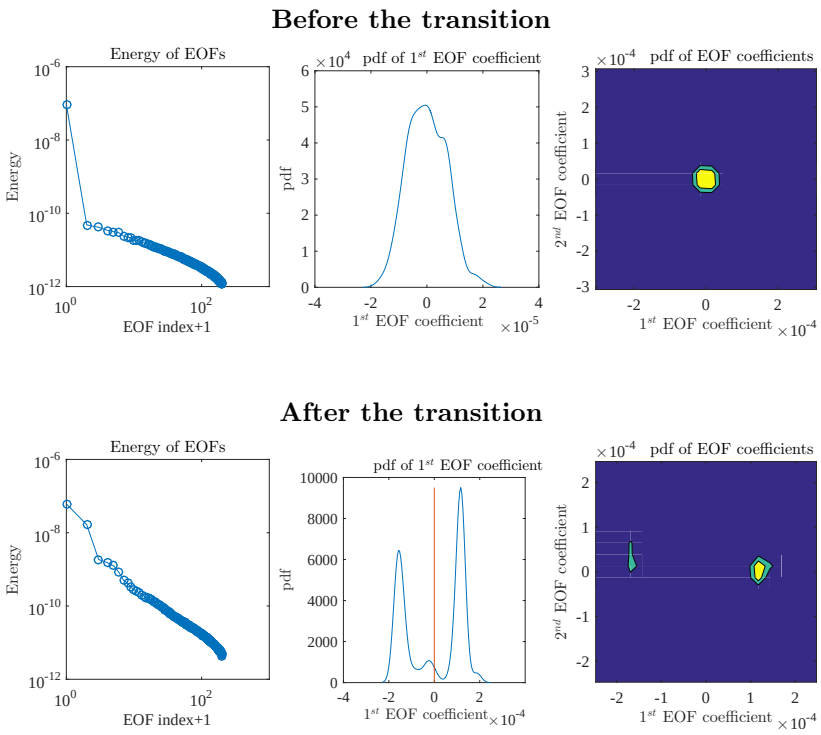


Figure 5.7: Energy of the EOFs ($m^2.s^{-4}$) (left), probability density function of the first EOF coefficient (middle) and joint probability density function for the two first EOF coefficients (right) after 30 days (top) and 70 days (bottom). The ensemble is simulated according to the SQG model under moderate uncertainty at resolution 128^2. At $t = 30$ days, the energy of the mean field, denoted as a 0-th order EOF (index+1 = 1), is much higher than the variance field, described by the other EOFs. On the contrary, at $t = 70$ days (after the symmetry breaking), the energy of the mean and variance have the same order of magnitude. The variance is mainly explained by the first EOF. Therefore, at the first order, the random component of the tracer can be approximated by this first EOF. The probability density function of the first EOF coefficient and the joint probability density function for the two first EOF coefficients are unimodal at $t = 30$ days and bimodal at $t = 70$ days. The breaking symmetry has created two likely scenarios, which are very different from one another. The scenario A corresponds to negative values of the first EOF coefficient (probability of 47%) whereas the scenario B corresponds to positive value (probability of 53%). The red line separates the probability density function between the two scenarios.

of advection. The two mean fields, $\hat{\mathbb{E}}(b|1)$ and $\hat{\mathbb{E}}(b|2)$, are calculated as:

$$\hat{\mathbb{E}}(b|1) = \hat{\mathbb{E}}(b) + \hat{\mathbb{E}}(c_1|1)\Psi_1 = \hat{\mathbb{E}}(b) + \left(\frac{1}{\#\left\{ i|c_1^{(i)} < 0 \right\}} \sum_{c_1^{(i)}<0} c_1^{(i)} \right) \Psi_1, \tag{5.4.9}$$

$$\hat{\mathbb{E}}(b|2) = \hat{\mathbb{E}}(b) + \hat{\mathbb{E}}(c_1|2)\Psi_1 = \hat{\mathbb{E}}(b) + \left(\frac{1}{\#\left\{ i|c_1^{(i)} \geqslant 0 \right\}} \sum_{c_1^{(i)}\geqslant 0} c_1^{(i)} \right) \Psi_1, \tag{5.4.10}$$

where # stands for the cardinality of a set. The scenario A is quite close to the scenario 1, which is the reference with this initial condition, whereas scenario B is close to scenario 2. The stochastic model has enabled the ensemble to track both scenarios and to describe them statistically. Let us point out that the shape of the isotropic small-scale velocity expression has been loosely fixed by an *a priori* form of the spectrum. Some learning procedures of the noise topology from past data could lead to express more informative heterogeneous random fields, and to statistically favor the most likely transition. Moreover, since the two scenarios are very different, the introduction of few observations, through an assimilation procedure, could very easily help to select the right scenario.

For sake of comparison, we also show results obtained using the deterministic SQG model, initialized with random conditions. Initial small-scale buoyancy perturbations are assumed Gaussian and sampled from a $(-\frac{5}{3})$ spectrum, as shown in Figure 5.9. These perturbations should not change the large-scale flow before the predictability time (about one month). Accordingly, we require those perturbations to be of small amplitudes and restrict them to small scales. The same analysis is performed, including the EOF decomposition and the distinction between two likely scenarios. Figure 5.10 gathers the results. After 70 days of advection, the scenario A is hardly visible in the probability density function of the first EOF. The ensemble estimates a probability of only 39% for the reference scenario. The SQG_{MU} ensemble estimated a probability of 47%. Moreover, the probability density function is very noisy. This suggests that the ensemble may not be converged, *i.e.* the empirical statistics of the ensemble will change if the ensemble size grows. The scenario A completely disappears in the joint probability density function for the two first EOF coefficients. The probability density function of the first EOF along time, before and after the symmetry breaking, exhibits very narrow branches associated with high probabilities, compared to Figure 5.8. It indicates that randomized initial conditions may lead to underdispersive ensemble. Furthermore, in Figure 5.10, trajectories of some realizations are still visible after the bifurcation. This confirms that the probability density function did not converge. Indeed, to estimate this density, we use the well-known Parzen-Rosenblatt estimator (Rosenblatt, 1956; Parzen, 1962): each realization is associated with a kernel and the estimator is the sum of those kernels. Here, some realizations or set of few realizations are isolated and create spikes in the estimator. More realizations would be needed to have almost continuously distributed realizations. In other words, the ensemble is not converged. This drawback could be expected for at least two reasons. First, the structure of the initial noise contains little physical information, while the dimension of the state space is huge. Without phase information, covering all the possibilities requires a very large number of realizations. Furthermore, the subgrid tensor diffuses the small-scales components of the tracer where the ensemble variability is encoded. This is a known features of ensemble forecasts: ensemble members tend to align with most unstable directions of the dynamics (Trevisan and Uboldi, 2004; Ng et al., 2011; Gottwald and Harlim, 2013; Bocquet et al., 2016). Since small scales are stabilized by the subgrid tensor, the ensemble shrinks to span a smaller large-scale unstable subspace (Sapsis, 2013). On the contrary, the stochastic model associates phase and intermittency with the noise and continuously injects it into the dynamics. The phase information or inhomogeneity as well as the non-Gaussianity come from the multiplicative structure. Even though the uncorrelated velocity is only prescribed by a spectrum, the tracer gradients have phase and dynamically constraint the regions of application of the noise. This process makes the stochastic forcing much more efficient. Hence, a smaller number of realizations are needed. In Figure 5.10, the convergence of the probability density function of scenario B (positive values of the first EOF coefficient) seems slightly better than the density of scenario A. Unfortunately, the bottom Figures shows that the scenario B is not the one followed by the high-resolution simulation. Let us note that the reference is deterministic. Accordingly, the

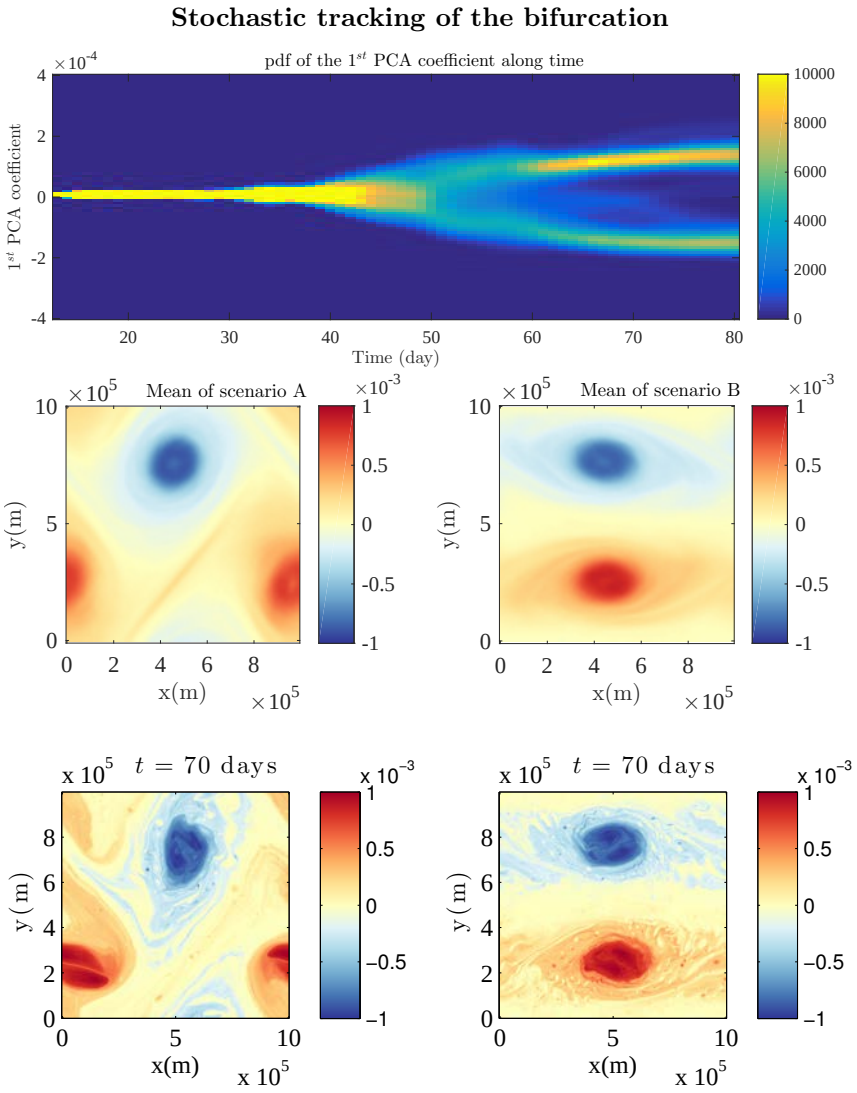


Figure 5.8: Probability density function of the first EOF coefficient along time (top), buoyancy ($m.s^{-2}$) after 70 days of advection for the mean of the two likely scenarios A and B of the $\mathrm{SQG_{MU}}$ model at resolution 128^2 (respectively middle left and middle right) and the reference scenarios 1 and 2 at high resolution 512^2 (respectively bottom left and bottom right). The bifurcation is clearly visible on the top plot. The two likely scenarios differ from the sign of the first EOF coefficient. They are almost associated with an equal probability of occurrence: a probability of 47% for scenario A and a probability of 53% for scenario B. The scenario A (respectively B) is similar to the scenario 1 (respectively 2).

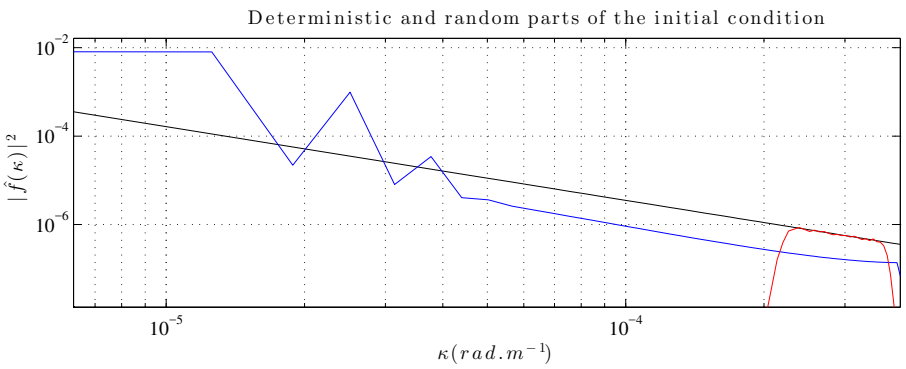


Figure 5.9: Spectrum ($m^2.s^{-4}/(\text{rad}.m^{-1})$), at the initial time, of the mean buoyancy, in blue, spectrum of its random perturbation, in red, and slope $-\frac{5}{3}$ in black. The initial perturbation is restricted to a narrow spectral band. This random initial condition has been used to simulate an ensemble with the deterministic SQG model.

reference probability density is a dirac measure. Indeed, the deterministic reference initial condition is assumed to be known and is used in all large-scale simulations. Moreover, the reference dynamic defined by the high-resolution SQG model is deterministic as the real ocean dynamics is. The bad description of the scenario A tends to confirm that the SQG model with randomized initial conditions fails to describe the bifurcation.

Another argument is the distance between the reference and each ensemble. In Figure 5.11 reports the error corresponding to the realization closest to the reference, *i.e.* with minimal error, for each ensemble. In geophysical data assimilation, large confidence is often given to observations. Thus, the maximum a posteriori estimator is almost equal to the minimal error realization. Figure 5.11 shows that the ensemble from the stochastic model is closer to the reference than the ensemble with randomized initializations. It suggests that the stochastic method should lead to a better maximum a posteriori estimator.

5.5 Conclusion

In this study, long-time forecasts of two different SQG models have been compared. The first one corresponds to the classical SQG equations with a random initial condition. The second one is a stochastic version of SQG, derived from a stochastic expression of the transport equations (Mémin, 2014). Both models are compared to a high-resolution simulation reference.

The chosen high-resolution reference is subject to a bifurcation after 40 days of advection. An infinitesimal modification of the initial condition determines the global shape of the flow two months later. Depending on the value of this initial modification, two different scenarios are isolated. For the same initial condition, the deterministic high-resolution and the low-resolution simulations do not follow the same transition. Indeed, the sub-grid tensor associated with the coarser resolution diffuses the crucial initial perturbation before the transition. This makes the deterministic forecast useless. This result questions the classical definition of predictability and associated error which only rely on initial perturbations (Lorenz, 1969). Note that before the symmetry breaking, the large-scale errors induced by slight modifications of the initial conditions are negligible in front of large-scale model errors. To recover, the true scenario at low resolution, a possible solution could be to randomize the initial condition. As shown, this solution would require a large number of realizations. At the opposite, the proposed stochastic model tracked both scenarios. Its efficiency is mainly explained by the continuous injection of multiplicative noise. This structure dynamically constrains both the phase and the intermittency of the noise. The model achieved to predict the likelihood and the point-wise tracer probability density in each case.

To identify and separate the scenarios, a simple threshold is sufficient regarding the high energy distributed along the first PCA axis. For more realistic flow, the number of scenarios is likely to

Same diagnostic based on a deterministic model with random initial conditions

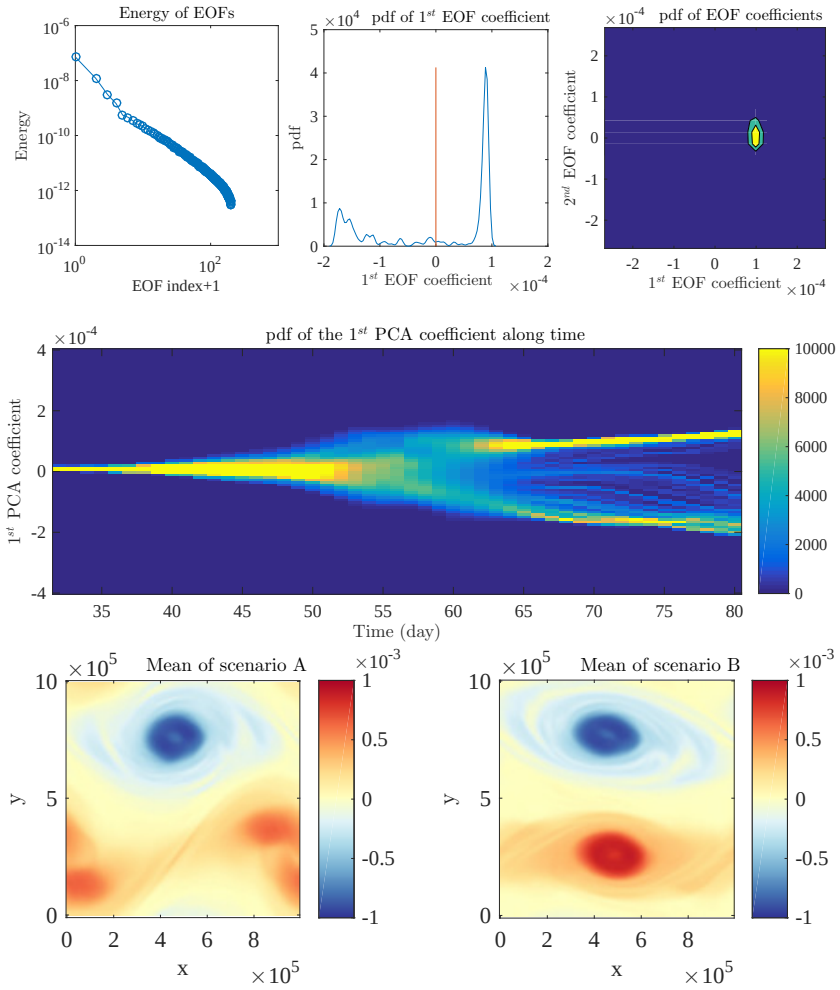


Figure 5.10: Energy of the EOFs ($m^2.s^{-4}$) (top left), probability density function of the first EOF coefficient (top middle), joint probability density function for the two first EOF coefficients (top right) after 70 days, probability density function of the first EOF coefficient along time (middle), buoyancy ($m.s^{-2}$) after 70 days of advection for the means of the two likely scenarios A and B (bottom left and bottom right respectively). The ensemble is simulated according to the usual deterministic SQG model with random initial conditions at resolution 128^2. The joint probability density function for the two first EOF coefficients suggests only one likely scenario. The central Figure confirms that this model fails to correctly describe the bifurcation. Indeed, the probability density function appears to be not converged in this case. The bottom Figures show that the worst resolved scenario (the scenario A) is the one similar to the true reference scenario (the scenario 1).

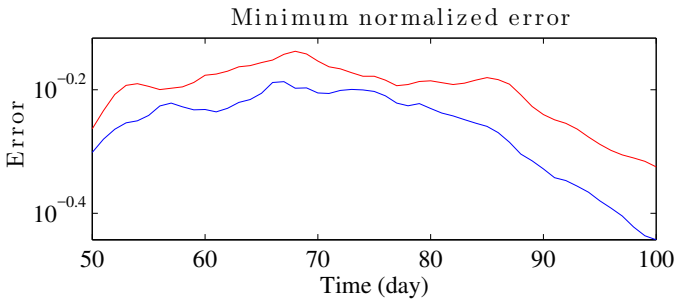


Figure 5.11: Minimum normalized error along time in the ensemble with random initial conditions (red) and in the ensemble with random dynamics (blue). The square error was integrated over the space and divided by the energy of the reference.

be larger. Hence, this classification may be inadequate and more advanced clustering methods (e.g. kmeans and expectation-maximization algorithms) are probably required. In this paper, this ensemble aggregation was used as a diagnosis. But, it could also improve filtering methods. Indeed, for geophysical flows, the state-space dimension being large, the ensemble size is generally too small to encode all possibilities. So, when an observation is assimilated, only few realizations of the ensemble are close enough, and all the others are considered useless. This often leads to filter degeneracy. Considering distance to likely scenarios rather than distance to realizations may help preventing this deficiency.

As expressed transport under location uncertainty involves an inhomogeneous and anisotropic diffusion, a drift correction and a multiplicative noise. All these terms are related to the sub-grid velocity statistics. This transport can also provide a systematic derivation method for stochastic large-scale fluid dynamics models. While the analysis of this paper is applied to a randomized version of the SQG model with a toy initial condition, the multiplicative structure of the noise and the balance with diffusion hold for any fluid dynamics models under location uncertainty. This suggests that similar conclusions could be expected in more complex problems. We shall consider this question in future works. Hence, this systematic derivation method may open for new ensemble forecasts methods, especially for climate projections where uncertainty quantification is a main issue (Allen and Stainforth, 2002).

Appendix

5.A Stochastic material derivative

The Reynolds transport theorem provides the expression of the rate of change of a scalar function, q, within a material volume, $\mathcal{V}(t)$, transported by the flow. Relying on a stochastic Lagrangian expression of the fluid particles' trajectory (5.2.5), a stochastic extension of this theorem has been derived by Mémin (2014):

$$\mathrm{d}\int_{\mathcal{V}(t)} q = \int_{\mathcal{V}(t)} \left(\mathrm{d}_t q + \Big[\boldsymbol{\nabla}\cdot\big(q(\underbrace{\boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T}_{\boldsymbol{w}^\star})\big) - \frac{1}{2}\sum_{i,j=1}^{d} \partial_i(a_{ij}\partial_j q)\Big]\mathrm{d}t + \boldsymbol{\nabla} q\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \right). \quad (5.A.1)$$

For an isochoric stochastic flow ($\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star = 0; \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$) it is easy to see that the change of rate expression (5.A.1) corresponds to the following stochastic version of the material derivative:

$$\begin{aligned}\mathrm{d}\int_{\mathcal{V}(t)} q &= \int_{\mathcal{V}(t)} D_t q, \\ &= \int_{\mathcal{V}(t)} \left(\mathrm{d}_t q + \Big[\boldsymbol{\nabla} q\cdot\boldsymbol{w}^\star - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{\nabla} q)\Big]\mathrm{d}t + \boldsymbol{\nabla} q\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \right). \end{aligned} \quad (5.A.2)$$

5.B Simulation of the random velocity

To simulate the $\mathrm{SQG_{MU}}$ model (equations (5.2.12) and (5.2.15)), the covariance of the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ must be specified. A simple choice consists in relying on a solenoidal homogeneous field similar to the Kraichnan model (Kraichnan, 1968, 1994; Gawędzki and Kupiainen, 1995; Majda et al., 1999):

$$\boldsymbol{\sigma}_H(x)\mathrm{d}\boldsymbol{B}_t = \int_{\mathbb{R}^2} d\boldsymbol{z}\breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z})\mathrm{d}\boldsymbol{B}_t(\boldsymbol{z}) = \int_{\mathbb{R}^2} d\boldsymbol{z}\breve{\boldsymbol{\sigma}}(\boldsymbol{x}-\boldsymbol{z})\mathrm{d}\boldsymbol{B}_t(\boldsymbol{z}), \tag{5.B.1}$$

$$\text{with } \boldsymbol{\nabla}\cdot\breve{\boldsymbol{\sigma}} = 0. \tag{5.B.2}$$

Note that the proposed model handles spatially inhomogeneous field, which can be much more physically relevant. However, homogeneity greatly simplifies the random field parameterization and simulation. Indeed, homogeneity in physical space implies independence between Fourier modes, $\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})$, in the half-space $\boldsymbol{k}\in\mathbb{R}\times\mathbb{R}^+$. As the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ is further assumed to be divergence-free and isotropic, it can be conveniently specified from its two-dimensional spectrum $\boldsymbol{k}\to\frac{1}{\Delta t}|\hat{\breve{\boldsymbol{\sigma}}}(\|\boldsymbol{k}\|)|^2$. To follow the SQG turbulence, the omni-directional spectrum slope is fixed to $-\frac{5}{3}$, which corresponds to the usual SQG spectrum slope (Blumen, 1978; Klein et al., 2008). According to our stochastic framework, the energy of this small-scale velocity is fixed by the diffusion coefficient a_H and the simulation time step, Δt:

$$\mathbb{E}\left\{\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)\left(\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right)^T\right\} = \frac{1}{\Delta t}\boldsymbol{a}_H = \frac{1}{\Delta t}\begin{pmatrix} a_H & 0 \\ 0 & a_H\end{pmatrix}. \tag{5.B.3}$$

The time step depends itself, through the CFL conditions, linked either to the advection term or to the diffusion terms, on the spatial resolution and on the maximum magnitude of the resolved velocity. The unresolved velocity should be energetic only where the dynamics cannot be resolved properly, meaning between the effective resolution and the true resolution $\frac{\pi}{\Delta x}$. Consequently, a smooth band-pass filter, f_{BP} is introduced which has non-zero value only between two wave-numbers κ_{min} and κ_{max}. This specifies the scale range over which the unresolved component spectrum is non zero. The parameter κ_{min} is inversely related to the spatial correlation length of the unresolved component. The Fourier transform of this component is finally defined as:

$$\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k}) \triangleq \frac{A}{\sqrt{\Delta t}}\, i\boldsymbol{k}^{\perp} f_{BP}\left(\|\boldsymbol{k}\|\right)\|\boldsymbol{k}\|^{-\alpha}\,\frac{\widehat{dB_t}}{\sqrt{\Delta t}}(\boldsymbol{k}) \text{ with } 3-2\alpha = -\frac{5}{3}, \tag{5.B.4}$$

where A is a constant, which is set such that $\mathbb{E}\left\|\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}\right\|_2^2 = 2\frac{a_H}{\Delta t}$ (see equation (5.B.3) above), $\widehat{dB}_t$ is the spatial Fourier transform of dB_t and $\frac{dB_t}{\sqrt{\Delta t}}$ is a discrete scalar white noise process of variance 1 in space and time. In such a parameterization, the small-scale velocity is easy to sample. We first sample $\frac{dB_t}{\sqrt{\Delta t}}$ then we get $\frac{\widehat{dB_t}}{\sqrt{\Delta t}}$ and finally $\widehat{\boldsymbol{\sigma}_H\dot{\boldsymbol{B}}}(\boldsymbol{k})$ with the above equation. Figure 5.4 illustrates the definition (5.B.4). Finally, only three parameters, a_H, κ_m and the spectral slope have too be specified in this homogeneous case. They are defined by the usual diffusion coefficient, the effective resolution and the dynamics of interest respectively. When the typical spectral slope is unknown, it may be identified to the instantaneous spectral slope of the resolved velocity.

For the inhomogeneous case, there are many ways of prescribing this unresolved velocity. Without prior learning, one could for instance make one or several of the previous parameters be space-dependent. Many deterministic realistic geophysical simulations already involve an anisotropic and inhomogeneous eddy diffusivity coefficient. The effective resolution and the spectrum slope could be defined based on the spectrum of the resolved velocity. If small-scale observations are available, useful statistics or their relations with the large-scale velocity could be learned (e.g. Gottwald et al., 2015).

The latter is specific to two-dimensional domains. In three dimension, the three-dimensional curl and 3 independent $\mathbb{I}_d$-cylindrical Wiener processes must be considered to simulate the small-scale velocity. If boundaries conditions are not periodic, the convolution (5.B.1) has to be implemented instead of the multiplication in Fourier space (5.B.4).

Chapter 6

Reduced flow models from a stochastic Navier-Stokes representation

V. Resseguier, E. Mémin, and B. Chapron. Reduced flow models from a stochastic Navier-Stokes representation. *Annales de l'ISUP*, 2015

Abstract

In large-scale Fluids Dynamics systems, the velocity lives in a broad range of scales. To be able to simulate its large-scale component, the flow can be decomposed into a finite variation process, which represents a smooth large-scale velocity component, and a martingale part, associated with the highly oscillating small-scale velocities. Within this general framework, a stochastic representation of the Navier-Stokes equations can be derived, based on physical conservation laws. In this equation, a diffusive sub-grid tensor appears naturally and generalizes classical sub-grid tensors.
Here, a dimensionally reduced large-scale simulation is performed. A Galerkin projection of our Navier-Stokes equation is done on a Proper Orthogonal Decomposition basis. In our approach of the POD, the resolved temporal modes are differentiable with respect to time, whereas the unresolved temporal modes are assumed to be decorrelated in time. The corresponding reduced stochastic model enables to simulate, at low computational cost, the resolved temporal modes. It allows taking into account the possibly time-dependent, inhomogeneous and anisotropic covariance of the small scale velocity. We proposed two ways of estimating such contributions in the context of POD-Galerkin.
This method has proved successful to reconstruct energetic *chronos* for a wake flow at Reynolds 3900, even with a large time step, whereas standard POD-Galerkin diverged systematically. This paper describes the principles of our stochastic Navier-Stokes equation, together with the estimation approaches, elaborated for the model reduction strategy.

Keywords

Stochastic calculus, fluid dynamics, large eddy simulation, Proper Orthogonal Decomposition, reduced order model, uncertainty quantification

6.1 Introduction

Modeling accurately and understanding geophysical fluid dynamics is a main issue in current researches. Indeed, beyond economic applications linked to weather forecasting, the need for accurate climate projections is becoming more and more important. Studying such systems using physics

is challenging, especially in regard to the non-linearity of the Navier-Stokes equations. Since these equations make large-scale velocities interact with small-scale velocity fluctuations, the main velocity tendency cannot be simulated alone (Slingo and Palmer (2011) and Palmer and Williams (2008)). The effects of the unresolved small-scale, so-called turbulent, fluctuations have to be taken into account.

A first way of modeling 3D turbulence is to only consider the direct energy cascade, which means that energy goes from the large scales to smaller scales, until it is dissipated by molecular viscosity (Vallis (2006)). This can be done by assuming that the large-scale velocity follows the Navier-Stokes equation with an additive diffusive term, parametrized by a so-called eddy viscosity (Kraichnan (1987)). Such types of additive terms are called sub-grid tensors, since they represent the effect of velocity living at smaller scale than the simulation grid. However, the diffusion matrix or diffusion coefficient and its temporal and spatial dependence have to be determined. In many cases, it is done empirically and/or using scaling assumptions. This is the case for LES (Large Eddies Simulation) where the large-scale flow is defined by a spatially low-pass-filtered velocity (Parnaudeau et al. (2008),Lesieur and Metais (1996)) and RANS (Reynolds Average Numerical Simulation) where the large-scale component is defined by a statistical average of the velocity. The same type of models are used for large-scale modeling of tracers evolution, using eddy diffusivity instead of eddy viscosity.

Another drawback of this approach is the assumption of a permanent direct energy cascade. In real systems, there are intermittent back-scattering of energy from smaller scales toward the larger one. Therefore, some authors proposed to include terms that artificially bring energy to the system. Sometimes, it is done by a locally negative eddy viscosity (Protas et al. (2015)). Another solution consists in setting up a system forced by a Gaussian process decorrelated in time (Kupiainen (2000)). The spatial covariance of this forcing is a parameter that has to be determined. A stationary assumption greatly simplifies the problem and the associated model.

Considering a random velocity is now widely used (Slingo and Palmer (2011), Palmer and Williams (2008) and Franzke et al. (2015)). In addition to theoretical physics constraints, it enables uncertainty quantification, and the use of ensemble based methods such as filtering (Doucet et al. (2001), Doucet and Johansen (2009) and Candy (2011)). The stochastic model, described in the previous paragraph, adds a random force to the equation without deep theoretical justification. This additive noise can be interpreted as an explicit error of the model. However, within this prospect, why would an additive noise be more adequate than a multiplicative noise or any other model of noise? According to Majda et al. (1999), in reduced models for geophysical fluid dynamics applications, the coupling of an additive and a multiplicative noise is a good choice, leading naturally to heavy tails processes. But as far as we know, no theoretical justifications of this choice have been provided.

The study Mémin (2014) and this paper follow another approach, introduced by Brzeźniak et al. (1991) and Mikulevicius and Rozovskii (2004). The aim is to bring up naturally a physically based uncertainty quantification and a sub-grid-tensor model without strong assumptions. The velocity is assumed to be random and partially decorrelated in time. From the Lagrangian point of view, it defines a general semimartingale flow. Using stochastic calculus and classical fluid dynamics principles, one can prove a stochastic representation of the so-called Reynolds transport theorem. It describes the time-space evolution of a scalar transported by this semimartingale flow. Using energy and mass conservation, it leads to time-space evolution of respectively the temperature and the density. Then, the transport theorem applied to momentum and the second Newton Law lead to a stochastic version of the Navier-Stokes equations. The corresponding scalar and velocity evolution laws involve an inhomogeneous time-dependent anisotropic diffusive sub-grid-tensor and additive and multiplicative noise.

For some industrial applications, the resolution of a system of partial differential equations may be too time consuming. A solution consists in deriving a model of reduced dimension, like in the case of the Proper Orthogonal Decomposition (POD) Holmes et al. (1998). Within this model, the velocity at a fixed time, t, is assumed to live in a small dimensional subspace of functions of space. The basis of this subspace, so-called spatial modes, is determined by a Principal Component Analysis (PCA) on a sequence of velocity snapshots. The coefficient of the velocity in the reduced basis are called temporal modes. Then, the partial differential equation of interest is projected on the function of this basis. It leads to a finite set of coupled ordinary differential equations which

describe the time evolution of the temporal modes. One problem for non-linear models such as Navier-Stokes equations is that keeping only a small number of modes destabilizes the system. In order to overcome this, some authors empirically add a diffusive term, parametrized by an eddy viscosity model to the reduced model. Several modeling of this eddy viscosity have been proposed. For instance, Aubry et al. (1988) proposed a constant coefficient, whereas Rempfer and Fasel (1994) introduced the modal model with one eddy viscosity per mode. Recently, Östh et al. (2014) and Protas et al. (2015) proposed an eddy viscosity model that depends on the instantaneous energy of the temporal modes. Other authors (Carlberg et al. (2011)) perform non-linear Galerkin methods, with the same spatial modes. It leads to another form of the reduced model, that will not be investigated in this paper.
In our approach, the unresolved temporal modes are assumed to be random and decorrelated in time whereas the resolved ones are deterministic. Thus, according to our stochastic Navier-Stokes model, an explicit sub-grid tensor appears both in the PDE and in the associate reduced model. The parameters of this sub-grid tensor can then be easily estimated on the residual velocity, through a statistical estimator. By residual velocity, we mean the part of the velocity snapshots which is not represented by the PCA. As will be demonstrated here, this sub-grid tensor successfully stabilizes the reduced system.
The paper is organized as follows. The first section presents the stochastic fluid dynamics model, on which we rely. The second section is a reminder of the classical POD approach. The third one presents our POD based reduced model under uncertainty. The fourth section presents some numerical results and comparisons. Finally, the last section concludes and provides perspectives.

6.2 The proposed stochastic model

In this work, an Eulerian stochastic description of the velocity and tracer evolution is used, as proposed in Mémin (2014). Unlike classical stochastic methods, a random part, encoding an uncertainty on the velocity expression, is added to the Lagrangian velocity before any model derivation. Thanks to this decomposition, a stochastic representation of the so-called Reynolds transport theorem, cornerstone of the deterministic fluid dynamic theory, can be derived. Thus, assuming a dynamical balance similarly to the second Newton law, a stochastic Navier-Stokes expression can be derived. It should be noticed that the equations, described below, are derived from fundamental physical laws only.
The time differentiation of a trajectory $\boldsymbol{X}_t$ of a particle is noted:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\mathrm{d}\boldsymbol{B}_t, \tag{6.2.1}$$

where $\boldsymbol{\sigma}(.,t)$ is an Hilbert-Schmidt operator on $(L^2(\mathbb{R}^d))^d$ defined by its kernel $\breve{\boldsymbol{\sigma}}(.,.,t)$: $\forall \boldsymbol{f} \in (L^2(\mathbb{R}^d))^d$, $\boldsymbol{\sigma}(.,t)\boldsymbol{f} \triangleq \int_\Omega \breve{\boldsymbol{\sigma}}(.,\boldsymbol{y},t)\boldsymbol{f}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$ and $t \mapsto \boldsymbol{B}(t)$ is a cylindrical I_d-Wiener process (see Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007) for more information on infinite dimensional Wiener process and cylindrical I_d-Wiener process). Then, $(\boldsymbol{x},t) \mapsto \boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t$ is a centered Gaussian process with the following covariance:

$$\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d, \mathbb{E}\left((\boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t)\left(\boldsymbol{\sigma}(\boldsymbol{y},t')\mathrm{d}\boldsymbol{B}_{t'}\right)^T\right) \triangleq \boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})\delta(t-t')\mathrm{d}t,$$

where:

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}t &= \int_\Omega \breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{z})\breve{\boldsymbol{\sigma}}^T(\boldsymbol{y},\boldsymbol{z})\mathrm{d}\boldsymbol{z}\mathrm{d}t \triangleq \boldsymbol{\sigma}(\boldsymbol{x})\boldsymbol{\sigma}(\boldsymbol{y})^T\mathrm{d}t \\ &= \mathrm{d}\left\langle \int_0^t \boldsymbol{\sigma}(\boldsymbol{x},t')\mathrm{d}\boldsymbol{B}_{t'}, \left(\int_0^t \boldsymbol{\sigma}(\boldsymbol{y},t'')\mathrm{d}\boldsymbol{B}_{t''}\right)^T\right\rangle.\end{aligned}$$

The notation $< f, g >$ is the quadratic cross-variation of f and g, used in stochastic calculus, and its expression is recall in Appendix 6.A. The term $\int_0^t \boldsymbol{w}\mathrm{d}t'$ represents the large-scale part of the flow whereas $\int_0^t \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}$ represents the small-scale part. The real physical small-scale flows are differentiable w.r.t. (with respect to) time. But, the time sampling used for large-scale modeling

or observation is often larger than the smaller physical time scale of the real velocity. Thus, at this large scale time sampling, the smallest scales of the flow are non differentiable almost everywhere w.r.t. time.
The semimartingale Lagrangian formulation (6.2.1) together with stochastic calculus theory allows us differentiating and integrating random physical quantities. Some basic notions of stochastic calculus, concerning finite variation processes, martingales and semimartingales, are provided in Appendix 6.A.

6.2.1 Stochastic representation of the Reynolds-transport theorem

Thanks to the previous decomposition, it is possible to derive a stochastic representation of the so-called Reynolds transport theorem. Unlike Mémin (2014), we will not assume that, for each $\boldsymbol{x}$,$\boldsymbol{y}$ and t, the matrix $\breve{\boldsymbol{\sigma}}(\boldsymbol{x},\boldsymbol{y},t)$ is symmetric. Furthermore, the time differentiable part of the flow, $\boldsymbol{w}(\boldsymbol{x},.)$, will not be assumed to be deterministic anymore, but rather, to be a continuous semimartingale. Nevertheless, if we exactly follow the same procedure, the very same stochastic transport theorem can be derived.

Theorem 6.2.1 *Stochastic Reynolds transport theorem*

Noting $\boldsymbol{\phi}$ the stochastic flow defined by:

$$\forall \boldsymbol{x} \in \Omega, t \in \mathbb{R}^+, \boldsymbol{\phi}(\boldsymbol{x},t) = \boldsymbol{x} + \int_0^t \boldsymbol{w}(\boldsymbol{\phi}(\boldsymbol{x},t'),t')\mathrm{d}t' + \int_0^t \boldsymbol{\sigma}(\boldsymbol{\phi}(\boldsymbol{x},t'),t')\mathrm{d}\boldsymbol{B}_{t'},$$

and denoting $V(t) = \boldsymbol{\phi}(V(0),t)$ a material volume transported by the stochastic flow, we have:

$$\mathrm{d}\int_{V(t)} q(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = \int_{V(t)} \left(\mathrm{d}_t q + \boldsymbol{\nabla}\cdot\left(q\mathrm{d}\boldsymbol{X}_t + q\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}q)^T\mathrm{d}t\right)\right)\mathrm{d}\boldsymbol{x}. \tag{6.2.2}$$

The mathematical equivalence between formulation (6.2.2) and the stochastic Reynolds transport theorem of Mémin (2014) is proven in Appendix 6.B. If q is a passive tracer, transported by the stochastic flow, $\mathrm{d}\int_{V(t)} q(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = 0$, and:

$$q(\boldsymbol{x},t) - q(\boldsymbol{x},0) = -\int_0^t \boldsymbol{\nabla}\cdot\left(q\boldsymbol{w} + q\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{a}q)^T\right)\mathrm{d}t' - \int_0^t \boldsymbol{\nabla}\cdot(q\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{B}_{t'}. \tag{6.2.3}$$

This equation is the unique decomposition of the continuous semimartingale q, into a finite variation process (the integral in $\mathrm{d}t$) and a local martingale (the integral in dB_t) Kunita (1997). Physically, the finite variation process varies slowly and is responsible of the large time-scale variation of q, whereas $\boldsymbol{\nabla}\cdot(q\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{B}_{t'}$ is decorrelated in time and null in average. From this point of view, the two components live in two different spaces, and hence the semimartingale decomposition is unique. If we make the hypothesis of a constant density ρ, the last equation applied to $q = \rho$ and the uniqueness of the decomposition leads to:

$$0 = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}, \tag{6.2.4}$$

$$0 = \boldsymbol{\nabla}\cdot\left(\boldsymbol{w} + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\right) = \boldsymbol{\nabla}\cdot\left(\boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\right). \tag{6.2.5}$$

Usually, the evolution of an intensive property, q, can be computed from equation (6.2.3), through the knowledge of the small-scale velocity characteristic, $\boldsymbol{\sigma}$, and of the large-scale drift, $\boldsymbol{w}$. Indeed, the evolution of all intensive property statistical moments can be formalized through equation (6.2.3). For instance, the equation of the conditional expectation of the scalar, given the velocity $\boldsymbol{w}$ for all time, $\bar{q} \triangleq \mathbb{E}(q|\boldsymbol{w})$, is:

$$\frac{\partial\bar{q}}{\partial t} + \boldsymbol{\nabla}\cdot(\bar{q}\boldsymbol{w}^*) = \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\bar{q}\right) \text{ where } \boldsymbol{w}^* = \boldsymbol{w} + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T. \tag{6.2.6}$$

This is a classical advection-diffusion equation. Indeed, since a is symmetric positive-semidefinite, $\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\bar{q}\right)$ leads only to diffusion. The expectation, $\bar{q}$, is advected by an effective drift, $\boldsymbol{w}^*$, and undergoes a diffusion through the tensor $\frac{1}{2}\boldsymbol{a}$. In the case of a constant density, $\boldsymbol{w}^*$ is naturally divergence-free (see equation (6.2.5)). For large-scale tracers, this advection-diffusion equation, derived from physical laws, has the same form as the widely used empirical advection-diffusion equation setup through an eddy diffusivity assumption (Vallis (2006)). However, unlike most of these classical models, the sub-grid diffusion we got is time-dependent, anisotropic and inhomogeneous.

6.2.2 Stochastic Navier-Stokes model

Similarly to the Newton second law, a dynamical balance between the temporal differentiation of the stochastic momentum, $\rho \mathrm{d}\boldsymbol{X}_t$, and general stochastic forces action is assumed. This leads, applying (6.2.2) to $\rho \mathrm{d}\boldsymbol{X}_t$ and ρ, to the following stochastic Navier-Stokes representation.

Theorem 6.2.2 *Stochastic Navier-Stokes representation*
If $\boldsymbol{w}$ is a finite variation process and f the integral of the pressure p along time can be decomposed as a general continuous semimartingale $\int_0^t (p'\mathrm{d}t + \mathrm{d}_t p_\sigma)$, then

$$\rho\left(\frac{\partial \boldsymbol{w}}{\partial t} + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\boldsymbol{w} + \boldsymbol{f}\times\boldsymbol{w}\right) = \tau(\boldsymbol{w}) + \rho\boldsymbol{g} - \boldsymbol{\nabla}p + f_V(\boldsymbol{w}), \tag{6.2.7}$$

$$\rho\left((\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla})\boldsymbol{w} + \boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = -\boldsymbol{\nabla}\mathrm{d}_t p_\sigma + f_V(\boldsymbol{\sigma})\mathrm{d}\boldsymbol{B}_t, \tag{6.2.8}$$

where

$$\left\{\begin{array}{rcl} f_V(h) & = & \mu\left(\boldsymbol{\nabla}^2 h + \frac{1}{3}\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot h\right)\right), \\ \forall k, \tau_k(\boldsymbol{w}) & = & \frac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\rho\boldsymbol{a}\boldsymbol{w}_k))^T - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\rho\boldsymbol{a}))^T\boldsymbol{w}_k - 2*\rho\left((\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)\boldsymbol{w}_k\right). \end{array}\right.$$

As a consequence, if the large-scale component, $\boldsymbol{w}$, is a finite variation process (i.e. if it is time differentiable) and if the density ρ is deterministic, then $\boldsymbol{w}$ is deterministic, knowing the initial conditions. It can be noted again that the kernel $\breve{\boldsymbol{\sigma}}$ is not assumed pointwise symmetric. The equivalence between formulation (6.2.7) and the stochastic Navier-Stokes model of Mémin (2014) is proven in Appendix 6.B.
Expression (6.2.7) can be seen as a generalization of several classical turbulence models. For instance, if the small-scale infinitesimal displacement $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ is isotropic and divergence free, and if the density is constant, the sub-grid tensor simplifies to $\tau(\boldsymbol{w}) = \rho\frac{\boldsymbol{a}}{2}\Delta\boldsymbol{w}$. We retrieve the simplest expression of the Boussinesq assumption, with a constant eddy viscosity given by $\frac{\boldsymbol{a}}{2}$. Generally speaking, we may wonder whether the sub-grid tensor, τ, is dissipative, like in a theoretical 3D direct energy cascade (Vallis (2006), Kupiainen (2000)). If ρ is assumed to be constant, and if $\boldsymbol{a}$ or $\boldsymbol{w}$ and their derivatives are assumed to be null on the border of Ω, then τ is dissipative. The proof is provided in Appendix 6.C.
The knowledge of small-scale physical flow realizations allows estimating $\boldsymbol{\sigma}$ and $\boldsymbol{a}$:

$$\boldsymbol{a}\mathrm{d}t = \mathbb{E}((\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T) \text{ and } (\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}^T\mathrm{d}t = \mathbb{E}((\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T). \tag{6.2.9}$$

Thus, the value of $\boldsymbol{a}(\boldsymbol{x},t)$ can be used in a large-scale simulation ruled by equation (6.2.7). The tensor $\boldsymbol{a}$ and functions of $\boldsymbol{\sigma}$ can also be estimated from a single realization thanks to stochastic calculus, as explained later, or by assuming local time or space ergodicity as in Harouna and Mémin (Preprint 2014). Another interesting way of using (6.2.9) is through Monte-Carlo small-scale simulations, such as particle filtering, where particles correspond to several probable values of the small-scale velocity. A third alternative to estimate $\boldsymbol{a}$ and $\boldsymbol{\sigma}$, without using (6.2.9), consists in relying directly on known statistical properties of small-scale measurements. Some works on this subject are currently ongoing. In all these methods, the estimation of the tensor $\boldsymbol{a}$ corresponds to a solution of a closure problem. Knowing the value of $\boldsymbol{a}$ should lead to a simulation of the drift, through (6.2.7), or of a tracer transport, through the stochastic transport theorem (Mémin (2014)), in which the small-scale actions are taken into account in a statistical way.
It is also possible to follow a dual strategy with a downscaling approach, like, for instance, mixing diagnostics, which is, at the moment, an important issue in Meteorology and Oceanography

(Nakamura (2001) and Mezić et al. (2010)). The evolution equation of both the averaged stochastic transport theorem (6.2.6) and the average stochastic Navier-Stokes model (6.2.7), applied to observed large-scale geophysical data can give information on the tensors $\boldsymbol{a}$ and $\boldsymbol{\sigma}$. Indeed, analysing how a tracer is advected and diffused may help computing these two tensors. Related information, such as the local small-scale energy, the anisotropy created by the matrix $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{x},t)$, or the local divergence, which is linked to $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}$, can be inferred. These information could teach us, for instance, the likelihood of locally strong velocity or tracer gradient and the principal directions of mixing created by the main variance directions.
Theorems 6.2.1 and 6.2.2 provide the foundations of a physically relevant stochastic Fluid Dynamics framework. In this paper, we will rely on them for a reduced model application.

6.3 Classical model reduction using POD

Dimensional reduction techniques are methods allowing simplification of Partial Differential Equations (PDE), using dedicated basis specified from observed data. The Proper Orthogonal Decomposition (POD) is one of these methods, and below are recalled its main principles.
Here, we consider an observed multivariate field such as a velocity $\boldsymbol{u}(\boldsymbol{x},t)$ depending on space $x \in \Omega$ and time $t \in [0,T]$. The goal consists in looking for a subspace of reduced dimension where $\boldsymbol{v}(.,t)$ is likely to live for all t. We thus seek a finite orthonormal set of function of space, which spans this subspace. These functions $(\boldsymbol{\phi}_i(\boldsymbol{x}))_{1\leqslant i\leqslant N}$ are called spatial modes or *topos* and are computed from a Karunen Loeve decomposition on a series of available velocity snapshots. In other words, a spectral analysis is done on the space (or time) autocorrelation tensor of observed data:

$$\bar{\boldsymbol{v}}(\boldsymbol{x}) \triangleq \frac{1}{T}\int_0^T \boldsymbol{v}(\boldsymbol{x},t)\mathrm{d}t, \tag{6.3.1}$$

$$\boldsymbol{cov}(\boldsymbol{x}_1,\boldsymbol{x}_2) \triangleq \frac{1}{T}\int_0^T (\boldsymbol{v}(\boldsymbol{x}_1,t)-\bar{\boldsymbol{v}}(\boldsymbol{x}_1))(\boldsymbol{v}(\boldsymbol{x}_2,t)-\bar{\boldsymbol{v}}(\boldsymbol{x}_2))^T\mathrm{d}t, \tag{6.3.2}$$

$$\int_\Omega \boldsymbol{cov}(\boldsymbol{x}_1,\boldsymbol{x}_2)\boldsymbol{\phi}_i(\boldsymbol{x}_1)\mathrm{d}\boldsymbol{x}_1 = \lambda_i\boldsymbol{\phi}_i(\boldsymbol{x}_2) \text{ with } \int_\Omega \boldsymbol{\phi}_i(\boldsymbol{x})\cdot\boldsymbol{\phi}_j(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \delta_{i,j}. \tag{6.3.3}$$

The *topos* are sorted such that $\lambda_1 > ... > \lambda_N$, where N is the number of observed snapshots (if the number of points of the spatial grid is larger than N). It leads to the decomposition:

$$\forall(\boldsymbol{x},t)\in\mathbb{R}^d\times\mathbb{R},\ \boldsymbol{v}(\boldsymbol{x},t)\approx\bar{\boldsymbol{v}}(\boldsymbol{x})+\sum_{i=1}^N b_i(t)\boldsymbol{\phi}_i(\boldsymbol{x}). \tag{6.3.4}$$

The values $(b_i(t))_{1\leqslant i\leqslant N}$ are called temporal modes or *chronos* and satisfy:

$$\forall i,j,\ \frac{1}{T}\int_0^T b_i(t)b_j(t)\mathrm{d}t = \lambda_i\delta_{i,j}. \tag{6.3.5}$$

In the following, $\bar{\boldsymbol{v}}$ will be denoted $\boldsymbol{\phi}_0$ and $b_0 \triangleq 1$. Then, since only the first temporal modes concentrate the most significant part of the energy, a second truncation approximation is usually performed:

$$\forall(\boldsymbol{x},t)\in\mathbb{R}^d\times\mathbb{R},\ \boldsymbol{v}(\boldsymbol{x},t)\approx\sum_{i=0}^n b_i(t)\boldsymbol{\phi}_i(\boldsymbol{x}) \text{ with } n\ll N. \tag{6.3.6}$$

A Galerkin projection enables us to look for an approximate solution of a PDE. The approximate solution at time t, $\boldsymbol{v}(.,t)$, defined in (6.3.6), is assumed to live in a finite-dimensional sub-space, spanned by $(\boldsymbol{\phi}_0,...,\boldsymbol{\phi}_n)$, instead of an infinite-dimensional one. The time-space evolution equation of $\boldsymbol{v}$ (a PDE) is then expressed as the time evolution equations (a finite set of coupled ODEs) of *chronos*. In fluid dynamics, PDE system describing the velocity evolution, such as the Navier-Stokes equations, have the general following abstract form:

$$\frac{\partial\boldsymbol{v}}{\partial t} = I + L(\boldsymbol{v}) + C(\boldsymbol{v},\boldsymbol{v}), \tag{6.3.7}$$

where L and C are respectively linear and bilinear differential operators. The first term, I, includes pressure and gravity. The second one, L, involves molecular viscosity and possibly Coriolis force. The last one, C, encodes the non-linear advection term. Projecting this PDE on each *topos* leads to:

$$\forall i \leqslant n,\ \frac{db_i}{dt} = \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot I\right)}_{\triangleq i_i} + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot L(\boldsymbol{\phi}_p)\right)}_{\triangleq l_{p,i}} b_p + \sum_{p,q=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot C(\boldsymbol{\phi}_p, \boldsymbol{\phi}_q)\right)}_{\triangleq c_{p,q,i}} b_p b_q.$$

Because of the non-linearity, the temporal modes strongly interact with each others. In particular, even though the original model (with $n = N$) is computationally stable for moderate Reynolds number, the reduced one is generally not so. This particularity of the Navier-Stokes equation is not restricted to the POD framework. Simulating a large-scale flow, considering only the largest Fourier modes, leads also to strong instabilities and numerical explosions. A rough truncation cannot be considered without introducing a dissipative term whose role is to drain the energy brought by the larger modes beyond the truncation and thus to avoid an energy accumulation. Eddy viscosity models, which consist in enforcing the fluid viscosity, are often used for that purpose. This principle, which dates back to Boussinesq (Kraichnan (1987)), is often used in large-scale simulation as well as in the context of POD (Aubry et al. (1988), Rempfer and Fasel (1994), Östh et al. (2014), Protas et al. (2015)). In practice, these methods introduce empirically an additional damping term to the Navier-Stokes equation. This leads to a modified linear term in (6.3.7). Unfortunately, since this term is built from an empirical thermodynamical analogy, its precise form is difficult to justify and its parametrization has to be tuned for each simulation to get optimal results. The method proposed in the next section allows us to tackle these drawbacks.

6.4 Stochastic POD

To overcome the difficulties developed previously, we suggest to use our stochastic Navier-Stokes model instead of the classical Navier-Stokes equations. Let us outline that both systems address the same physics. They both rely on mass and momentum conservation and differ only in how they are taking into account small-scale missing information.

6.4.1 Model

The reduced dynamic system we propose is based on the stochastic Navier-Stokes model developed in (6.2.2), assuming that the density ρ is constant and the smooth part of the flow, $\boldsymbol{w}$, is of bounded variations. To tackle the problem of modes interactions, Mémin (2014) proposed to decompose $\boldsymbol{v}$ as follows : $\boldsymbol{v}\mathrm{d}t = \boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ with $\boldsymbol{w} = \sum_{i=0}^{n} b_i \boldsymbol{\phi}_i$ (projection on the truncated subspace) and $\sum_{i=n+1}^{N} b_i \boldsymbol{\phi}_i \mathrm{d}t$ a realization of $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ (projection on the complementary "small-scale" subspace). Since $\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0$, for all i, $\boldsymbol{\nabla}\cdot\boldsymbol{\phi}_i = 0$ and, then, $\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0$. The drift, $\boldsymbol{w}$, follows the finite variation part of the stochastic Navier-Stokes equation (6.2.7) in the incompressible case. Projecting on the divergence-free functions space, we have:

$$\frac{\partial \boldsymbol{w}}{\partial t} + \boldsymbol{\mathcal{P}}\left((\boldsymbol{w}\cdot\boldsymbol{\nabla})\boldsymbol{w}\right) = \boldsymbol{\mathcal{P}}\left(\frac{1}{\rho}\tau(\boldsymbol{w})\right) + \boldsymbol{g} + \nu\Delta\boldsymbol{w}, \tag{6.4.1}$$

where

$$\frac{1}{\rho}\tau_k(\boldsymbol{w}) = \frac{1}{2}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{w}_k))^T - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\boldsymbol{w}_k\right), \tag{6.4.2}$$

$$\boldsymbol{\mathcal{P}} \triangleq \mathbb{I}_d - \Delta^{-1}\boldsymbol{\nabla}\boldsymbol{\nabla}^T. \tag{6.4.3}$$

In Fourier space, the projector on divergence-free functions space reads $\hat{\boldsymbol{\mathcal{P}}} = \mathbb{I}_d - \frac{\boldsymbol{k}\boldsymbol{k}^T}{\|\boldsymbol{k}\|_2^2}$. Equation (6.4.1) can be rewritten as:

$$\frac{\partial \boldsymbol{w}}{\partial t} = I + L(\boldsymbol{w}) + C(\boldsymbol{w}, \boldsymbol{w}) + F(a, \boldsymbol{w}), \tag{6.4.4}$$

where F is a bilinear differential operator. Projecting this equation along $\boldsymbol{\phi}_i$ for each $i \in [\![1, n]\!]$, gives the evolution equation of $b \triangleq (b_i)_{1 \leqslant i \leqslant n}$.

$$\forall i \in [\![1, n]\!], \frac{db_i}{\mathrm{d}t} = i_i + \left(\boldsymbol{l}_{\cdot i} + \breve{\boldsymbol{f}}(\boldsymbol{a})_{\cdot i}\right)^T \boldsymbol{b} + \boldsymbol{b}^t \boldsymbol{c}_{\cdot\cdot i} \boldsymbol{b}, \tag{6.4.5}$$

$$\text{with } \breve{f}(\boldsymbol{a})_{j,i} \triangleq \int_\Omega \boldsymbol{\phi}_i \cdot F(\boldsymbol{a}, \boldsymbol{\phi}_j), \tag{6.4.6}$$

where the coefficients $(i_i)_{1 \leqslant i \leqslant n}, (\breve{f}(a)_{j,i})_{1 \leqslant i,j \leqslant n}, (l_{j,i})_{1 \leqslant i,j \leqslant n}$ and $(c_{k,j,i})_{1 \leqslant i,j,k \leqslant n}$ are computed through the integration over the whole space of the terms of (6.4.1). Those dynamical coefficients depend on both the resolved *topos* and the unresolved velocity variance tensor, $\boldsymbol{a}$. This system includes a natural small-scale dissipation mechanism, through the tensor τ. To fully specify this system, we need to estimate the quadratic variance tensor $\boldsymbol{a}$. This important issue is developed in subsection 6.4.3. But first we will elaborate further on the choice of a characteristic time step related to the truncation operated.

6.4.2 Choice of the time step

For several applications, the simulation of the most energetic large-scale component of the solution is sufficient. However, this simulation needs to be fast, implying a low complexity evolution model and a large time step. The structure of our stochastic model enables to reach both goals.
Indeed, as long as the resolved modes, which represent $\boldsymbol{w}$, are differentiable w.r.t. time, our stochastic reduced model is valid. Thus, the time step can be chosen as large as desired, as long as these modes remain smooth. The Shannon-Nyquist sampling theorem provides a natural upper bound to fix this time step. This theorem states that a function can be sampled, without loss of information, if the sampling frequency is twice as large as the largest frequency of the original function. Otherwise, the sampled function undergoes an aliasing artifact characterized by a back folding of the Fourier spectrum. If the resolved POD modes and their evolution equations are not affected by aliasing phenomena, the required smoothness is assumed to be reached. Since the evolution equations are quadratic, a sufficient condition for the necessary smoothness is:

$$\frac{1}{\Delta t} \geqslant 4\ max_{i \leqslant n} \left(f_{max}\left(b_i\right)\right), \tag{6.4.7}$$

where $f_{max}\left(b_i\right)$ is the maximum frequency of the i-th temporal mode.
Of course, aliasing will occur in the unresolved temporal modes, associated with smaller time scales. However, our stochastic model is derived from a decorrelation assumption of the small-scale unresolved part of the velocity. A strong subsampling of these components strengthens the decorrelation property of these modes.

6.4.3 Estimation of the uncertainty variance tensor

After having estimated the *topos* and fixed the time step, we need to estimate the uncertainty variance tensor $\boldsymbol{a}$. This estimation will enable us to get a full expression of the dynamical coefficients of the *chronos* evolution equations (6.4.5). To that end, additional modeling assumptions must be imposed. The first natural hypothesis consists in assuming an uncertainty field that is stationnary in time – and spatially non homogenous. In this stationary case, the uncertainty variance tensor is constant in time.

The uncertainty variance tensor is constant in time

This case corresponds to the assumption used in Mémin (2014). To understand the consequence of this hypothesis, we recall that $\frac{\boldsymbol{a}}{\Delta t}$ is the variance of the residual velocity $\boldsymbol{v} - \boldsymbol{w}$. The process

is decorrelated in time and Gaussian. The snapshots are hence independent. Therefore, the $(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{v}-\boldsymbol{w})^T(t_i)$ are independent and identically distributed. So, the expectation, $\frac{\boldsymbol{a}}{\Delta t}$, can be computed by averaging the snapshots $(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{v}-\boldsymbol{w})^T(t_i)$. In other words, the process is ergodic.

$$\frac{\boldsymbol{a}(\boldsymbol{x})}{\Delta t} = \lim_{N\to+\infty}\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{v}-\boldsymbol{w})(x,t_i)((\boldsymbol{v}-\boldsymbol{w})(x,t_i))^T, \tag{6.4.8}$$

where N is the number of snapshots after time sub-sampling. The convergence is almost sure by the strong law of large numbers but, here, only the convergence in probability is used. One can notice that this relation is straightforward when using stochastic calculus, as explained hereafter. Thanks to the expression above, one can see several advantages of such an assumption. First of all, the construction of $\boldsymbol{a}$ is straightforward and easy to compute. Secondly, $\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t$ itself is a $\boldsymbol{a}$-Wiener process, since $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})$ is a trace class operator constant in time (see Prévôt and Röckner (2007) and Da Prato and Zabczyk (1992) for more details on Q-Wiener processes). The spectrum of this operator, which is central in the model reduction process, enables us to use a diagonalized version of $\boldsymbol{\sigma}$ and $\boldsymbol{a}$ as in (Mémin (2014)):

$$\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y}) = \sum_{k=n+1}^{\infty}\lambda_k\Delta t\ \boldsymbol{\phi}_k(\boldsymbol{x})\boldsymbol{\phi}_k(\boldsymbol{y})^T \approx \sum_{k=n+1}^{N}\lambda_k\Delta t\ \boldsymbol{\phi}_k(\boldsymbol{x})\boldsymbol{\phi}_k(\boldsymbol{y})^T, \tag{6.4.9}$$

$$\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t = \sum_{k=n+1}^{\infty}\sqrt{\lambda_k\Delta t}\ \boldsymbol{\phi}_k(\boldsymbol{x})d\beta_t^{(k)} \approx \sum_{k=n+1}^{N}\sqrt{\lambda_k\Delta t}\ \boldsymbol{\phi}_k(\boldsymbol{x})d\beta_t^{(k)}, \tag{6.4.10}$$

where the $(\beta^{(k)})_{k>n}$ are independent standard one-dimensional Brownian motions. As a result, it is very easy in this context to generate realizations of the small-scale uncorrelated component. The knowledge of the leading eigenfunctions of the POD complementary space allows us to access directly to the spectral representation of the diffusion tensor.

The uncertainty variance tensor is time varying

Assuming a constant value for $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{x})$ means that the turbulence is not intermittent. In the context of POD, it would mean that all the unresolved modes have a constant variance. It is a good first approximation. But, one may wonder whether it is possible to do better.
If $\boldsymbol{a}$ does depend on time, the estimation is more involved. Since only one realization of the small-scale velocity is available, some time-ergodicity hypothesis would be necessary, at least locally, to use (6.2.9) as in (Harouna and Mémin (Preprint 2014)). Otherwise, parametric and non parametric estimation of $\boldsymbol{a}(\boldsymbol{X}_t,t)\mathrm{d}t = \mathrm{d}<\boldsymbol{X},\boldsymbol{X}>_t$ are studied in the literature (Rao (1999), Genon-Catalot et al. (1992), Florens-Zmirou (1993), Genon-Catalot and Jacod (1993), Hofmann et al. (1999) and Comte et al. (2007)). Parametric ones use for instance maximum likelihood estimation. Indeed, denoting θ the parameters, the Girsanov theorem (Oksendal (1998)) leads, as explained in Rao (1999), to the following log-likelihood:

$$\begin{aligned} l(\boldsymbol{X}_t-\boldsymbol{X}_0|\boldsymbol{\theta}) &= \int_0^t \boldsymbol{w}(t',\boldsymbol{X}_{t'})^T\left(\boldsymbol{a}^{op}(t',\boldsymbol{X}_{t'}|\boldsymbol{\theta})\right)^{-1}\mathrm{d}\boldsymbol{X}_{t'} \\ &\quad -\frac{1}{2}\int_0^t \boldsymbol{w}(t',\boldsymbol{X}_{t'})^T\left(\boldsymbol{a}^{op}(t',\boldsymbol{X}_{t'}|\boldsymbol{\theta})\right)^{-1}\boldsymbol{w}(t',\boldsymbol{X}_{t'})\mathrm{d}t', \end{aligned}$$

where:

$$\boldsymbol{a}^{op}(t,\boldsymbol{x}|\boldsymbol{\theta})(\boldsymbol{f}) \triangleq \int_\Omega \boldsymbol{a}(t,\boldsymbol{x},\boldsymbol{y}|\boldsymbol{\theta})\boldsymbol{f}(\boldsymbol{y})\mathrm{d}\boldsymbol{y}.$$

It is a very powerful tool because it can be used on Lagrangian data. However, the knowledge of the inverse of the infinite dimensional operator $\boldsymbol{a}^{op}$ is required. Moreover, here, we look rather for a non-parametric technique. These methods assume that $\boldsymbol{\sigma}$ is either constant in time or in space (Genon-Catalot et al. (1992), Florens-Zmirou (1993), Genon-Catalot and Jacod (1993), Hofmann et al. (1999) and Comte et al. (2007)). But, contrary to the main application domain of the literature

applications (finance), we have here access to an Eulerian realization of the stochastic flow. For all $\boldsymbol{x} \in \mathbb{R}^d$, it is then possible to build a spatially homogeneous local martingale $\tilde{\boldsymbol{X}}_t^{\boldsymbol{x}} \triangleq \int_0^t \boldsymbol{\sigma}(\boldsymbol{x},t)\mathrm{d}\boldsymbol{B}_t$. Its realization, $\int_0^t (\boldsymbol{v}(\boldsymbol{x},t)-\boldsymbol{w}(\boldsymbol{x},t))\mathrm{d}t$, after time sub-sampling, enables to estimate, for all functions h_k,

$$\begin{aligned}
\int h_k(t)\boldsymbol{a}(\boldsymbol{x},t)\mathrm{d}t &= \int h_k(t)\mathrm{d}\left\langle \tilde{\boldsymbol{X}}^{\boldsymbol{x}}, \left(\tilde{\boldsymbol{X}}^{\boldsymbol{x}}\right)^T \right\rangle_t, \\
&= \mathbb{P}-\lim_{\Delta t\to 0}\sum_{t_i=0}^{T} h^k(t_i)(\tilde{\boldsymbol{X}}_{t_{i+1}}^{\boldsymbol{x}} - \tilde{\boldsymbol{X}}_{t_i}^{\boldsymbol{x}})(\tilde{\boldsymbol{X}}_{t_{i+1}}^{\boldsymbol{x}} - \tilde{\boldsymbol{X}}_{t_i}^{\boldsymbol{x}})^T, \\
&\approx (\Delta t)^2 \sum_{t_i=0}^{T} h^k(t_i)(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{x},t_i)((\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{x},t_i))^T,
\end{aligned}$$

where $\mathbb{P}-\lim_{\Delta t\to 0}$ stands for the limit in probability as the time step, Δt, approaches 0. The functions h^k can be a orthonormal basis of $L^2([0,T])$ such as wavelets (Genon-Catalot et al. (1992)). In Genon-Catalot et al. (1992), it is shown that such estimators have good statistical properties: local asymptotic normality of the integrated square errors, together with the rate of convergence of its bias and variance. Therefore, the influence of the sub-grid tensor on $\frac{d}{\mathrm{d}t}b_i$ will be represented by time-dependent, linear coefficients as shown in (6.4.5).

The uncertainty variance tensor is in the span of the chronos

The *chronos* reduced basis $(b_i/\|b_i\|_{L^2([0,T])})_{1\leqslant i\leqslant n}$ provides a much better solution than a wavelet basis. As a matter of fact, this choice has three main advantages. First of all, since we are studying the time evolution of $(b_i)_{1\leqslant i\leqslant n}$, the slow time variations of $\boldsymbol{a}$, which are consistent with the time variations of $(b_i)_{0\leqslant i\leqslant n}$, are the information most needed. The number of wavelets needed to represent these time variations would be a priori much larger than $n+1$. Secondly, we do not need to reconstruct $\boldsymbol{a}$. Indeed, noting $\boldsymbol{z}_i(\boldsymbol{x}) = \frac{\overline{b_i(.)\boldsymbol{a}(\boldsymbol{x},.)}}{\lambda_i}$ and using the fact that $\boldsymbol{a} \to f(\boldsymbol{a})$ (defined by (6.4.6)) is linear,

$$\breve{f}(\boldsymbol{a}) = \breve{f}\left(\sum_{k=0}^{n} b_k \boldsymbol{z}_k\right) = \sum_{k=0}^{n} b_k \breve{f}\left(\boldsymbol{z}_k\right). \tag{6.4.11}$$

Thus, (6.4.5) becomes:

$$\forall i \in [\![1,n]\!], \frac{db_i}{\mathrm{d}t} = i_i + \boldsymbol{l}_{\cdot i}^T \boldsymbol{b} + \boldsymbol{b}^T\left(\boldsymbol{c}_{\cdot\cdot i} + \boldsymbol{f}_{\cdot\cdot i}\right)\boldsymbol{b}, \tag{6.4.12}$$

$$\text{with } f_{pqi} \triangleq \breve{f}_{qi}\left(\boldsymbol{z}_p\right). \tag{6.4.13}$$

If one chooses the basis $(b_i)_{0\leqslant i\leqslant n}$, only $(\boldsymbol{z}_i)_{0\leqslant i\leqslant n}$ is needed to compute f and, hence, to simulate the influence of $\boldsymbol{a}$ in the evolution of b_i. The calculation is thus more direct. One may notice that equations (6.4.12) define again a quadratic evolution system of *chronos*. Thus, the criterion for time step choice, developed in 6.4.2, remains the same. The third advantage of this basis is that f does not depend on time, unlike the term $\breve{f}(a)$ in equation (6.4.5). It is thus faster to compute than $\breve{f}(\boldsymbol{a}(t))$ at all time t, and requires a lower memory capacity to store it. But, above all and unlike any other basis, although the variance tensor is time dependent, the evolution system of $(b_i)_{1\leqslant i\leqslant n}$ remains autonomous. This is an unavoidable requirement for a forecasting task (Protas et al. (2015)).
To simplify the equation and to be more precise in what follows, we will remove the constant balance in (6.4.12). In permanent regime, since the system is stable, one can assume that $\overline{\frac{db_i}{\mathrm{d}t}} = 0$.

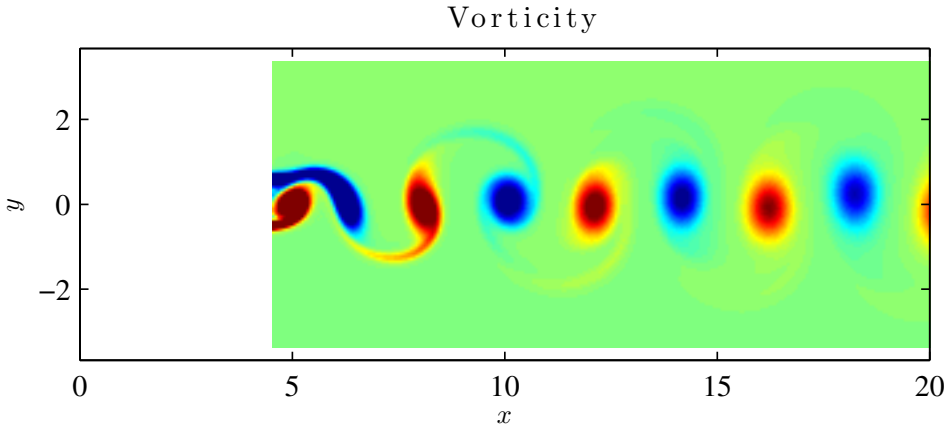


Figure 6.1: Vorticity along z of a wake behind a cylinder at Reynolds 300.

Thus, noting $\boldsymbol{b}_{1:n} = (b_i)_{1 \leqslant i \leqslant n}$, we get:

$$\forall i \in [\![1, n]\!], \frac{db_i}{\mathrm{d}t} = -\sum_{k=1}^{n} \lambda_k \left(c + f\right)_{k,k,i} + \boldsymbol{b}_{1:n}^T \left(\boldsymbol{c} + \boldsymbol{f}\right)_{1:n,1:n,i} \boldsymbol{b}_{1:n} + \left(\boldsymbol{l}_{1:n,i} + \left(\boldsymbol{c} + \boldsymbol{f}\right)_{0,1:n,i}^T + \left(\boldsymbol{c} + \boldsymbol{f}\right)_{1:n,0,i}\right)^T \boldsymbol{b}_{1:n}. \tag{6.4.14}$$

This model leads, as you will see later on, to an improved accuracy and stability of the system. We however lose the direct sampling capability of the previous simple ergodic assumption.

6.5 Numerical results

The different variations of the proposed approach have been assessed and compared numerically on numerical data of a wake behind a cylinder at Reynolds 300 and 3900 (Parnaudeau et al. (2008)).

6.5.1 Characteristics of the data

The fluid is incompressible: $\boldsymbol{\nabla} \cdot \boldsymbol{v} = 0$. At $x = 0$, there is a constant velocity $U = 1$ directed along $x > 0$. At $(x, y) = (5, 0)$, there is a motionless cylinder with an axis along the z axis. In permanent regime, it creates a Von Kármán vortex street behind the cylinder. A clockwise vortex is created at the bottom right of the cylinder, it breaks away from it and moves downstream. Then, a counter-clockwise vortex is created at the top right of the cylinder, breaks away from the first one and moves downstream, and so on. This periodic physical process makes the two first *chronos* almost sinusoidal.
Figures 6.1 and 6.2 show the z component of the vorticity $\boldsymbol{\nabla} \times \boldsymbol{v}$ on horizontal section of the fluid. In Figure 6.1, the cylinder is cropped. The vorticity is a measurement of the rotation of the fluid on this plane. A positive vorticity (in red) means a counter-clockwise rotation. A negative vorticity (in blue) means a clockwise rotation. At this point, one can see, in both Figures, a counter-clockwise is breaking away at the right of the cylinder and a clockwise one is enlarging at the bottom right.

In Figure 6.2, Kelvin-Helmholtz instabilities at the top right and bottom right of the cylinder can be observed just before the vortex creation zone. At the top and bottom of the cylinder, the velocity is close to the inflow velocity $U = 1$ along x whereas, at the right of the cylinder, close to it, the velocity is close to zero. Thus, there are two mixing layers at the boundaries, at the top right and bottom right of the cylinder. These Kelvin-Helmholtz instabilities as well as the Von Kármán vortex street creates a turbulent wake downstream of the cylinder.
At Reynolds 300, there are only few small-scales features. Most of the energy and most of the

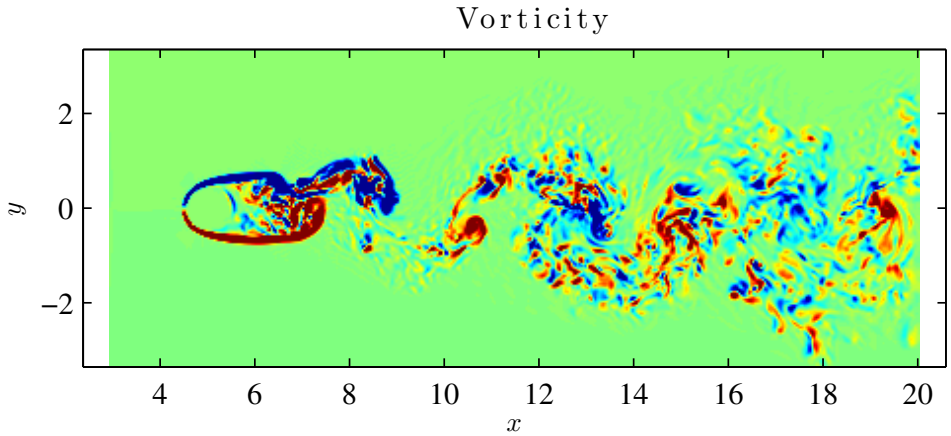


Figure 6.2: Vorticity along z of a wake behind a cylinder in the horizontal section $z = 0$, at Reynolds 3900.

dynamic is in large-scale structures. We use 80 vortex shedding. At Reynolds 3900, the turbulence is relatively important. Therefore, the spectrum support of the velocity is quite large, meaning that the velocity exists at several space and time scales. Indeed, one can see both small and large structures on Figure 6.2. Thus, in the context of POD, the *chronos* live at different time scales. Since the spectrum is more energetic for lower wave-number, the first *chronos*, i.e. the most energetic ones, have larger time-scale. Due to the quasi-periodic behavior of the flow along time, the *chronos* are closed to the Fourier modes but not exactly equal. This analysis of the *chronos* time scale is hence just a rough tendency. It explains nevertheless why our stochastic model, based on a separation between smooth and highly oscillating parts of the velocity is relevant.
Compared to the data of Parnaudeau et al. (2008), we slightly filtered and sub-sampled them spatially in order to reduce by two the number of gridded points by axes. The Gaussian filtering is used here only to reduce a potential spatial aliasing. To speed-up and facilitate the computations, we also removed part of the space where the vorticity is negligible: at $|y| > 3.5$ and $x < 3$, as seen in Figure 6.2. We use $N = 251$ time steps to observe 3 vortex shedding.

6.5.2 Reconstruction of *chronos*

To reconstruct the *chronos*, the reduced order dynamical system (6.4.14) is used. The modes mean energy, $(\lambda_i)_{1\leqslant i\leqslant n}$, and the *topos*, $(\boldsymbol{\phi}_i)_{0\leqslant i\leqslant n}$, are computed from the whole sequence of snapshots ($N = 3999$ for Reynolds 300 and $N = 251$ for Reynolds 3900). As for the initial condition, we used the referenced values of the *chronos* computed from the scalar product of initial velocity with the *topos*. Then, regarding the *chronos* spectra, an optimal time sub-sampling is chosen, as explained in subsection 6.4.2. Afterwards, using the *topos*, the residual velocity and possibly the *chronos*, the variance tensor, $\boldsymbol{a}$, or its decomposition is estimated. The coefficients of the reduced order dynamical system of *chronos* (see equation 6.4.14) are computed, using discrete derivation schemes and integration. Finally, the *chronos* are recomputed, integrating (6.4.14) with a 4-th order Runge-Kutta method, with $(b_i^{ref})(t = 0)_{1\leqslant i\leqslant n}$ as initial condition.

Figures 6.3 and 6.4 show examples of the reconstruction of the *chronos* for $n = 2$, at Reynolds 300, and $n = 10$, at Reynolds 3900, with the classical POD method (blue lines) and our method with a variance tensor defined as a linear combination of *chronos* (red lines). At Reynolds 300, the first two modes explain most of the energy. That is why we consider only $n = 2$. The reference $(b_i^{ref})_{1\leqslant i\leqslant n}$ (black dots) are superimposed for comparison purpose. It can be observed that our model follows the references quite well whereas the deterministic model blows up. The divergence occurs very quickly at Reynolds 3900. It may be pointed out that here both reduced models are parameter free. No constant had to be tuned to adapt any viscosity model.

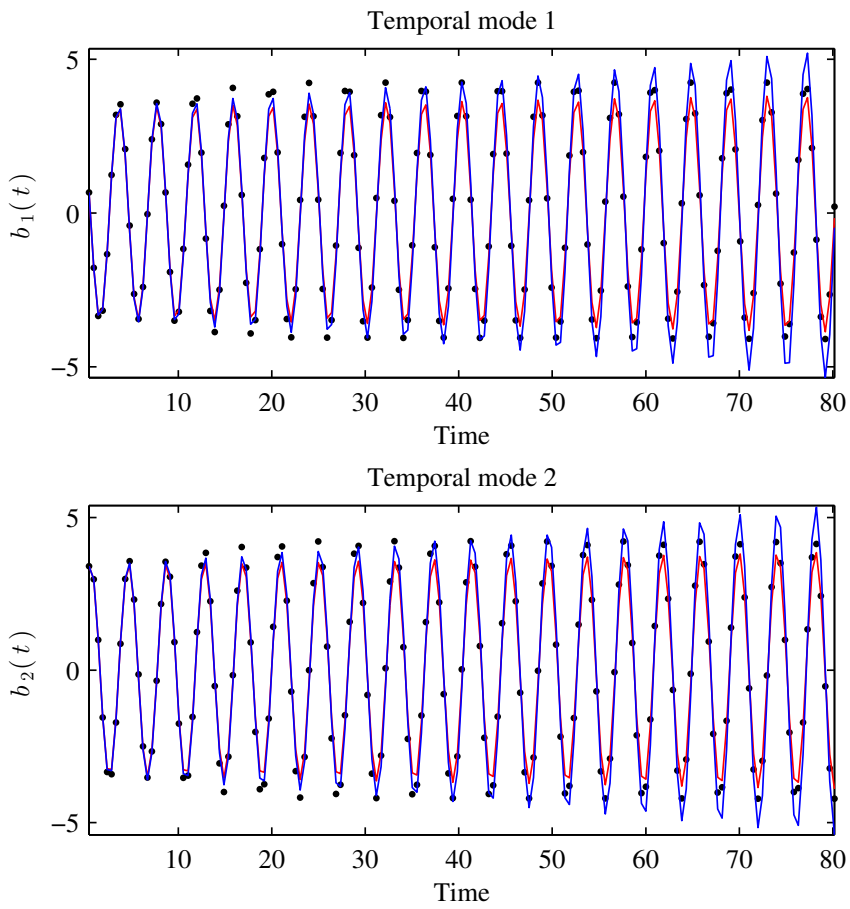


Figure 6.3: Reconstruction of the first two modes ($n = 2$), of a wake flow at Reynolds 300, with a variance tensor constant in time. The black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation, without any corrective coefficient. The initial condition, at $t = 0$, is common.

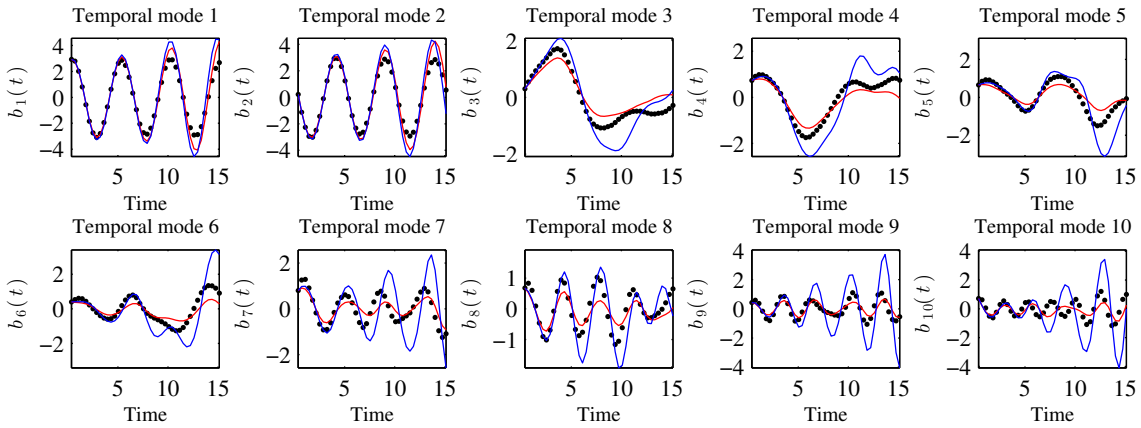


Figure 6.4: Reconstruction of the first ten modes ($n = 10$), of a wake flow at Reynolds 3900, with a variance tensor expressed as a linear function of the *chronos*. The black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation, without any corrective coefficient. The initial condition, at $t = 0$, is common.

Figures 6.5 and 6.6 show the error of the solution along time. The error is defined as follows:

$$\begin{aligned} err(t) &= T\frac{\left\|\boldsymbol{v}^{ref} - \boldsymbol{v}\right\|_{L^2(\Omega)}}{\|\boldsymbol{v}^{ref}\|_{L^2(\Omega\times[0,T])}}, \\ &= T\frac{\left\|\sum_{i=1}^{n}\left(b_i^{ref} - b_i\right)\boldsymbol{\phi}_i + \sum_{i=n+1}^{N} b_i^{ref}\boldsymbol{\phi}_i\right\|_{L^2(\Omega)}}{\left\|\sum_{i=0}^{n} b_i^{ref}\boldsymbol{\phi}_i\right\|_{L^2(\Omega\times[0,T])}}, \\ &= \left(\frac{\sum_{i=1}^{n}\left(b_i^{ref} - b_i\right)^2 + \sum_{i=n+1}^{N}\left(b_i^{ref}\right)^2}{\|\bar{\boldsymbol{w}}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}. \end{aligned}$$

Approximating the square of the real unresolved modes, $\left(\left(b_i^{ref}\right)^2\right)_{n+1\leqslant i\leqslant N}$, by their time average, $(\lambda_i)_{n+1\leqslant i\leqslant N}$, the error simplifies to:

$$err(t) \approx \left(\frac{\sum_{i=1}^{n}\left(b_i^{ref} - b_i\right)^2 + \sum_{i=n+1}^{N}\lambda_i}{\|\bar{\boldsymbol{w}}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}, \tag{6.5.1}$$

which is greater than the minimal error associated to the modal truncation:

$$err(t) \geqslant \left(\frac{\sum_{i=n+1}^{N}\lambda_i}{\|\bar{\boldsymbol{w}}\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}. \tag{6.5.2}$$

Equation (6.5.1) defines the criterion error plotted in Figures 6.5 and 6.6, whereas (6.5.2) constitutes a lower bound of this error.
Here, we use $\boldsymbol{v}^{ref} \triangleq \boldsymbol{v} - \boldsymbol{V}$ as the reference solution, and $\boldsymbol{V}$ is the constant inflow velocity. The reference velocity is null, far from the cylinder and the integration of its energy, on the domain, does not depend on the size of the domain.
Figures 6.5 and 6.6 illustrate the error obtained for the standard POD Galerkin model without

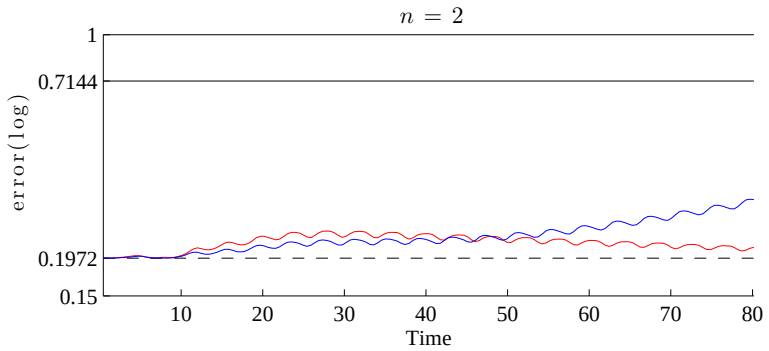


Figure 6.5: Normalized error for $n = 2$ modes, of a wake flow at Reynolds 300. The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue line corresponds to the standard POD Galerkin. The red one stands for our model with a constant variance tensor along time. The magenta one represents our model with linear representation of the variance tensor. The doted line indicates the error associated with the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error considering only the time mean velocity.

sub-grid dissipative term, and our model for a variance tensor which is either fixed constant along time or expressed as a linear combination of the *chronos*. For the Reynolds 300, only the model with constant variance has been used. Indeed, this fluid dynamics system has only few degrees of freedom. For this Reynolds number, the model with variance tensor varying in time is overparameterized. The doted line represents the minimal error associated with the reduced subspace truncation error. The black solid line is the error considering only the time mean velocity – if we set all the *chronos* to 0.
In this case:

$$\begin{aligned} err_{|b=0}(t) &= T \frac{\left\| \boldsymbol{v}^{ref} - \bar{\boldsymbol{v}} \right\|_{L^2(\Omega)}}{\| \boldsymbol{v}^{ref} \|_{L^2(\Omega \times [0,T])}}, \\ &= \left(\frac{\sum_{i=1}^{N} \left(b_i^{ref} \right)^2}{\| \bar{\boldsymbol{w}} \|^2_{L^2(\Omega)} + \sum_{i=1}^{N} \lambda_i} \right)^{1/2}, \end{aligned}$$

can be finally approximated as

$$err_{|b=0}(t) \approx \left(\frac{\sum_{i=1}^{N} \lambda_i}{\| \bar{\boldsymbol{w}} \|^2_{L^2(\Omega)} + \sum_{i=1}^{N} \lambda_i} \right)^{1/2}.$$

This term does not constitute an upper bound of the error. However, if this limit is crossed it means that the model is completely useless. In Figures 6.5 and 6.6, the fast exponential divergence of the standard POD reduced order (in blue) is clearly visible. Conversely, our methods, based on a physically relevant stochastic representation of the small scale component, have much weaker errors, without tuning any additional parameters on the data. There is only a slight difference between a constant and a linear representation of the variance tensor. A drawback of the second method is that $\boldsymbol{a}(\boldsymbol{x}, t)$ is not ensured to be a positive definite matrix. When the number of modes increases, the basis used for the projection of $\boldsymbol{a}$, $(b_i)_{0 \leqslant i \leqslant n}$, is larger. Thus, the projection approximates better the identity, and the estimation of $\boldsymbol{a}(\boldsymbol{x}, t)$ becomes closer to a positive matrix and close to $\boldsymbol{a}$. This may explain the difference between the two methods.
Whatever their differences, both methods provide very encouraging results. These representations clearly enable the construction of autonomous sub-grid models. This constitutes an essential point for the devising of autonomous reduced order dynamical systems.

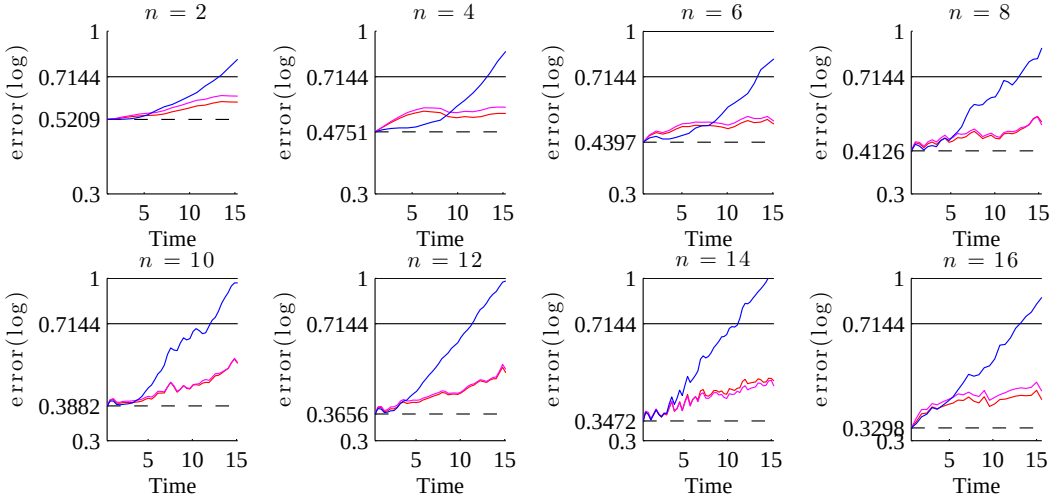


Figure 6.6: Normalized error for $n = 2, 4, 6, 8, 10, 12, 14$ and 16 modes, of a wake flow at Reynolds 3900. The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue line corresponds to the standard POD Galerkin. The red one stands for our model with a constant variance tensor along time. The magenta one represents our model with linear representation of the variance tensor. The doted line indicates the error associated with the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error considering only the time mean velocity.

6.6 Conclusion

In this paper, a fluid dynamics model built from fundamental physical principles applied to a stochastic representation of the flow has been used. In this representation, the fluid velocity is random and partially decorrelated in time. This time decorrelation can be interpreted as coming from a subsampling in time of a fast oscillating part of the velocity. In this framework, mass and momentum conservation principles can be constituted from stochastic calculus to derive a complete fluid flow dynamics model. This framework brings a strong theoretical support to classical empirical models, while generalizing them through the incorporation of an anisotropic, inhomogeneous and time-dependent diffusion. Compared to the original stochastic model, introduced in Mémin (2014), some initial assumptions have been removed. The diffusion tensor $\boldsymbol{\sigma}$ does not need to be symmetric anymore, $\boldsymbol{w}$ can be any semimartingale for the stochastic version of the Reynolds transport theorem and any finite variation process for the stochastic Navier-Stokes model. It has also been proved that the sub-grid tensor is diffusive when the density is constant. Thanks to our stochastic representation of fluid dynamics, a reduced model, describing the resolved modes evolution, has been derived. This model takes explicitly into account the unresolved modes influence. Since our stochastic model enables to deal with aliasing effects, we have chosen a time step as large as possible to simplify the reduced model simulation. A criterion based on Shannon-Nyquist theorem has been proposed to set the time step. Two different methods have been proposed to estimate the variance tensor. The first one relies on the assumption of a constant variance tensor along time, whereas the second one decomposes this tensor as a linear combination of the *chronos* basis. From both methods, closed autonomous reduced systems have been derived. Finally, in section 6.5, both methods have been tested on numerical data from DNS simulation at Reynolds 300 and LES simulation at Reynolds 3900 of wake flow. The two kinds of reduced models have been compared to POD Galerkin reduced system. The standard reduced system exhibits very fast diverging trajectories. On the contrary, our models have shown to provide much better results without any parameter tuning.
Those results are very encouraging. Indeed, we have written basic physical conservation laws in a stochastic framework where Itō formalism is interpreted as scale separation. This new methodology has yielded to a powerful stochastic fluid dynamics model. It is true that the variance tensor $\boldsymbol{a}$ remains to be estimated or modeled. Nevertheless, we have proposed two estimation methods of this tensor, based on stochastic calculus, in the context of reduced order model. These simple

estimation methods were sufficient to illustrate the potential of our new fluid dynamics model. Needless to say, a lot of improvements are possible and may be considered. Here, the variance has been assumed to be constant in time or in the span of *chronos*. However, it can also be assumed that it is a quadratic or a cubic function of the temporal modes. To obtain a more sophisticated time dependence for the variance tensor, a dynamical model can also be used. Many dynamical models of the sub-grid velocity variance exist in the literature, based mainly on heuristic observations or statistical estimations. For instance, the RANS equations are closed by empirical turbulent kinetic energy evolution equations (Menter (1992)). More recent works reveal new sub-grid dynamical models. Stochastic superparametrization (Grooms and Majda (2014)) is one example of such models. Using the so-called point approximation, the large-scale influence on the evolution equation of the small-scale velocity becomes constant and uniform. Then, a Gaussian closure decouples the small-scale Fourier modes and enables solving the small-scale variance dynamic. To go further, the sub-grid velocity can be non-Gaussian. Modified Quasilinear Gaussian (MQG) closure (Sapsis and Majda (2013c,b)) can be used instead of a Gaussian one. The Quasilinear Gaussian (QG) method neglects third order moments. The MQG algorithm enables simulating accurately the two first moments by modeling the third order moments (Sapsis and Majda (2013c)). This model is based on energy transfer principles and estimations on long time. The MQG-DO algorithm manages to also simulate accurately the other moments in a reduced subspace (Sapsis and Majda (2013b)). QG, MQG and MQG-DO closures will also lead to a dynamical model of the small-scale variance tensor $\boldsymbol{a}$. Another improvement of our algorithm could consist in using several time steps. One time step by resolved mode will involve one value of the variance tensor by resolved mode. This should make the most of the time-decorraleted unresolved velocity explicit influence. Therefore, the variance tensor would be different for each resolved modes. Finally, it could be suitable to remove the finite variations assumption for the large-scale drift, $\boldsymbol{w}$. This yields to a new model, that will be exposed in future works. In such a model, an evolution equation determines the partially time-correlated sub-grid velocity component: the martingale part of $\boldsymbol{w}$. This component is random, centered and not differentiable w.r.t. time. Therefore, it lives at a smaller scale than the finite-variation velocity component, studied previously in this paper. However, since the martingale component is continuous w.r.t. time, its evolution is smoother than the time-uncorrelated component. This new evolution equation, on the martingale part of the drift, may bring several advantages. First, this is a linear stochastic partial differential equation with additive and multiplicative noises, which is both easy to handle and physically pertinent (Majda et al. (1999)). Then, the noises covariances are naturally linked to the covariance of the time-uncorrelated velocity component $a(x, y)$ and to the sub-grid tensor. It implies a lot of interesting properties such as energy conservation, up to molecular viscous effect. Moreover, this stochastic equation is only inferred from fundamental physical laws. Therefore, neither tuning nor ad hoc model assumption is needed. For all these reasons, this new stochastic fluid dynamics model should be very helpful to built relevant uncertainty quantification (UQ) and sub-grid stochastic dynamic models, with reduced complexity. Such UQ methods can be used for stochastic reduced order models, filtering or probabilistic closures. And the associated sub-grid stochastic dynamics can be used, for instance, for stochastic superparametrization. Some work on this new model is currently ongoing.

Appendix

6.A Basic notions of stochastic calculus

We recall here some basic definitions and properties of stochastic calculus. Here, for simplicity, we only deal with functions of a compact set of time: $t \in [0, T]$ with $T \in \mathbb{R}_+^*$. However, everything can be generalized easily to functions of $\mathbb{R}_+ \times \Omega$ with $\Omega \subset \mathbb{R}^d$ (see Kunita (1997), Da Prato and Zabczyk (1992) and Prévôt and Röckner (2007)).
We use a sample space $\breve{\Omega}$, a probability measure $\mathbb{P}$, a Wiener process, $(B_t)_{t\geqslant 0}$, its filtration $(\mathcal{F})_{t\geqslant 0}$ (the set of $\boldsymbol{\sigma}$-algebra generated by each B_t), the whole σ-algebra, $\mathcal{F} \triangleq \mathcal{F}_\infty \triangleq \bigcup_{t\geqslant 0} \mathcal{F}_t$, and the resulting filtered probability space $(\breve{\Omega}, \mathcal{F}, (\mathcal{F})_{t\geqslant 0}, \mathbb{P})$

Defintion 6.A.1 *Finite variation function*
$t \to f(t)$ is a finite variation function if and only if for all $a < b$ and all partition $a = t_0 < ... < t_n = b$ of $[a, b]$, $\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| < \infty$.

Defintion 6.A.2 *Finite variation process*
$(t, \omega) \to f(t, \omega)$ has finite variations if and only if:

- *f is adapted (i.e. $f(t, .)$ is $\mathcal{F}_t$ measurable),*
- *For each trajectory ω, $f(., \omega)$ is a finite variation function .*

Characterization:
f is a finite variation process if and only if $\exists g, f(t, .) = f(0, .) + \int_0^t g(t', .)\mathrm{d}t'$.

Defintion 6.A.3 *Martingale*
$(t, \omega) \to f(t, \omega)$ is a martingale if and only if:

- *f is adapted ,*
- *$f(t, .) \in L^1_{\tilde{\Omega}} \triangleq \{Y : \mathbb{E}|Y| < \infty\}$,*
- *$\forall s < t,\ \mathbb{E}(f(t, .)|\mathcal{F}_s) = f(s, .)$.*

In particular, if $f = 0$ at $t = 0$, then f is a centered process.
Characterization:
f is a martingale if and only if $\exists g, f(t, .) = f(0, .) + \int_0^t g(t', .)\mathrm{d}B_{t'}$.

Defintion 6.A.4 *Continuous semimartingale*
f is a continuous semimartingale if and only if it the sum of a finite variation process and a martingale

Stochastic calculus deals only with semimartingales. In our fluid dynamics representation, we also deal with time-decorrelated processes, formally, the differentiation along time of a martingale.

Defintion 6.A.5 *Quadratic variation and quadratic cross-variation*
If f and g are semimartingale and $f(t = 0) = g(t = 0) = 0$, then, their quadratic cross-variation, noted $< f, g >$, is the unique finite variation process such $fg- < f, g >$ is a martingale and $< f, g >_{t=0}= 0$.

Characterization:

- If $f(t, .) = \int_0^t f_1(t')\mathrm{d}t' + \int_0^t f_2(t')\mathrm{d}B_{t'}$ and $g(t, .) = \int_0^t g_1(t')\mathrm{d}t' + \int_0^t g_2(t')\mathrm{d}B_{t'}$, then $< f, g >_t= \int_0^t f_2(t')g_2(t')\mathrm{d}t'$.
  It should be noticed that, if f_2 and g_2 are random, $< f, g >$ is also random.
- $< f, g >_t= \mathbb{P} \underset{\Delta t \to 0}{lim} \sum_{t_i=0}^{t_n=t} (f(t_i) - f(t_{i-1}))(g(t_i) - g(t_{i-1}))$.

Thus, $< f, g >$ may be interpreted as a kind of "covariance along time".

Theorem 6.A.1 *Itô-Wentzell Formula*
If $(t, x) \to f(t, x)$ and $(t, y) \to g(t, y)$ are semimartingale (as function of time), and $x \to f(., x)$ is twice differentiable, then

$$\mathrm{d}_t\left(f(t, g(t, y))\right) = \mathrm{d}_t f + \partial_x f \mathrm{d}_t g + \frac{1}{2}\partial^2_{xx} f \mathrm{d}_t < g, g > + \mathrm{d}_t < \partial_x f, g > .$$

6.B Equivalence of formulation for the stochastic transport and Navier-Stokes theorem

6.B.1 Stochastic Reynolds transport model

Mémin (2014) describes the stochastic transport theorem as follows:

$$\mathrm{d}\int_{V(t)} q(\boldsymbol{x},t)\mathrm{d}\boldsymbol{x} = \int_{V(t)} \left(\mathrm{d}_t q + \left[\boldsymbol{\nabla}\cdot(q\boldsymbol{w}) + \frac{1}{2}\|\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\|^2 q - \frac{1}{2}\sum_{i,j} \partial^2_{ij}(a_{ij}q)_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0} \right] \mathrm{d}t + \boldsymbol{\nabla}\cdot(q\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \right) \mathrm{d}\boldsymbol{x},$$

where:

$$\begin{aligned}
\sum_{i,j} \partial^2_{ij}(a_{ij}q)_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0} &= \sum_{i,j} (\partial^2_{ij}(\boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.}))_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0} q + 2(\partial_i(\boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.}))_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0})\partial_j q + a_{ij}\partial^2_{ij} q, \\
&= \sum_{i,j} \partial_j \boldsymbol{\sigma}_{i.}\partial_i \boldsymbol{\sigma}^t_{j.} q + 2\boldsymbol{\sigma}_{i.}\partial_i \boldsymbol{\sigma}^t_{j.}\partial_j q + a_{ij}\partial^2_{ij} q, \\
&= \sum_{i,j} \partial^2_{ij}(a_{ij}q) - 2\partial^2_{ij}\boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.} q - \partial_i \boldsymbol{\sigma}_{i.}\partial_j \boldsymbol{\sigma}^t_{j.} q - 2\partial_i \boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.}\partial_j q.
\end{aligned}$$

So,

$$\begin{aligned}
\frac{1}{2}\|\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\|^2 q - \frac{1}{2}\sum_{i,j} \partial^2_{ij}(a_{ij}q)_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0} &= \sum_{i,j} -\frac{1}{2}\partial^2_{ij}(a_{ij}q) + \partial^2_{ij}\boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.} q \\
&\qquad + \partial_i \boldsymbol{\sigma}_{i.}\partial_j \boldsymbol{\sigma}^t_{j.} q + \partial_i \boldsymbol{\sigma}_{i.}\boldsymbol{\sigma}^t_{j.}\partial_j q, \\
&= -\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}q)^T) + \boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T q), \\
&= \boldsymbol{\nabla}\cdot\left(-\frac{1}{2}(\boldsymbol{\nabla}\cdot(\boldsymbol{a}q)^T) + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T q \right).
\end{aligned}$$

6.B.2 Stochastic Navier-Stokes model

Mémin (2014) describes the r-th coordinate of the diffusion tensor of the stochastic Navier-Stokes model as:

$$\begin{aligned}
\tau_r(\boldsymbol{w}) &= \sum_{i,j} \frac{1}{2}\rho a_{ij}\partial^2_{ij}(\boldsymbol{w}_r) + \partial_i(\rho a_{ij})_{|\boldsymbol{\nabla}\boldsymbol{\sigma}=0}\partial_j \boldsymbol{w}_r, \\
&= \sum_{i,j} \frac{1}{2}\rho a_{ij}\partial^2_{ij}(\boldsymbol{w}_r) + \partial_i(\rho a_{ij})\partial_j \boldsymbol{w}_r - \rho\partial_i(\boldsymbol{\sigma}_{i.})\boldsymbol{\sigma}_{j.}\partial_j \boldsymbol{w}_r, \\
&= \frac{1}{2}(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\rho\boldsymbol{a}\boldsymbol{w}_r)^T)) - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\rho\boldsymbol{a})^T)))\boldsymbol{w}_r - \rho\boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\boldsymbol{\sigma}^T\boldsymbol{\nabla}\boldsymbol{w}_r.
\end{aligned}$$

6.C Dissipative effect of sub-grid tensor τ

If the density, ρ, is assumed to be constant, then $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$ by the martingale part of the mass conservation. Moreover, $\boldsymbol{a}$ or $\boldsymbol{w}$ and its derivatives are assumed to be null in the border of Ω, then

with two integrations by parts,

$$\begin{aligned}
\sum_k \int_\Omega \boldsymbol{w}_k \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}\boldsymbol{w}_k))^T \,\mathrm{d}\boldsymbol{x} &= \int_\Omega \boldsymbol{w}\cdot\sum_{i,j}\partial^2_{ij}\left(a_{i,j}\boldsymbol{w}\right)\mathrm{d}\boldsymbol{x} \\
&= -\int_\Omega \sum_{i,j}\partial_i \boldsymbol{w}^t \partial_j\left(a_{i,j}\boldsymbol{w}\right)\mathrm{d}\boldsymbol{x}, \\
&= -\int_\Omega \sum_{i,j}\left(\partial_i \boldsymbol{w}^T \partial_j a_{i,j}\boldsymbol{w} + \partial_i \boldsymbol{w}^T a_{i,j}\partial_j \boldsymbol{w}\right)\mathrm{d}\boldsymbol{x}, \\
&= -\int_\Omega \sum_{i,j}\left(\frac{1}{2}\partial_i \|\boldsymbol{w}\|_2^2 \partial_j a_{i,j} + \partial_i \boldsymbol{w}^T a_{i,j}\partial_j \boldsymbol{w}\right)\mathrm{d}\boldsymbol{x}, \\
&= \int_\Omega \left(\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\|\boldsymbol{w}\|_2^2 - \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\right)\mathrm{d}\boldsymbol{x},
\end{aligned}$$

where $\|\boldsymbol{\nabla}\boldsymbol{w}^t\|_{\boldsymbol{a}}^2 \triangleq \sum_k \|\boldsymbol{\nabla}\boldsymbol{w}_k\|_{\boldsymbol{a}}^2 \triangleq \sum_k \boldsymbol{\nabla}\boldsymbol{w}_k^T\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{w}_k = tr\left((\boldsymbol{\nabla}\boldsymbol{w}^T)^T\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{w}^T\right)$.

$$\begin{aligned}
2\int_\Omega \boldsymbol{w}\cdot\tau\,\mathrm{d}\boldsymbol{x} &= \rho\int_\Omega\left(\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\|\boldsymbol{w}\|_2^2 - \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2 - \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\|\boldsymbol{w}\|_2^2\right)\mathrm{d}\boldsymbol{x}, \\
&= -\rho\int_\Omega\left(\frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T\|\boldsymbol{w}\|_2^2 + \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\right)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

Using now the finite variation part of the mass conservation, which is $\boldsymbol{\nabla}\cdot\boldsymbol{w} = \frac{1}{2}\boldsymbol{\nabla}\cdot(2\boldsymbol{\nabla}\cdot\boldsymbol{a})^T$, we get:

$$\begin{aligned}
2\int_\Omega \boldsymbol{w}\cdot\tau\,\mathrm{d}\boldsymbol{x} &= -\rho\int_\Omega\left(\boldsymbol{\nabla}\cdot\boldsymbol{w}\|\boldsymbol{w}\|_2^2 + \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\right)\mathrm{d}\boldsymbol{x}, \\
&= \rho\int_\Omega\left((\boldsymbol{w}\cdot\boldsymbol{\nabla})\|\boldsymbol{w}\|_2^2 - \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\right)\mathrm{d}\boldsymbol{x}, \\
&= \rho\int_\Omega\left(2\,\boldsymbol{w}\cdot((\boldsymbol{w}\cdot\boldsymbol{\nabla})\boldsymbol{w}) - \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\right)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

Considering together the advection term and the sub-grid term of the Navier Stokes equation we have for the energy:

$$\int_\Omega \boldsymbol{w}\cdot\left(-\rho\left(\boldsymbol{w}\cdot\boldsymbol{\nabla}\right)\boldsymbol{w} + \tau\right)\mathrm{d}\boldsymbol{x} = -\frac{\rho}{2}\int_\Omega \|\boldsymbol{\nabla}\boldsymbol{w}^T\|_{\boldsymbol{a}}^2\mathrm{d}\boldsymbol{x} < 0, \tag{6.C.1}$$

which is the sought result. It should be noted that in the incompressible deterministic equation, $\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0$ and thus the advection term $(\boldsymbol{w}\cdot\boldsymbol{\nabla})\boldsymbol{w}$ does not influence the global energy. Here however, it is not the case anymore and this term has to be taken into account, as above.

Chapter 7

Stochastic modeling and diffusion modes for POD models and small-scale flow analysis

V. Resseguier, E. Mémin, and B. Chapron. Stochastic modeling and diffusion modes for POD models and small-scale flow analysis. Under consideration for Journal of Fluid Mechanics, 2017d

ABSTRACT

We introduce a stochastic modelling in the constitution of fluid flow reduced order models. This framework introduces a spatially inhomogeneous random field to represent the unresolved small-scale velocity component. Such a decomposition of the velocity in terms of a smooth large-scale velocity component and a rough, highly oscillating, component gives rise, without any supplementary assumption, to a large-scale flow dynamics that includes a modified advection term together with an inhomogeneous diffusion term. Both of those terms, related respectively to *turbophoresis* and *mixing* effects, depend on the variance of the unresolved small-scale velocity component. They bring to the reduced system an explicit subgrid term enabling to take into account the action of the truncated modes. Besides, a decomposition of the variance tensor in terms of *diffusion modes* allows us to provide a meaningful statistical representation of the stationary or nonstationary structuration of the small-scale velocity and of its action on the resolved modes. This supplies a useful tool for turbulent fluid flows data analysis. We apply this methodology to circular cylinder wake flow at Reynolds numbers $Re = 300$ and $Re = 3900$, respectively. The finite dimensional models of the wake flows reveal the energy and the anisotropy distributions of the small-scale diffusion modes. These distributions identify critical regions where corrective advection effects as well as structured energy dissipation effects take place. In providing rigorously derived subgrid terms, the proposed approach yields accurate and robust temporal reconstruction of the low-dimensional models.

7.1 Introduction

Surrogate empirical models of flow dynamics with a reduced set of degrees of freedom are widely used in fluid mechanics for control applications or physical analysis (Noack et al., 2010). Within such modelling a few numbers of modes extracted from experimental or numerical measurements are used to represent the main dynamical behaviour of a flow. The modes in themselves may help unveiling recurrent dynamical patterns. Spectral approaches are quite natural for that purpose. Fourier representation has been used for a long time to characterize hydrodynamics instabilities. Proper Orthogonal Decomposition (POD) and the spectral representation of the velocity auto-correlation matrix is used to extract descriptive empirical spatial or temporal basis of the flow (Aubry et al., 1988; Holmes et al., 1996; Sirovich, 1987). More recently the Dynamic Modes Decomposition (DMD) relying on the eigenvectors of the Koopman operator (Koopman, 1931)

and Takens delay embedding theorem (Takens, 1981) has been proposed to represent, from the evolution of observations, the principal modes of the dynamical system's attractor (Mezic, 2005; Rowley et al., 2009; Schmid, 2010). Combination of both representations can be used as well (Cammilleri et al., 2013) to provide a suitable energy spectrum representation. In all those modal representations the construction of the reduced order dynamics requires a truncation operation in which the most "influential" modes – with respect to a given criterion – are kept to describe the flow. In general, the action of the discarded modes must be modeled to get accurate and stable dynamical systems. The effect of those neglected processes encompasses dissipation effects but are also responsible of some energy redistribution and backscattering (Piomelli et al., 1991).

In most of the flow low-order dynamics, the unresolved small-scale processes are represented on the basis of an eddy viscosity assumption (Boussinesq, 1877). This takes the form of a damping term in the reduced order dynamical system. In Galerkin POD reduced models, this extra dissipation, which adds up to the linear molecular diffusion, is modeled by a constant coefficient (Aubry et al., 1988) or through a modal constant vector (Cazemier et al., 1998; Rempfer and Fasel, 1994). Recently, nonlinear functions have been proposed for a bluff body wake flow (Östh et al., 2014). Although those models have demonstrated their efficiency in numerous situations, the estimation of the associated parameters and/or the choice of the nonlinear dependency between the eddy-viscosity coefficients and the modal coefficients constitute a sensible issue. Furthermore, from a physical interpretation point of view, the action of the small-scale velocity component is interpreted only with regard to an homogeneous stationary dissipation effect. Neither preferential local direction of diffusion related to the flow physics is considered, nor energy redistribution action by the small scales.

Robust techniques based on optimal control strategies have also been proposed for building reduced dynamical models from noisy data (Artana et al., 2012; D'Adamo et al., 2007; Tissot et al., 2013; Semaan et al., 2016) and incomplete knowledge of the actual flow dynamics (i.e. unknown initial condition, partially known forcing terms, etc.). Those techniques accurately estimate low-order dynamical systems in the temporal windows on which the data are available. These methods unfortunately experience some limitations for forecasting new states of the system. Furthermore, the physical interpretation of the unresolved velocity component remains difficult since their contribution is distributed in an unknown manner over all the coefficients of the dynamical system and on the error function (when weak dynamical constraint (Artana et al., 2012) is considered).

In this work, to take into account the unresolved modes in the surrogate dynamic model, we will rely on a recently proposed stochastic framework (Mémin, 2014; Resseguier et al., 2017a). In this context, an advection of the large-scale component due to the action of the unresolved random component emerges naturally, together with an inhomogeneous nonstationary diffusion. This will lead us to consider corrective advection and diffusion terms driven by the turbulence inhomogeneity whose local effects can now now physically interpreted.

After presenting the stochastic model in section 7.2, section 7.3 describes the derivation of the associated POD reduced order model. We also propose a method to estimate the additional components of the dimensional reduced system from the residual velocity. Then, the data benchmarks are detailed in section 7.4. From the estimated additional components, we analyse in section 7.5 the influence of the residual velocity on the large-scale flow and reconstruct the temporal modes of the reduced order models.

7.2 Dynamics stochastic modelling

The proposed stochastic principle relies on a Lagrangian random description of the flow velocity:

$$\frac{\mathrm{d}\boldsymbol{X}_t}{\mathrm{d}t} = \boldsymbol{w}(\boldsymbol{X}_t, t) + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)\dot{\boldsymbol{B}}. \tag{7.2.1}$$

The first right-hand term, $\boldsymbol{w}$, stands for the large-scale velocity component. It is a smooth component along time. For turbulent flows, it is associated with a much larger time-scale than the unresolved small-scale velocity component. This latter, $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}/\mathrm{d}t$, is associated with fast modes that are rapidly decorrelating at the resolved time scale. Based on this observation, we will assume that such a component can be ideally represented through a spatially smooth incompressible (divergence-free) Gaussian random field uncorrelated in time. This (possibly inhomogeneous)

random field is formally built from an infinite-dimensional Brownian motion. It is associated with a covariance tensor denoted:

$$Q_{ij}(\boldsymbol{x},\boldsymbol{y},t,t') = \mathbb{E}((\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_i(\boldsymbol{x},t)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_j(\boldsymbol{y},t')) = c_{ij}(\boldsymbol{x},\boldsymbol{y},t)\delta(t-t')\mathrm{d}t. \tag{7.2.2}$$

In the following, the diagonal of the covariance tensor, which plays a central role in our setting, will be denoted as: $\boldsymbol{a}(\boldsymbol{x}) \triangleq \boldsymbol{c}(\boldsymbol{x},\boldsymbol{x},t)$. This tensor, that may depend on time, will be referred to as the small-scale variance tensor. It is a symmetric positive definite matrix at all spacial points, $\boldsymbol{x}$ (excluding degenerate cases) with dimension in $m^2.s^{-1}$. It corresponds thus to an eddy viscosity term.

This stochastic formulation is related in spirit to the Lagrangian stochastic models based on Langevin equations that have been intensively used for turbulent dispersion (Sawford., 1986) or in probability density function (PDF) modelling of turbulent flows (Haworth and Pope, 1986; Pope, 1994, 2000). However, here our interest focuses more on the associated large-scale Eulerian representations of the flow dynamics. This Eulerian description of the resolved velocity component is obtained through a formulation of the Reynolds transport theorem adapted to such a stochastic flow.

7.2.1 Stochastic conservation equations

Considering the flow decomposition (7.2.1), the rate of change of a scalar quantity (in the absence of random forcing) within a material volume is given by the following expression (Mémin, 2014; Resseguier et al., 2017a):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{V(t)} q\,\mathrm{d}\boldsymbol{x} = \int_{V(t)}\left(\frac{\partial q}{\partial t} + \boldsymbol{\nabla}\cdot(q\boldsymbol{w}^*) - \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q\right) + \boldsymbol{\sigma}\dot{\boldsymbol{B}}\cdot\boldsymbol{\nabla}q\right)\mathrm{d}\boldsymbol{x}, \tag{7.2.3}$$

where the effective advection velocity is given as:

$$\boldsymbol{w}^* \triangleq \boldsymbol{w} - \frac{1}{2}\big(\boldsymbol{\nabla}\cdot\boldsymbol{a}\big)^T. \tag{7.2.4}$$

Equation (7.2.3) provides a stochastic representation of the so-called Reynolds transport theorem. It is important to outline that at a given grid point, q is a random value which depends among other things on the Brownian component of the particles flowing through that point. The second term corresponds to the large-scale advection by an effective drift, $\boldsymbol{w}^*$, that includes a contribution related to the divergence of the small-scale velocity variance tensor (7.2.4). The third term is a diffusion expressing the mixing effect exerted by the small-scale velocity component. The final term corresponds to the scalar advection by the small-scale velocity field. From this expression a conservation of an extensive property, $\int_{V(t)} q$, such as mass or internal energy (neglecting diabatic and compressive effects) reads immediately as the following intensive property evolution equation

$$\frac{\partial q}{\partial t} + \boldsymbol{\nabla}\cdot(q\boldsymbol{w}^*) + \boldsymbol{\sigma}\dot{\boldsymbol{B}}\cdot\boldsymbol{\nabla}q = \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}q\right). \tag{7.2.5}$$

As the right-hand term is a smooth temporal component, we observe immediately that the Brownian terms associated, on the one hand, to the scalar temporal variation and, on the other hand, to the small-scale advection necessarily compensate each other. A fluid with a constant density ρ, requires naturally a divergence-free constraint on the effective advection:

$$0 = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma}\dot{\boldsymbol{B}}, \tag{7.2.6}$$

$$0 = \boldsymbol{\nabla}\cdot\boldsymbol{w}^* = \boldsymbol{\nabla}\cdot\left(\boldsymbol{w} - \frac{1}{2}(\boldsymbol{\nabla}\cdot\boldsymbol{a})\right). \tag{7.2.7}$$

This is the case we are dealing with in this study. The two constraints (7.2.6-7.2.7) correspond to the incompressibility conditions associated with the stochastic representation. For isochoric flows with variable density as in geophysical fluid dynamics, interested readers can refer to Resseguier et al. (2017a,b,c).

7.2.2 Navier-Stokes equations associated with a stochastic representation of the small-scales

Similarly to the Newton second law, a dynamical balance between the temporal differentiation of the stochastic momentum, $\rho \mathrm{d}\boldsymbol{X}_t$, and the action of the forces is assumed. Applying the stochastic representation of the Reynolds transport theorem (7.2.3) leads to the following Navier-Stokes equations (Mémin, 2014):

$$\frac{\partial \boldsymbol{w}}{\partial t} + (\boldsymbol{w}^* \cdot \boldsymbol{\nabla})\boldsymbol{w} = -\frac{1}{\rho}\boldsymbol{\nabla} p + \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(\frac{1}{2} a_{ij} \frac{\partial \boldsymbol{w}}{\partial x_j}\right) + \nu \triangle \boldsymbol{w}. \tag{7.2.8}$$

This equation corresponds to the large-scale momentum equation. This expresssion differs from the classical Reynolds decomposition formulation mainly by the introduction of both a large-scale dissipation term and a correction term in the large-scale advection. The dissipative term plays a role that is similar to the eddy viscosity models introduced in classical large scale representations (Bardina et al., 1980; Lilly, 1992; Smagorinsky, 1963) or to the spectral vanishing viscosity (Karamanos and Karniadakis, 2000; Pasquetti, 2006; Tadmor, 1989). It is also akin to numerical regularization models considered in implicit models (Aspden et al., 2008; Boris et al., 1992; Lamballais et al., 2011). The small-scale stochastic representation principle is nevertheless more general as it does not rely on *a priori* fixed shapes of the subgrid tensor (e.g. Boussinesq assumption) nor does it presuppose a given numerical scheme (e.g. implicit models). The subgrid term takes a general diffusion form whose matrix coefficients are given by the small-scale variance tensor. The diffusion principal directions are thus aligned with this tensor principal directions.

The advection correction term is much less intuitive. It is related here to an advection bias due to the inhomogeneity of the small-scale variance tensor. This corresponds to the eddy-induced velocity introduced for tracer mean transport in oceanic or atmospheric circulation models (Andrews and McIntyre, 1976; Gent et al., 1995) and more generally to the *turbophoresis* phenomenon associated with small-scale inhomogeneity, which drives inertial particles toward the regions of lower diffusivity (Brooke et al., 1992; Caporaloni et al., 1975; Reeks, 1983; Sehmel, 1970). Qualitatively, this drift correction can be understood as follow. Fluid parcels with higher turbulent kinetic energy (TKE) move faster. It ensues that at large scales, areas associated with maximum of TKE spread whereas areas associated with minimum of TKE shrink. Hence, a large-scale drift oriented toward these maxima/minimum emerges. This orientation suggests an anticorrelation with the TKE gradient. Since the turbulent velocity variations are multidimensional, they are better described by the variance tensor. The drift correction is consequently proportional to the opposite of the variance tensor divergence. For homogeneous turbulence, the small-scale variance tensor is constant and this corrective advection does not come into play. It can be noted that this advection correction is of the same form as the one proposed in Caporaloni et al. (1975); MacInnes and Bracco (1992); Reeks (1983).

The small-scale random field can be freely defined and be in a shape that goes from isotropic stationary models up to inhomogeneous non-stationary random fields. However, in the inhomogeneous case (such as the Smagorinsky model) the advection correction term comes into play. A stochastic representation of the unresolved scales differ thus significantly from classical large-scale modelling. It relies on less strict assumptions, which enable to cope naturally with inhomogeneous anisotropic turbulence.

This stochastic representation relies on a scale gap assumption, which is coherent with deterministic justifications of the eddy viscosity (Kraichnan, 1987). The stochastic transport expression (7.2.3) and the momentum equation (7.2.8) provide the foundations of a physically relevant large-scale fluid dynamics formulation. It opens a new paradigm for large-scale modelling adapted to turbulence inhomogeneity in involving a general subgrid diffusion together with a small-scale drift correction. In the next section, we will rely on this model for the construction of reduced order dynamical systems.

7.3 Reduced order models

Dimensional reduction techniques enable the constitution of simplified lower dimensional representations of partial differential equations (PDE). They are usually specified from a Galerkin projection onto data-based dedicated basis. The proper orthogonal decomposition, also called Empirical Orthogonal Functions (EOF) in geophysics, is one of those methods for turbulent flows. In §7.3.1 the POD model reduction is briefly presented. Then, in §7.3.2 we introduce the derivation of the reduced order model from the stochastic representation principle described in §7.2. In §7.3.3, different characteristic time scales are introduced for the different modes, leading to the concept of modal characteristic time steps. Finally, in \$ 7.3.4 a precise specification of the small-scale variance tensor is proposed with two different estimation methods.

7.3.1 POD model reduction

POD reduced order models rely on the linear decomposition of the velocity $\boldsymbol{w}$ on a reduced number of orthogonal spatial modes (Holmes et al., 1996):

$$\boldsymbol{w}(\boldsymbol{x},t) \approx \sum_{i=0}^{n} b_i(t)\boldsymbol{\phi}_i(\boldsymbol{x}). \tag{7.3.1}$$

The number of modes, n, is assumed to be much lower than the state space dimension. The functions $(\boldsymbol{\phi}_i(\boldsymbol{x}))_{1\leqslant i\leqslant N}$ encoding the spatial flow variations are referred to as *topos* and are computed from a Karunen-Loeve decomposition on a series of $N+1$ available velocity snapshots. The *topos* are sorted by decreasing order of the snapshots empirical covariance eigenvalues: $\lambda_1 > ... > \lambda_N$. The $(b_i(t))_{1\leqslant i\leqslant N}$ denote the temporal modes; they are called *chronos*. The *chronos* are the eigenvectors of the spatially averaged temporal correlation matrix, whereas the *topos* constitute the eigenvectors of the temporally averaged spatial correlation matrix. They are both computed from the snapshots covariance. Function ϕ_0 corresponds to the time average velocity and $b_0 \triangleq \lambda_0 \triangleq 1$. We also denote by T the time between the first and the last snapshot. The Navier-Stokes equations can be written in the general following form:

$$\frac{\partial \boldsymbol{w}}{\partial t} = \boldsymbol{I} + \boldsymbol{L}(\boldsymbol{w}) + \boldsymbol{C}(\boldsymbol{w},\boldsymbol{w}), \tag{7.3.2}$$

where $\boldsymbol{L}$ and $\boldsymbol{C}$ stand respectively for linear and bilinear differential operators. The first term, $\boldsymbol{I}$, gathers the pressure and the external forces such as gravity. The second one, $\boldsymbol{L}$, includes the molecular friction term and possibly the Coriolis force. The last one, $\boldsymbol{C}$, encodes the nonlinear advection term. Projecting this PDE on each *topos* leads to:

$$\frac{\mathrm{d}b_i}{\mathrm{d}t} = \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot \boldsymbol{I}\right)}_{\triangleq i_i} + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot \boldsymbol{L}(\boldsymbol{\phi}_p)\right)}_{\triangleq l_{p,i}} b_p + \sum_{p,q=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot \boldsymbol{C}(\boldsymbol{\phi}_p,\boldsymbol{\phi}_q)\right)}_{\triangleq c_{p,q,i}} b_p b_q. \tag{7.3.3}$$

Due to nonlinearity, the temporal modes strongly interact with one another. In particular, even though the original model (with $n = N$) is computationally stable for moderate Reynolds number, a strongly reduced model ($n \ll N$) appears unstable in general. A frequency shift is also often observed. Those artefacts are extensively documented in the literature (Artana et al., 2012; Aubry et al., 1988; Rempfer and Fasel, 1994; Östh et al., 2014; Protas et al., 2015). The introduction of a damping eddy viscosity term to mimic the truncated modes' dissipation leads to a modified linear term in (7.3.3). Unfortunately, as this term is built on empirical grounds its precise form is difficult to justify. Furthermore, its parametrization has to be tuned for each simulation to achieve good results. When large wake domains are considered the influence of the pressure term (and of the boundaries) is in general negligible (Deane et al., 1991; Ma et al., 2002; Noack et al., 2005). We will also rely on this assumption, although several authors have shown that neglecting the pressure term was a source of uncertainty regarding an accurate representation of the flow dynamics (Kalb and Deane, 2007; Noack et al., 2005). To take into account the effect of the outflow boundary, corrective terms are introduced by some authors through modifications of the linear (Galetti et al., 2007) or quadratic terms (Noack et al., 2005).

7.3.2 Reduced order modelling associated with the stochastic representation

To overcome the difficulties evoked previously, we propose to derive the reduced order model from the stochastic representation principle described previously. To account for the effect due to the modal truncation, we will assume that the whole field $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{\sigma}\dot{\boldsymbol{B}}$ can be decomposed in such a way that the large-scale component lives on the subspace endowed with the reduced POD basis $\boldsymbol{w} = \sum_{i=0}^{n} b_i \boldsymbol{\phi}_i$ while realizations of the small-scale component belong to the orthogonal complement subspace $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \sum_{i=n+1}^{N} b_i \boldsymbol{\phi}_i$. Since $\boldsymbol{\nabla} \cdot \boldsymbol{v} = 0$, for all i, $\boldsymbol{\nabla} \cdot \boldsymbol{\phi}_i = 0$ and, then, $\boldsymbol{\nabla} \cdot \boldsymbol{w} = 0$. The dynamics of the large-scale component, $\boldsymbol{w}$, is given by the incompressible Navier-Stokes equations (7.2.8). Projecting this equation onto the *topos* $\boldsymbol{\phi}_i$ leads to:

$$\frac{\mathrm{d}b_i}{\mathrm{d}t} = i_i + \left(\boldsymbol{l}_{\bullet i} + \breve{\boldsymbol{f}}(\boldsymbol{a})_{\bullet i}\right)^T \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i} \boldsymbol{b} \,, \tag{7.3.4}$$

$$\text{where } \breve{\boldsymbol{f}}(\boldsymbol{a})_{j,i} \triangleq \int_\Omega \sum_{k=1}^{d} \boldsymbol{\phi}_i^{(k)} \cdot \left(\underbrace{-\frac{1}{2}(\boldsymbol{\nabla} \cdot \boldsymbol{a})\boldsymbol{\nabla}\phi_j^{(k)}}_{\text{Advection}} + \underbrace{\boldsymbol{\nabla} \cdot \left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\phi_j^{(k)}\right)}_{\text{Diffusion}} \right). \tag{7.3.5}$$

Let us note that $\breve{\boldsymbol{f}}(\boldsymbol{a})$ is linear and is the only function that depends on the variance tensor $\boldsymbol{a}$. This system includes now a natural small-scales dissipation mechanism, through the diffusion term. But it also corrects the frequency shift through the additional advective term brought by the variance tensor inhomogeneity. To fully define this system, we need to specify the small-scale variance tensor $\boldsymbol{a}$. This issue is developed in subsection 7.3.4. But before that, we will elaborate further on the choice of the characteristic times related to the modal truncation.

7.3.3 Time scale characterisation

Very efficient flow simulations are obtained by reducing as much as possible the number of modes of the associated surrogate model. An even higher efficiency can be obtained by increasing the evolution time step. This time step can be naturally chosen as a single constant for the whole system. However, as we shall see, different characteristic time scales can be fixed for the different modes, leading to the concept of modal characteristic time steps.

Single time step

As long as the resolved modes, representing $\boldsymbol{w}$, are smooth w.r.t. time, the assumption pertaining to our reduced model construction is valid. The time-step must thus be fixed as the largest value that guarantees that all the *chronos* remain smooth. The characteristic time scale associated with the fastest resolved mode (which is often the least energetic mode) is a good target for that purpose. This time scale is associated with the highest frequency of the *chronos* Fourier modes. Quantitatively, the Shannon-Nyquist theorem provides us with a natural upper bound to fix its value. This theorem states that a function can be sampled, without loss of information, if the sampling frequency is two times bigger than the largest frequency of the original function. Otherwise, the sampled function undergoes an aliasing artifact characterized by a folding of the Fourier spectrum and a loss of regularity. We will thus assume that the required regularity condition is fulfilled if the modes are not affected by aliasing phenomena. Since the evolution equations are quadratic, a sufficient condition thus reads:

$$\frac{1}{\Delta t} \geqslant 4 \max_{i \leqslant n} \left(f_{max}\left(b_i\right)\right). \tag{7.3.6}$$

where $f_{max}\left(b_i\right)$ is the maximum frequency of the i-th mode. Aliasing takes place in the unresolved temporal modes, which are associated with smaller time scales. However, the stochastic representation is precisely built from a decorrelation assumption of the small-scale unresolved part of the velocity. A strong subsampling of those components strengthens further the decorrelation property of the unresolved modes.

Modal characteristic times

The resolved *chronos* are associated with different time scales. Hence, the previous criterion can be immediately extended to get a characteristic time scale attached to each *chronos*, b_i:

$$\frac{1}{\Delta t_i} \geqslant 4 f_{max}\left(b_i\right). \tag{7.3.7}$$

A modal variance tensor field for each *chronos* can now be defined as

$$\boldsymbol{a}^{(i)}(\boldsymbol{x}) \triangleq \Delta t_i \mathbb{E}\left\{\boldsymbol{\sigma}\dot{\boldsymbol{B}}\left(\boldsymbol{\sigma}\dot{\boldsymbol{B}}\right)^T\right\}(\boldsymbol{x}) = \frac{\Delta t_i}{\Delta t}\tilde{\boldsymbol{a}}(\boldsymbol{x}) \text{ with } \tilde{\boldsymbol{a}}(\boldsymbol{x}) = \Delta t\ \mathbb{E}\left\{\boldsymbol{\sigma}\dot{\boldsymbol{B}}\left(\boldsymbol{\sigma}\dot{\boldsymbol{B}}\right)^T\right\}(\boldsymbol{x}). \tag{7.3.8}$$

The modal variance tensor $\boldsymbol{a}_i$ corresponds to the small-scale velocity variance during a given time step at the time-scale of the *Chronos* b_i (i.e. it corresponds to an eddy viscosity associated with the neglected modes expressed with respect to the characteristic time associated with b_i). The *chronos* evolution equation (7.3.4) thus becomes:

$$\frac{\mathrm{d}b_i}{\mathrm{d}t} = i_i + \left(\boldsymbol{l}_{\bullet i} + \frac{\Delta t_i}{\Delta t}\breve{\boldsymbol{f}}(\tilde{\boldsymbol{a}})_{\bullet i}\right)^T \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i}\boldsymbol{b}. \tag{7.3.9}$$

The time step Δt corresponds in practice to the simulation time step.

7.3.4 Estimation of the small-scale variance tensor

The full definition of the reduced order model requires a precise specification of the small-scale variance tensor. We compare here two different estimation methods for this tensor. A first method will rely on a stationarity assumption while a second technique will allow us to define a time-varying tensor.

Stationary small-scale variance tensor

This case corresponds to the model developed in Mémin (2014). The small-scale velocity variance, $\boldsymbol{a}/\Delta t$ can be computed through a temporal averaging of the residual velocity second moment $(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{v}-\boldsymbol{w})^T(t_i)$ at all spatial locations. This simple scheme thus provides a representation of a spatially varying stationary variance tensor.

Small-scale variance tensor in the chronos subspace

A stationary model has obvious limitations in terms of turbulence intermittency modelling. A time dependent variance tensor is nevertheless more involved to estimate as in this case only a single realization of the small-scale velocity trajectory, $(\boldsymbol{v}(\boldsymbol{x},t)-\boldsymbol{w}(\boldsymbol{x},t))$, is available. However considering a temporal basis it is possible to estimate, at a fixed point, the matrix coefficients, $\boldsymbol{z}_i(\boldsymbol{x})$, of the tensor, $\boldsymbol{a}(\boldsymbol{x},t)$ (Genon-Catalot et al., 1992). We term those coefficients the *diffusion modes*, as they correspond to a modal decomposition of the principal diffusion directions. With the *chronos* reduced basis we get:

$$\tilde{\boldsymbol{a}}(\boldsymbol{x},t) = \sum_{j=0}^{n} b_j(t)\boldsymbol{z}_j(\boldsymbol{x}). \tag{7.3.10}$$

Note that even though the residual velocity, $(\boldsymbol{v}-\boldsymbol{w})$, lives in the subspace orthogonal to the *chronos* reduced basis, its one-point one-time covariance – and hence the variance tensor – do not. So, it seems natural to introduce the decomposition (7.3.10). Using the orthogonality of the *Chronos*, $\int_0^T b_k b_l \mathrm{d}t = \delta_{kl}\lambda_k T$, leads to:

$$\boldsymbol{z}_j(\boldsymbol{x}) = \int_0^T \frac{b_j(t)}{T\lambda_j}\tilde{\boldsymbol{a}}(\boldsymbol{x},t)\mathrm{d}t \approx \frac{\Delta t}{N+1}\sum_{k=0}^{N}\frac{b_j(t_k)}{\lambda_j}(\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{x},t_k)((\boldsymbol{v}-\boldsymbol{w})(\boldsymbol{x},t_k))^T, \tag{7.3.11}$$

where $N = \frac{T}{\Delta t}$. It can be noticed that keeping only the zero-diffusion mode and canceling the others: $z_i = 0,\ \forall i \geqslant 1$, brings us back to the stationary variance tensor model. The non-zero modes introduce a time varying variance. Yet, it is important to outline that the reduced order model (7.3.4) remains a quadratic autonomous system. As a matter of fact from (7.3.10), we get the following system:

$$\frac{\mathrm{d}b_i}{\mathrm{d}t} = i_i + \boldsymbol{l}_{\bullet i}^T \boldsymbol{b} + \boldsymbol{b}^T \left(\boldsymbol{c}_{\bullet\bullet i} + \boldsymbol{f}_{\bullet\bullet i}\right) \boldsymbol{b}, \text{ where } f_{pqi} \triangleq \frac{\Delta t_i}{\Delta t} \breve{f}_{qi}\left(\boldsymbol{z}_p\right). \tag{7.3.12}$$

7.4 Flow configuration and numerical simulations

To evaluate the pertinence of the modelling developed in the previous section for the specification of a low order dynamical system and to analyse the contribution of the small-scale component, we consider three-dimensional incompressible flows past a circular cylinder at Reynolds number $Re = 300$ and $Re = 3900$. We performed direct numerical simulations (DNS) using Incompact3d, a high-order flow solver, based on the discretization of the incompressible Navier–Stokes equations with finite-difference sixth-order schemes on a cartesian mesh (Laizet and Lamballais, 2009). A second-order Adams-Bashforth scheme was used for the time advancement. The incompressibility condition is treated with a fractional step method based on the resolution of a Poisson equation in spectral space, allowing here for the velocity field the use of periodic boundary conditions in the two lateral directions y and z. A constant flow is imposed at the inlet of the computational domain and a simple convection equation is solved at the exit. Using the concept of the modified wavenumber, the divergence free condition is ensured up to machine accuracy. The pressure field is staggered from the velocity field by half a mesh to avoid spurious oscillations. The modelling of circular cylinder of diameter D inside the computational domain was performed here with a simple Immersed Boundary Method (IBM). It is based on a direct forcing to ensure zero velocities boundary condition at the wall and inside the solid body. We also used data extracted from a large-eddy numerical simulation (LES) performed by Parnaudeau et al. (2008) using a former version of Incompact3d. This code solved incompressible Navier–Stokes equations on a regular cartesian grid in nonstaggered configuration and using the customized IBM technique of Parnaudeau et al. (2003) to avoid discontinuities on the velocity field, leading to the creation of spurious oscillations when high-order centered schemes are used. This LES was carried out with a high spatial resolution, close to the one expected for a DNS, and with the subgrid-scale model proposed by Métais and Lesieur (1992) combined with a fixed filter length estimated as the cubic root of the mesh volume.

Three different cases were considered: a DNS at Reynolds number $Re = 300$, a high resolution LES (HR LES) at Reynolds number $Re = 3900$ (Parnaudeau et al., 2008) and a low resolution DNS (LR DNS) at Reynolds number $Re = 3900$. We performed the DNS at Reynolds number $Re = 300$ on a domain extending over $20D \times 20D \times 0.5D$ with $512 \times 321 \times 8$ points in the streamwise, perpendicular and spanwise directions, respectively. This reduced spanwise length corresponds to the minimum domain size usable with Incompact3d and led to a three dimensional wake flow simulation with a very short periodicity in the spanwise direction. This choice was done to reduce the computational cost and to simulate a longer time series, necessary for the POD analysis. The simulation was run on $N = 3999$ time steps to observe 80 vortex shedding cycles. The high resolution LES of Parnaudeau et al. (2008) we used was computed on a domain size of $20D \times 20D \times \pi D$ with $960 \times 960 \times 48$ points in the streamwise, perpendicular and spanwise directions, respectively. We extracted from this simulation 219 equidistant snapshots over 5 vortex shedding cycles. In order to have longer time series we performed a low resolution DNS for the same Reynolds number $Re = 3900$ and with the same domain size and mesh grid. This DNS was run on 2499 time steps to observe 80 vortex shedding, with the last parallel version of Incompact3d (Laizet and Li, 2011). Since the new version of Incompact3d did not implement the subgrid-scale model used in the study of Parnaudeau et al. (2008), we performed a low-resolution DNS for the same Reynolds number. The main difference in the results between HR LES and LR DNS are discussed below.

In the following (with the exception of §7.5.1), non-dimensional quantities are considered, and calculated using the cylinder diameter D and the inflow velocity U_0. Dimensionless quantities will be identified by lower-case symbols, e.g. (x, y, z) for the coordinate system and t for the time. In

Case	Re	$(L_x \times L_y \times L_z)/D$	$n_x \times n_y \times n_z$	Snapshots	Shedding cycles
DNS	300	$20 \times 20 \times 0.5$	$512 \times 321 \times 8$	3999	80
HR LES	3900	$20 \times 20 \times \pi$	$961 \times 960 \times 48$	219	5
LR DNS	3900	$20 \times 20 \times \pi$	$961 \times 960 \times 48$	2499	80

Table 7.1: Summary of simulations and of extracted data.

this frame reference, the inflow velocity vector at $x = 0$ is $(u, v, w) = (1, 0, 0)$ and the cylinder is located at $(x, y, z) = (5, 0, z)$. Details of the three cases are provided in table 7.1. Figure 7.1 shows the spanwise vorticity component in the plane $z = 0$ for the DNS at Reynolds number $Re = 300$, the high resolution LES at Reynolds number $Re = 3900$ and the low resolution DNS at Reynolds number $Re = 3900$, respectively. At Reynolds number $Re = 300$ and due to the quasi two-dimensional configuration of the simulation, there are only few small-scale features. Most of the energy is gathered in the large-scale vortical structures. In this regime two topos modes are sufficient to describe faithfully the flow. At Reynolds number $Re = 3900$, a sustained turbulence can be observed in the far wake of the cylinder and in the recirculation zone just behind the cylinder. The boundary layer on the body is laminar and transition to turbulence takes place in the shear layers. The near wake flow is mostly driven by those two shear layers (Ma et al., 2000). Their oscillations trigger the Von Karman vortex shedding and determine the size of the recirculation area. For this wake flow regime a higher number of modes must be kept. From figure 7.1 we can compare the predictions of both simulations at Reynolds number $Re = 3900$. It is worth noting that HR LES exhibits a statistical behaviour consistent with experimental measurements (Parnaudeau et al., 2008), especially in the very near wake for the prediction of the length of recirculation zone. On the contrary, LR DNS yields shorter formation length close to the one obtained at higher Reynolds number, i.e. for $Re \sim 10000$. This behaviour of low resolution wake flow simulations is classical and associated with the so called V-shape of the mean streamwise velocity profiles for $x \leq 8$ (Ma et al., 2000). Hence, both simulations provide distinct vorticity distributions in the very near wake. Farther downstream, i.e. for $x \geq 8$, such pronounced differences are less noticeable. LR DNS exhibits some wiggles, which seem to increase sligthly the vorticity spreading along the wake.

7.5 Diffusion modes results

We now apply the novel POD modelling based on a stochastic small-scale representation presented in §7.3 to the cylinder flow configuration described in §7.4. The reduced order dynamics of the cylinder wake flow is known to be sensitive to the unresolved small-scale velocity component. In §7.5.1, the small-scale energy and anisotropy is linked to small-scale diffusion modes. Assessment of the stochastic modelling is then performed in the following at the Reynolds number $Re = 300$ and $Re = 3900$, respectively. In §7.5.2, contributions of small-scale diffusion modes to large-scale flow are described and interpreted to determine which physical mechanisms of the wake flow are concerned. In §7.5.3, we assess the performance of the subgrid term that was introduced by the stochastic representation of the small-scales by comparing the *chronos* trajectories to the reference.

7.5.1 Estimation and decomposition of the turbulent velocity components

It can be noticed in the decomposition (7.3.10), that the diffusion modes $\boldsymbol{z}_i(x)$ are $d \times d$ symmetric matrices (with $\boldsymbol{z}_i = 0$ for $i > 0$ in the stationary case) at all spatial points. They can be diagonalized in a local orthonormal basis. Let us note however this decomposition does not ensure that $\boldsymbol{a}(x, t) = \sum_{k=0}^{n} b_k(t)\boldsymbol{z}_k(x)$ is positive definite since $b_i(t)$ takes positive and negative values. In practice though the stationary coefficient dominates largely the others (which gives a positive definite estimation), it would be necessary to project the variance tensor on the manifold of positive definite matrices. In the following section, to analyse the small-scale energy and anisotropy we visualize the absolute values of the eigenvalues associated with the matrix modes, $\boldsymbol{z}_i$. Due to *chronos* normalization, the

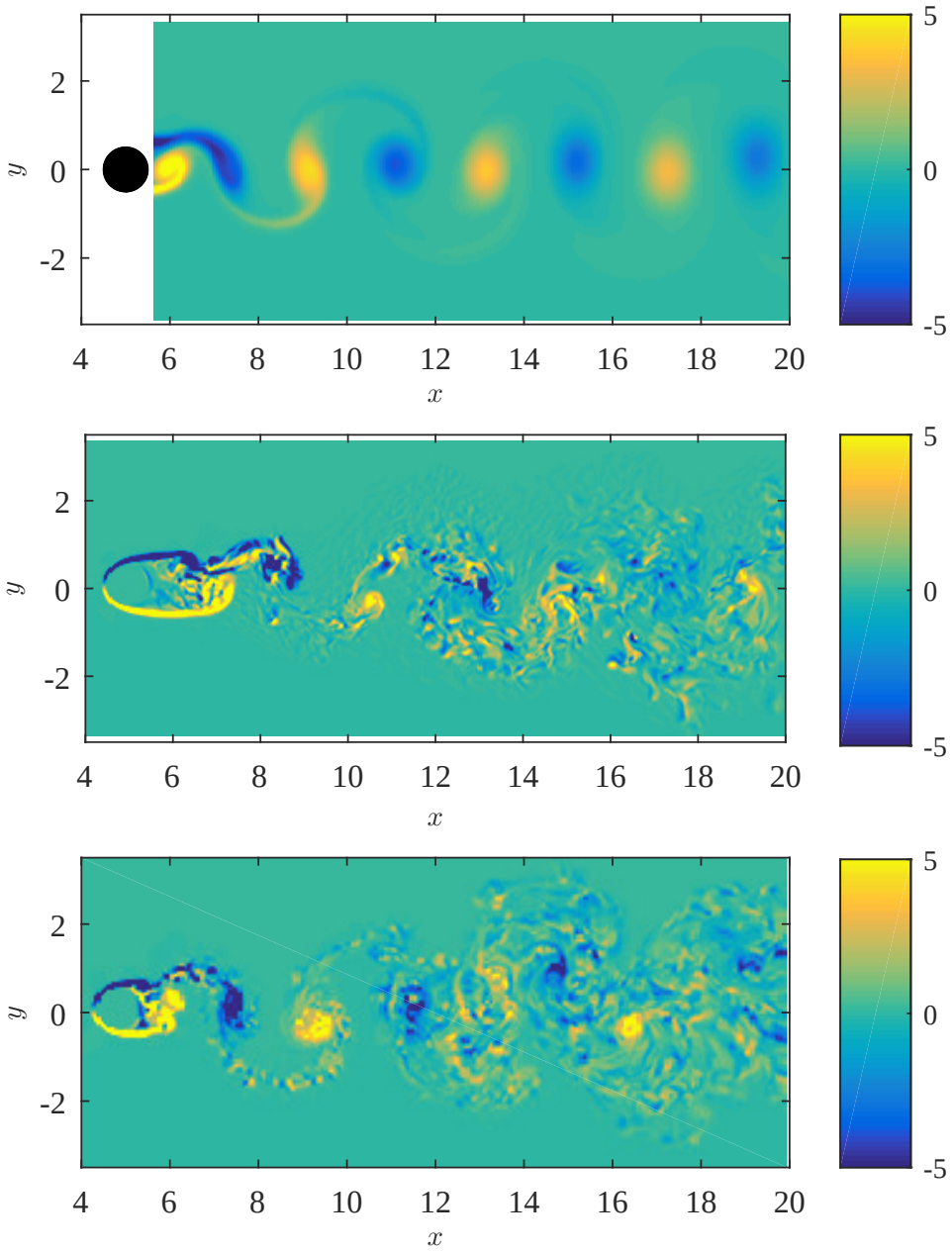


Figure 7.1: Spanwise vorticity componant in a circular cylinder wake flow at Reynolds number, from top to bottom, $Re = 300$ (DNS), $Re = 3900$ (high resolution LES, Parnaudeau et al. (2008)) and $Re = 3900$ (low resolution DNS), respectively.

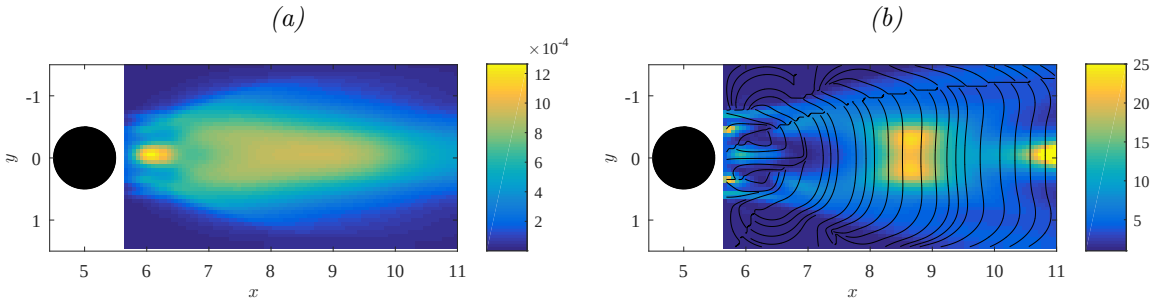


Figure 7.2: Local spectral representations of the matrix $\boldsymbol{a}$ in a cylinder wake flow at $Re = 300$, for $n = 2$ POD modes and in the plane $z = 0$: *(a)*, Turbulent kinetic energy of the diffusion modes $\boldsymbol{z}_0$ (square root of the sum of the squared eigenvalues); *(b)*, Small-scale anisotropy of the diffusion modes $\boldsymbol{z}_0$ (ratio of the absolute value of the largest eigenvalue to the absolute value of the smallest eigenvalue). The streamlines represent the first proper diffusion direction (i.e. the eigenvectors corresponding to the largest eigenvalues of the small-scale variance tensor).

variance tensor diffusion modes must be also normalized by the *chronos*' square root eigenvalues $\sqrt{\lambda_i}$, as:

$$\tilde{\boldsymbol{a}}(x,t) = \sum_{k=0}^{n} \frac{b_k(t)}{\sqrt{\lambda_k}} \left(\sqrt{\lambda_k} z_k(x)\right) \text{ and } \frac{1}{T}\int_0^T \left(\frac{b_k}{\sqrt{\lambda_k}}\right)^2 = 1. \tag{7.5.1}$$

We note this normalization put an even stronger emphasis on the stationary dissipation zero-mode. Then, by (7.2.4) the corrective drift reads

$$\boldsymbol{w}^\star - \boldsymbol{w} = \sum_{k=0}^{n} \frac{b_k(t)}{\sqrt{\lambda_k}} \, \boldsymbol{v}_c^{(k)}(\boldsymbol{x}) \text{ with } \boldsymbol{v}_c^{(k)} = -\frac{1}{2}\boldsymbol{\nabla}\cdot\left(\sqrt{\lambda_k}\boldsymbol{z}_k\right)^T. \tag{7.5.2}$$

Before dealing with *chronos* reconstruction, we propose in §7.5.1 and 7.5.2 a new type of POD data analysis involving the information contained in the residual velocity. Algorithm 1 summarizes the steps of our data analysis, including the POD and the diffusion modes computation.

7.5.2 Small-scale energy density, stationarity and anisotropy

The turbulent kinetic energy density (TKE) was computed by the sum of the diffusion modes eigenvalues, since small-scale TKE is represented (up to a time scale) by the norm of that tensor. The bigger the TKE, the more important the diffusion is. The diffusion zero-mode energy is plotted in figure 7.2 *(a)* for the circular cylinder wake flow at $Re = 300$, with two POD modes and in the plane $z = 0$. The other diffusion modes are much weaker (not shown) and the small-scale velocities component is thus almost stationary. The colormap yields regions of high TKE in the recirculation zone and in an arrow shape just downstream in the transitional region, i.e. for $7 \leq x \leq 10$.

To measure small-scale anisotropy, we computed the ratio between largest and smallest eigenvalues, corresponding to the condition number of the local small-scale velocity variances. The bigger this quantity the more aligned toward the first local proper direction the small-scale velocity is, i.e. the more anisotropic the small-scale velocity and the diffusion are. Figure 7.2 *(b)* shows the colormap of this quantity for the circular cylinder wake flow at $Re = 300$. In regions where the unresolved velocity component is largely anisotropic, the small-scale velocity is mainly directed in the direction of the eigenvector associated with the largest eigenvalue of the small-scale variance tensor. The small-scale component imposes a diffusion of the resolved velocity in the same direction. The streamlines in figure 7.2 *(b)* shows the principal local diffusion directions defined by the largest eigenvectors. The orthogonal to the streamlines would depict the directions of least diffusion of the large-scale velocity by the small-scale component. These directions can be interpreted as the ones of least small-scale uncertainty. The streamlines clearly show the vortex formation region with the symmetric vortex rolling zone. The two pivotal locations at $y = \pm 0.5$ just before station $x = 6$ where both shear layers start to roll into vortices are precisely indicated by high values of

Algorithm 1 POD and diffusion modes data analysis

function STOCHASTIC_POD_ANALYSIS($n, \boldsymbol{v}(\boldsymbol{x}, t_0), ..., \boldsymbol{v}(\boldsymbol{x}, t_N)$))

1. Usual POD: Resolved velocity component

$$\boldsymbol{w}(\boldsymbol{x}, t) = \sum_{i=0}^{n} b_i(t)\boldsymbol{\phi}_i(\boldsymbol{x}). \tag{7.5.3}$$

2. Optimal time step

$$\frac{1}{\Delta t} = 4 \max_{i \leqslant n} \left(f_{max}\left(b_i\right)\right). \tag{7.5.4}$$

3. Diffusion modes analysis: study of the residual velocity component influence

- Residual velocity component

$$\boldsymbol{v} - \boldsymbol{w}. \tag{7.5.5}$$

Decomposition of the residual velocity influence

- **for** $j = 0$ to n **do**

Component of the residual velocity influence associated with the time variability of the *chronos* b_j (note that $b_0 = \lambda_0 = 1$)

  - Diffusion mode computation

    Projection of the squared residues on the resolved *chronos* b_j

$$\boldsymbol{z}_j(\boldsymbol{x}) = \frac{\Delta t}{N+1} \sum_{k=0}^{N} \frac{b_j(t_k)}{\lambda_j} (\boldsymbol{v} - \boldsymbol{w})(\boldsymbol{x}, t_k)((\boldsymbol{v} - \boldsymbol{w})(\boldsymbol{x}, t_k))^T. \tag{7.5.6}$$

  - Analysis of the diffusion of the resolved velocity $\boldsymbol{w}$ by the residual velocity
    - Local diagonalization of the symmetric matrix $\boldsymbol{z}_i(\boldsymbol{x})$

$$\sqrt{\lambda_j}\boldsymbol{z}_j(\boldsymbol{x}) = \boldsymbol{P}_j(\boldsymbol{x})\boldsymbol{\Lambda}^{(j)}(\boldsymbol{x})\boldsymbol{P}_j^T(\boldsymbol{x}), \tag{7.5.7}$$

with

$$\boldsymbol{P}_j(\boldsymbol{x})\boldsymbol{P}_j(\boldsymbol{x})^T = \boldsymbol{P}_j(\boldsymbol{x})^T\boldsymbol{P}_j(\boldsymbol{x}) = \mathbb{I}_d \text{ and } \Lambda_{pq}^{(j)}(\boldsymbol{x}) = \delta_{pq}\Lambda_{pp}^{(j)}(\boldsymbol{x}). \tag{7.5.8}$$

    - Inhomogeneity of the turbulent diffusion of the resolved velocity (proportional to the small-scale kinetic energy)

$$\sum_{p=1}^{d} \left|\Lambda_{pp}^{(j)}(\boldsymbol{x})\right|. \tag{7.5.9}$$

    - Anisotropy of the turbulent diffusion of the resolved velocity (equal to the anisotropy of the small-scale kinetic energy)

$$\frac{\max_p \left|\Lambda_{pp}^{(j)}(\boldsymbol{x})\right|}{\min_p \left|\Lambda_{pp}^{(j)}(\boldsymbol{x})\right|}. \tag{7.5.10}$$

  - Corrective drift

$$\boldsymbol{v}_c^{(j)}(\boldsymbol{x}) = -\frac{1}{2}\boldsymbol{\nabla} \cdot \left(\sqrt{\lambda_j}\boldsymbol{z}_j\right)^T(\boldsymbol{x}). \tag{7.5.11}$$

    - Vorticity of the corrective drift

$$\boldsymbol{\nabla} \times \boldsymbol{v}_c^{(j)}. \tag{7.5.12}$$

    - Divergence of the corrective drift

$$\boldsymbol{\nabla} \cdot \boldsymbol{v}_c^{(j)}. \tag{7.5.13}$$

- **end for**

end function

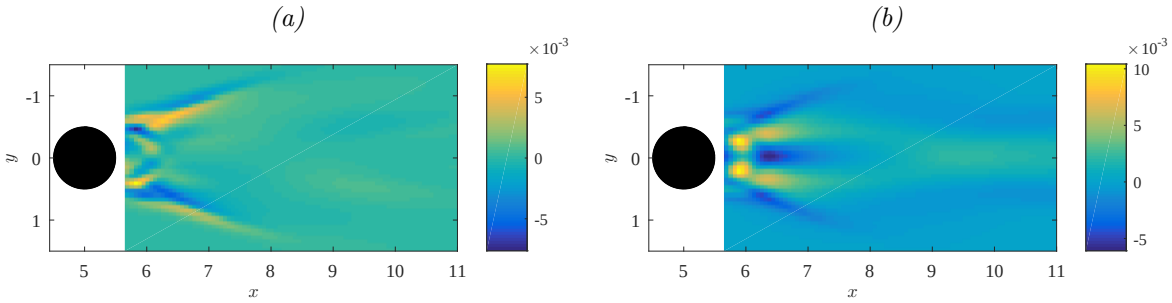


Figure 7.3: Spanwise vorticity *(a)* and divergence *(b)* of the drift correction $-\frac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)^T$, for the diffusion zero-mode $\boldsymbol{z}_0$ and $n = 2$ POD modes, in a cylinder wake flow at $Re = 300$.

the small-scale anisotropy. Downstream, just before station $x = 9$ in the middle of the transitional region, the anisotropy reaches high values in an approximately square area of length D. To this maximum corresponds the zone where vortices detach from the cylinder and reorganise into the Karman vortex street further downstream. Near station $x \approx 11$ at the beginning of this pattern of regular and aligned vortices, small-scale velocity anisotropy is maximum on the centreline.

Another interesting feature of the small-scales stochastic representation principle concerns the emergence of the small-scale effective velocity (7.2.4), also called drift correction, related to the variance tensor inhomogeneity. Though at $Re = 300$, this contribution is weak as the flow is well captured with only two POD modes, it is nevertheless interesting to observed the velocity component that is induced by the neglected POD modes. In figure 7.8 we plot the vorticity and divergence of this advection correction term (7.2.4) for the diffusion zero-mode $\boldsymbol{z}_0$. The small-scale vorticity induced by the neglected modes is 3 order of magnitude weaker than the whole flow vorticity (figure 7.1), which confirms its minor effect on the large-scale flow. However some interesting patterns emerge from those figures. In the divergence map (figure 7.3 *b*) we observe high divergence zones at station $x = 6$ corresponding to the vortex formation and just downstream at station $x = 6.5$ where the vortices are shed. Convergence zones are also shown at the same station but on the centreline and on both sides just outside of the recirculation region. In the vorticity map (figure 7.3 *a*) several small-scale vorticity spots take place within the recirculation area, whereas elongated vorticity patterns can be observed, like for the convergence, just outside of the recirculation region. Such flow corrections, though being weak, take place in the region of the flow where physical mechanisms that give rise to vortex shedding are active and may have significant contributions if the flow is sensitive in these regions. One interesting feature, here, is the presence of high values of vorticity, corresponding to the maximum of anisotropy, at the two pivotal locations of the shear layers rolling into vortices.

Then it is of particular interest to analyse how the proposed small-scale stochastic modelling behaves with a higher turbulent wake flow. We then consider the cylinder wake flow at Reynolds number $Re = 3900$ (HR LES). It should be noted that the low number of velocity snapshots available for that case (see table 7.1) leads to insufficient convergence of the POD modes. To avoid to much wiggles on the diffusion modes results the matrix $\boldsymbol{a}$ was computed only in two-dimensions in fixed z-planes and then averaged along the z-axis. Hence the contribution of the small-scale spanwise velocity component was neglected and the analysis was based on an averaged view of $\boldsymbol{a}$ along the z-axis. Colormaps of the energy density for the diffusion modes $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are plotted in figure 7.4 *(a)*, for 4 POD modes. We observe that the turbulent energy of the diffusion zero-mode is three times larger than for the nonstationary modes. These spatial small-scale energy distributions show that the largest magnitudes are reached at the end of the recirculation region and further downstream in the transitional region. Examination of the small-scale anisotropy spatial distribution together with the streamlines of the principal local diffusion directions plotted in figure 7.4 *(b)*, indicates for $\boldsymbol{z}_0$ diffusion mode the two pivotal regions where the shear layers start to roll into vortices, and for $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ diffusion modes more complex structures of the anisotropy high values in the whole wake flow. In figures 7.5 *(a)* and 7.5 *(b)*, respectively the vorticity norm and the divergence of the drift correction for the diffusion mode $\boldsymbol{z}_0$ is plotted. Like for the $Re = 300$ case,

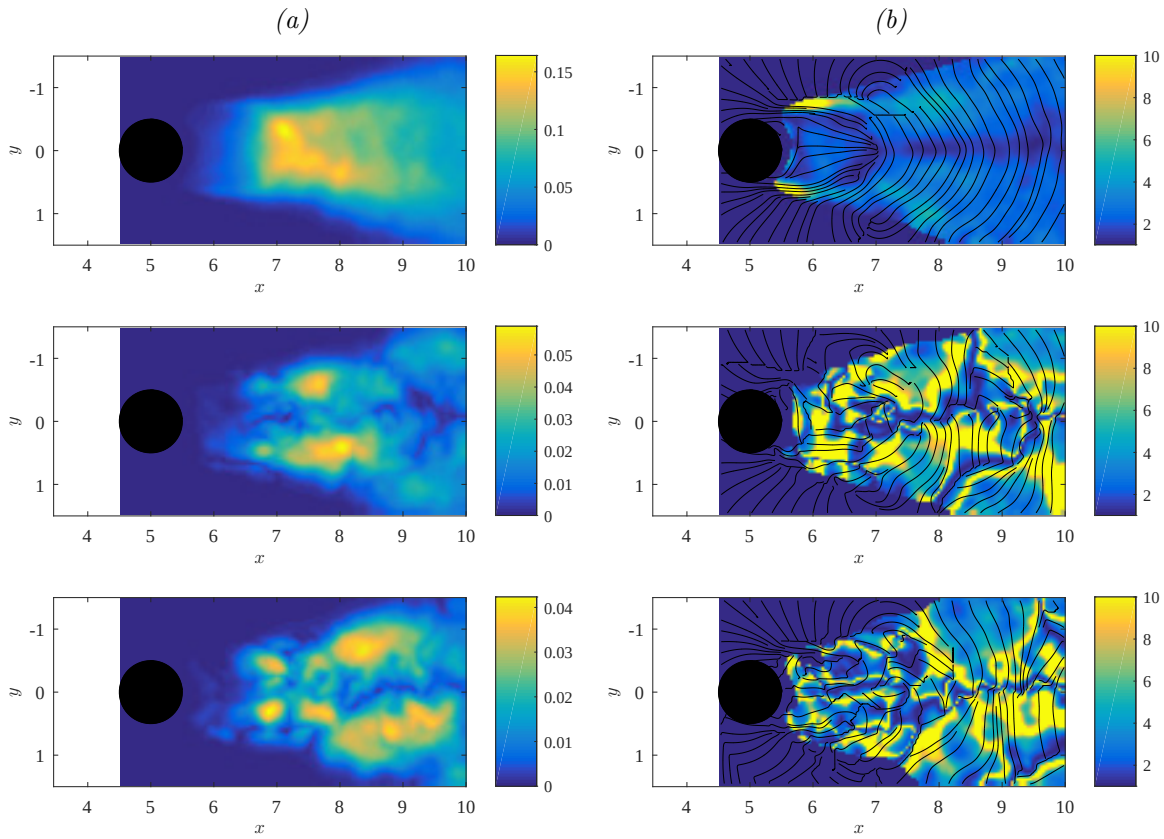


Figure 7.4: Local spectral representation of the matrix $\boldsymbol{a}$ in a cylinder wake flow at $Re = 3900$ (HR LES), for $n = 4$ POD modes: From top to bottom, diffusion modes $\boldsymbol{z}_0$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ respectively; *(a)*, Turbulent kinetic energy (square root of the sum of the squared eigenvalues); *(b)*, Small-scale anisotropy (ratio of the absolute value of the largest eigenvalue to the absolute value of the smallest eigenvalue). The streamlines represent the first proper diffusion direction (i.e. the eigenvectors corresponding to the largest eigenvalues of the small-scale variance tensor).

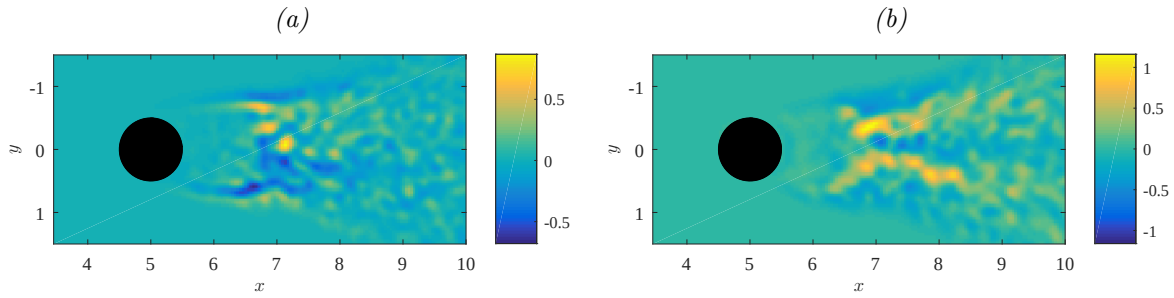


Figure 7.5: Spanwise vorticity component *(a)* and divergence *(b)* of the drift correction $-\frac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)^T$ due to the diffusion mode $\boldsymbol{z}_0$, in a cylinder wake flow at $Re = 3900$ (HR LES) for $n = 4$ POD modes.

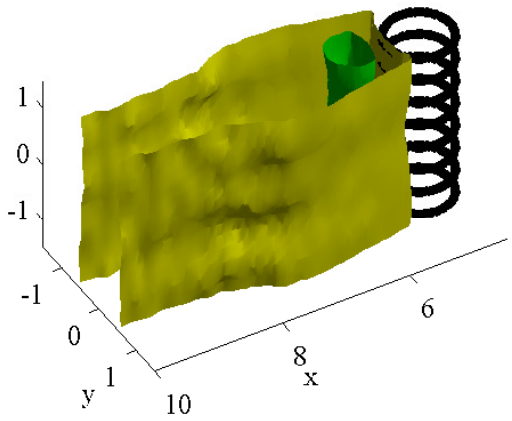
vorticity and divergence corrections are observed near the shear layers, outside and at the end of the recirculation zone, and just downstream in the launching area. Here again, high magnitudes of vorticity just downstream the maximum of anisotropy, are associated with the two pivotal regions of the shear layers rolling into vortices. Nevertheless, in the present HR LES case, results show some wiggles, which are here the footprint of insufficient convergence of the modes due to the low number of velocity snapshots available (see table 7.1). In that context the velocity induced by the neglected POD modes is less informative since the energy is less precisely spread over the modes.

We now consider the case LR DNS with sufficient velocity snapshots to get converged POD modes thus allowing us to analyse more precisely in three-dimensions the behaviour of the diffusion modes. Note however that due to the marginal resolution of LR DNS the flow yields a short recirculation zone length of the order $0.5D$, lower than the distance of $1.56D$ obtained with HR LES and expected for the Reynolds number $Re = 3900$ (Parnaudeau et al., 2008). Figure 7.6 *(a-d)* is a mapping of three-dimensional iso-surfaces of the energy density for the diffusion modes $\boldsymbol{z}_0$, $\boldsymbol{z}_1$, $\boldsymbol{z}_2$, $\boldsymbol{z}_3$, and $\boldsymbol{z}_4$, in a cylinder wake flow at $Re = 3900$ (LR DNS). We observe that the diffusion mode $\boldsymbol{z}_0$ is one order of magnitude larger than the nonstationary modes. The flow turbulence exhibits hence a strong stationary behaviour. On the stationary diffusion mode $\boldsymbol{z}_0$ we see clearly the recirculation zone and the external delimitation of the cylinder wake as regions of higher and lower dissipation, respectively. The second and third nonstationary diffusion modes are related to structures with a characteristic length of the size of the Von Karman vortices. The fourth and fifth nonstationary diffusion modes correspond to much finer structures in the near cylinder wake. When compared with the case at $Re = 300$, the stationary diffusion mode $\boldsymbol{z}_0$ exhibits similar spatial distribution of the small-scale turbulent kinetic energy, with the highest magnitudes reached at the end of the recirculation region and further downstream in the transitional zone with an arrow shape. Where the small-scale energy is high, the unresolved velocity and the diffusion are strong. The results we present next analyse the three-dimensionnal anisotropy of the diffusion modes.

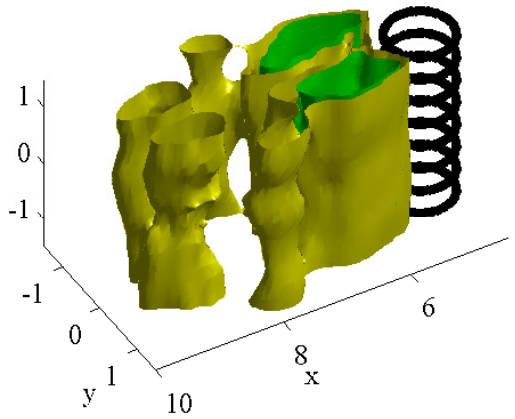
In figures 7.7 *(a-c)* we plot the small-scale anisotropy isosurfaces of modes $\boldsymbol{z}_0$ and $\boldsymbol{z}_1$, respectively. No particular structure shows up on the nonstationary first modes (see e.g. anisotropy of mode $\boldsymbol{z}_1$ in figure 7.7 *b*). As a consequence, the anisotropic structures and consequently the diffusion geometry is principally described by the diffusion mode $\boldsymbol{z}_0$. Like for the two other cases, the diffusion mode $\boldsymbol{z}_0$ exhibits clearly, through local high magnitudes of the small-scale anisotropy, the two three-dimensional pivotal locations of the shear layers close to the cylinder just before station $x = 5.5$ at $y = \pm 0.5$ (figures 7.7 *a* and *c*). Downstream near station $x = 8 - 9$ in the middle of the transitional region where vortices detach from the cylinder and start to reorganise into the Karman vortex street, local high magnitudes of the anisotropy is indicated in figure 7.7 *(c)*. Farther downstream, near station $x \approx 12$ at the beginning of this pattern of regular and aligned vortices, the diffusion mode anisotropy is also maximum on the centreline.

In figure 7.8 *(a-e)* we plot vorticity and divergence iso-surfaces of the advection correction term for the three first diffusion modes. The stationary diffusion mode $\boldsymbol{z}_0$ unveils two large-scale effective contra-rotative vortices (figure 7.8 *a*). Like for the two other cases, the drift correction of the resolved (large-scale) velocity by the unresolved (small-scale) velocity is associated with high small-scale anisotropy and with vorticity corrections at the pivotal locations of the shear layers. Just behind the cylinder in the recirculation zone figure 7.8 *(b)* reveals a divergent area and two flat convergent zones near both shear layers. Thinner vorticity and diverging structures can be observed in the near cylinder wake on the nonstationary modes $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. The vorticity and divergence magnitudes involved in those modes are one order weaker than for the stationary diffusion mode. Downstream no significant structures are observed, indicating rather an homogeneous character of the small-scale velocity.

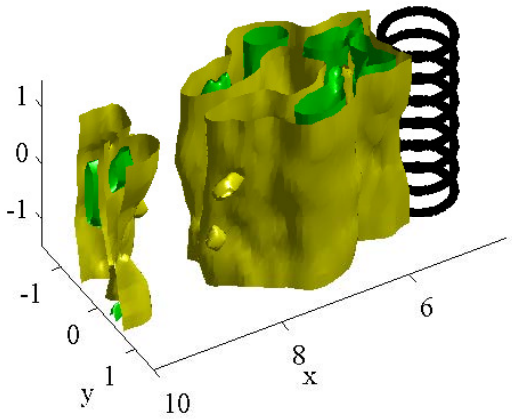
The diffusion modes analysis developed in the present paper identifies critical regions of the wake flow: the anisotropy exhibits mainly the pivotal location of the shear layers which are associated with large-scale vorticity corrections by the small-scale unresolved velocity and large-scale divergence corrections take also place in the vortex formation zone. For the wake flow considered, the results indicate that outside the recirculation area and the shear layer zones an eddy viscosity assumption is likely valid. However in the near wake, such an assumption is too strong and corrective advection effects as well as structured energy dissipation effects must be taken into account. These findings support the recent results of Chandramouli et al. (2016) who demonstrated the good contributions of such novel stochastic small-scale modelling in the context of coarse-grid large eddy



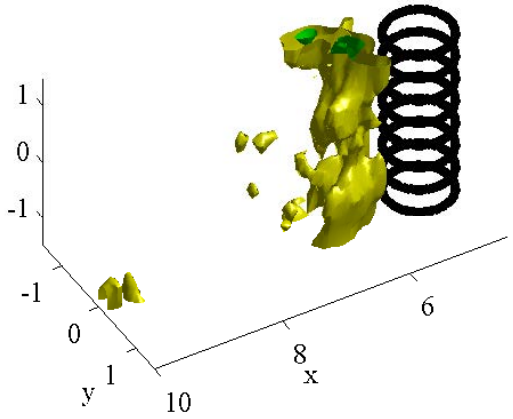
(a) $\sqrt{\lambda_0} tr(\boldsymbol{z}_0) = (0.25; 0.75)$



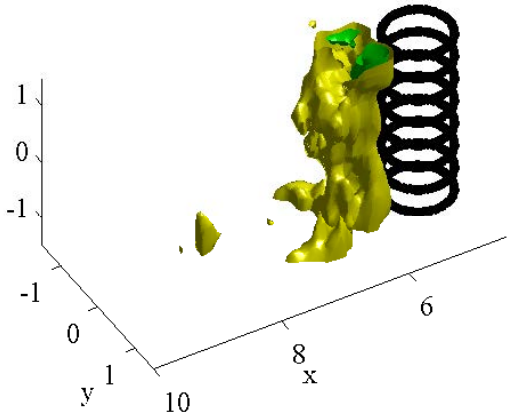
(b) $\sqrt{\lambda_1} tr(\boldsymbol{z}_1) = (0.02; 0.0125)$



(c) $\sqrt{\lambda_2} tr(\boldsymbol{z}_2) = (0.02; 0.0125)$



(d) $\sqrt{\lambda_3} tr(\boldsymbol{z}_3) = (0.012; 0.009)$



(e) $\sqrt{\lambda_4} tr(\boldsymbol{z}_4) = (0.012; 0.009)$

Figure 7.6: Local spectral representation of the matrix $\boldsymbol{a}$ in a cylinder wake flow at $Re = 3900$ (LR DNS), for $n = 4$ POD modes: *(a-e)*, Small-scale turbulent kinetic energy isosurfaces of the diffusion modes $\boldsymbol{z}_0$ to $\boldsymbol{z}_4$, respectively. At places where the energy is high, the unresolved velocity and the diffusion are strong. The green isosurfaces are associated with higher values than the yellow isosurfaces.

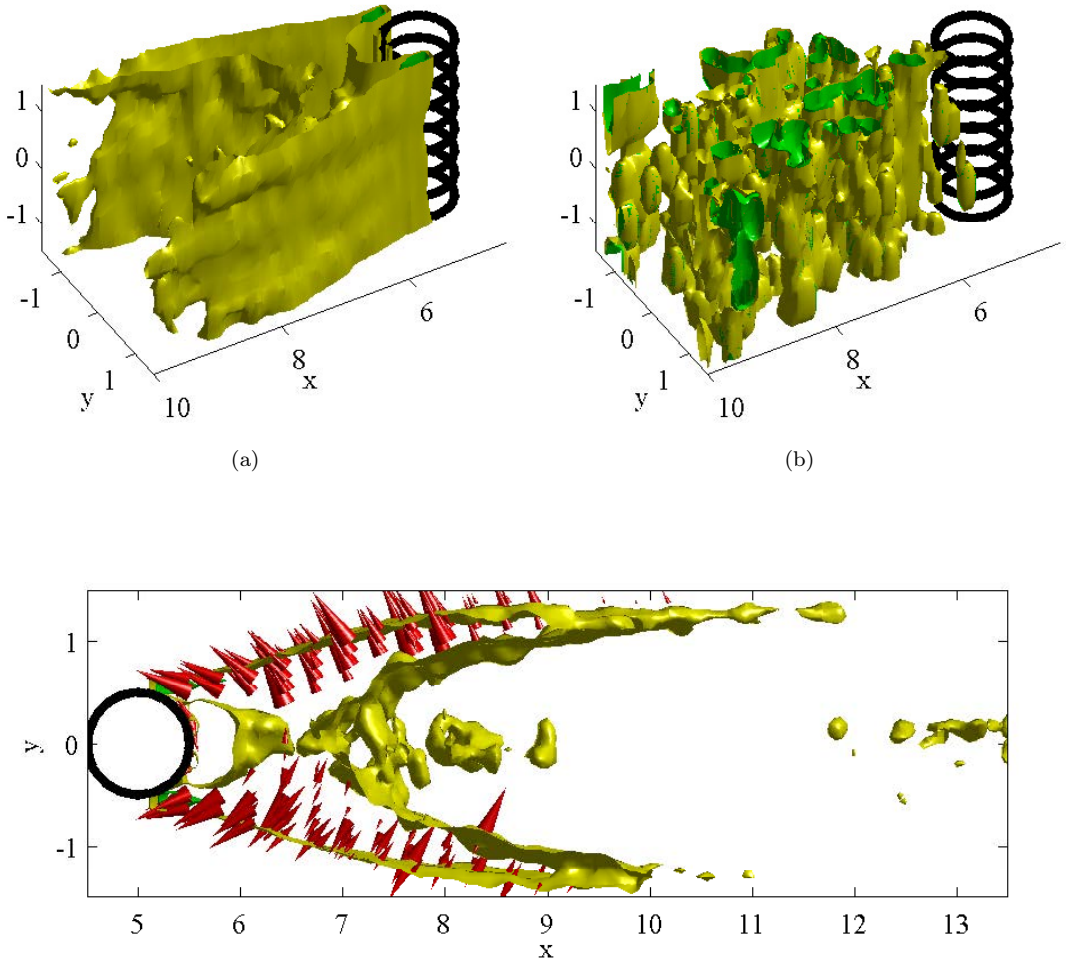
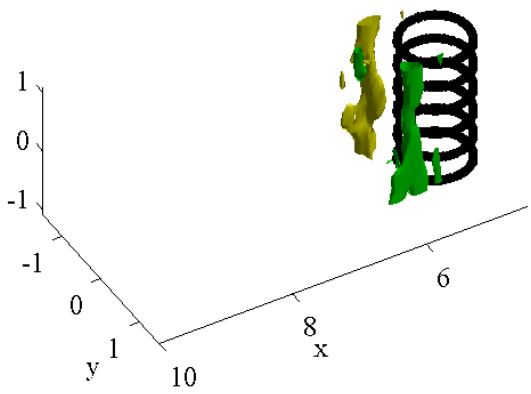
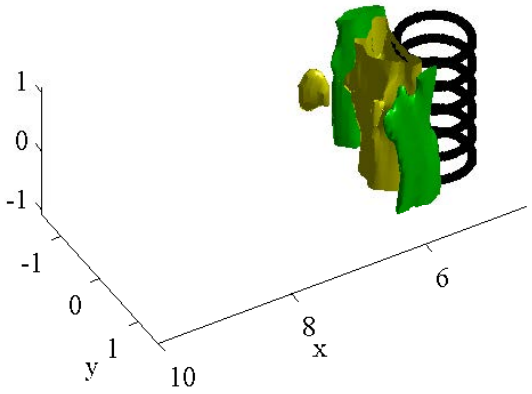


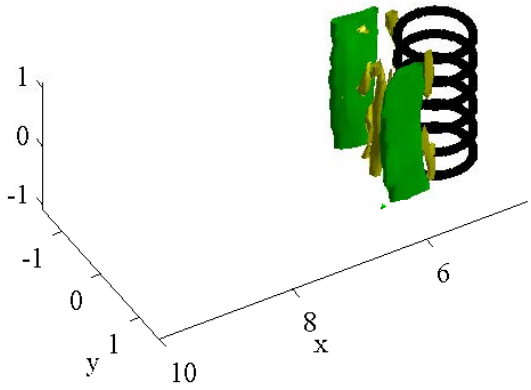
Figure 7.7: Local spectral representation of the matrix $\boldsymbol{a}$ in a cylinder wake flow at $Re = 3900$ (LR DNS), for $n = 4$ POD modes: *(a,b)*, Small-scale anisotropy isosurfaces of modes $\boldsymbol{z}_0$ and $\boldsymbol{z}_1$, respectively; *(c)*, Top view of small-scale anisotropy isosurface of modes $\boldsymbol{z}_0$. The red cones represent the preferential diffusion directions (i.e. the eigenvectors corresponding to the largest eigenvalues of the small-scale variance tensor). The green surface ($\sigma = 6$) is associated with a higher anisotropy of both the small-scale velocity and the diffusion, than the yellow surface ($\sigma = 3$).



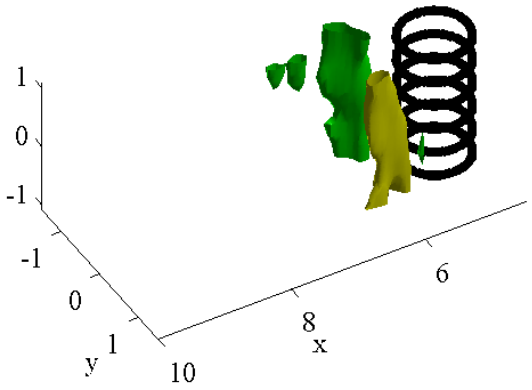
(a) $\frac{1}{2}|\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \cdot \boldsymbol{z}_0)^T| = 0.7$



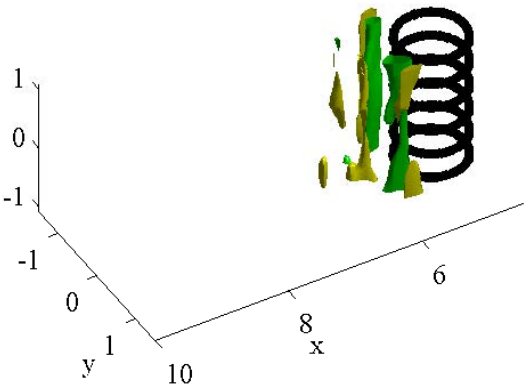
(b) $\frac{1}{2}|\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{z}_0)^T| = 0.4$



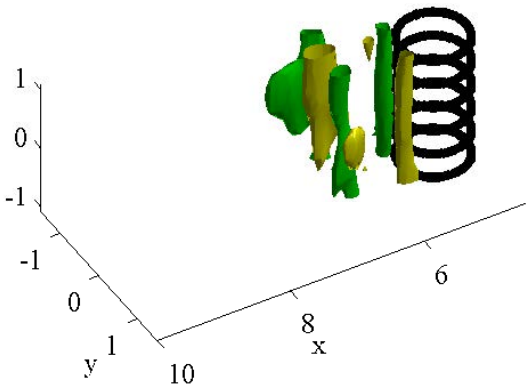
(c) $\frac{1}{2}|\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \cdot \boldsymbol{z}_1)^T| = 0.07$



(d) $\frac{1}{2}|\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{z}_1)^T| = 0.06$



(e) $\frac{1}{2}|\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \cdot \boldsymbol{z}_2)^T| = 0.07$



(f) $\frac{1}{2}|\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \cdot \boldsymbol{z}_2)^T| = 0.06$

Figure 7.8: Vorticity *(Left)* and divergence *(Right)* isosurfaces of the drift correction $-\frac{1}{2}\left(\boldsymbol{\nabla} \cdot \boldsymbol{a}\right)^T$ in a cylinder wake flow at $Re = 300$, for $n = 4$ POD modes: *(a,b)*, Diffusion mode $\boldsymbol{z}_0$; *(c,d)*, Diffusion mode $\boldsymbol{z}_1$; *(e,f)*, Diffusion mode $\boldsymbol{z}_2$. On the left column, the green iso-surfaces are associated with a vorticity vector aligned downward whereas the yellow iso-surfaces are associated with a vorticity vector aligned upward. On the right column the green surfaces stand for areas with iso-negative divergence (convergent zone) whereas the yellow iso-surfaces correspond to a positive divergence (divergent zone).

simulation of a wake flow.

7.5.3 Chronos reconstruction

In this section we aim at assessing the performance of the subgrid term introduced by the stochastic representation of the small-scales. We compare the *chronos* trajectories that were directly reconstructed from the reduced order dynamical system (7.3.12) to the observed *chronos*. Let us note that almost long time perfect trajectories could be recovered through data assimilation strategies (Artana et al., 2012; D'Adamo et al., 2007). However with such techniques it would be difficult to identify the intrinsic role of the subgrid scheme compared to a least-squares adaption of all the dynamics coefficients along the whole sequence. We therefore prefer to rely on a more difficult direct reconstruction strategy in which no coupling mechanism with the data nor any least squares estimation procedure of the dynamical coefficients are introduced.

The results presented so far on the diffusion mode analysis did not necessitate any knowledge of the Reynolds number to compute both the diffusion and the drift correction of the large-scale by the unresolved small-scale. The assessments have been conducted in particular with the LR DNS data for which the flow does not fit with the theoretical wake flow at Reynolds number $Re = 3900$. We now turn to reduced order dynamical systems which, in contrast, need for their evaluation flow data sets that are physically consistent with the Reynolds number considered. Hence in the following we consider the DNS case at Reynolds number $Re = 300$ and the HR LES case at Reynolds number $Re = 3900$.

The modes mean energy, λ_i, and the *topos*, ϕ_i, are computed from the whole sequence of snapshots ($N = 3999$ for Reynolds 300 and $N = 251$ for Reynolds 3900). As for the initial condition, we used the referenced values of the *chronos*, denoted b_i^{ref}, computed directly from the snapshots covariance diagonalization. Then, regarding the *chronos* spectra, an optimal time sub-sampling is chosen, as explained in §7.3.3. Afterward, using the residual velocity and the *chronos*, the variance tensor, a, is estimated. The coefficients of the reduced order dynamical system of *chronos* (7.3.12) are directly computed using discrete derivation schemes. The *chronos* trajectories are simulated with a 4-th order Runge-Kutta integration method, with $b^{ref}(t = 0)$ as initial condition.

Figures 7.9 and 7.10 show examples of the reconstruction of the *chronos* for $n = 2$ at Reynolds number $Re = 300$ and $n = 10$ at Reynolds number $Re = 3900$, respectively, for the classical POD method (blue plot) and for the proposed modelling with respectively a stationary and a nonstationary variance tensor defined on the subspace associated with the *chronos* basis (red plot). At Reynolds number $Re = 300$, the first two modes carry most of the energy. The reference b_i^{ref} (black plot) are superimposed for comparison purposes. It can be observed that our stochastic model follows the references quite well whereas the deterministic model blows up. Let us point out that here both reduced models are completely parameter free. No constant has been tuned to adapt any viscosity model. Figures 7.11 and 7.12 describe the error evolution along time. Approximating the square of the actual unresolved *chronos*, by the time average of their squares, we defined the error as follows:

$$\begin{aligned} err(t) &= T\frac{\left\|u^{ref} - u\right\|_{L^2(\Omega)}}{\left\|u^{ref}\right\|_{L^2(\Omega\times[0,T])}}, \\ &= T\frac{\left\|\sum_{i=1}^{n}\left(b_i^{ref} - b_i\right)\phi_i + \sum_{i=n+1}^{N} b_i^{ref}\phi_i\right\|_{L^2(\Omega)}}{\left\|\sum_{i=0}^{N} b_i^{ref}\phi_i\right\|_{L^2(\Omega\times[0,T])}}, \\ &\approx \left(\frac{\sum_{i=1}^{n}\left(b_i^{ref} - b_i\right)^2 + \sum_{i=n+1}^{N}\lambda_i}{\left\|\phi_0\right\|^2_{L^2(\Omega)} + \sum_{i=1}^{N}\lambda_i}\right)^{1/2}, \end{aligned} \tag{7.5.14}$$

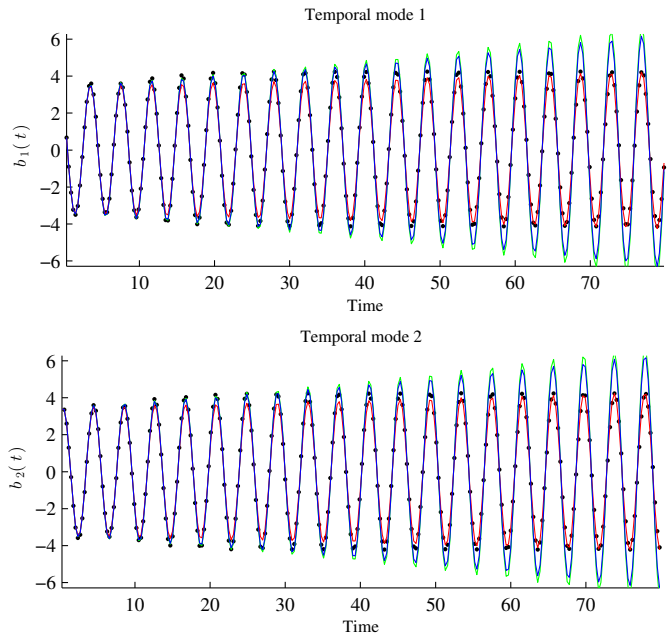


Figure 7.9: Reconstruction of the two first modes ($n = 2$) of a wake flow at Reynolds number $Re = 300$, with a stationary variance tensor. The black dots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation. The initial condition, at $t = 0$, is identical for all methods.

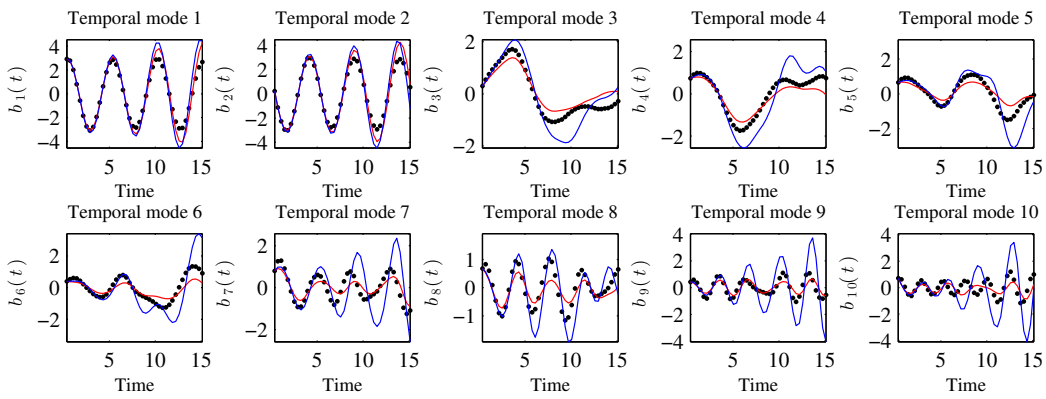


Figure 7.10: Reconstruction of the 10 first modes ($n = 10$) of a wake flow at Reynolds number $Re = 3900$ (HR LES), with a variance tensor expressed as a linear function of the *chronos*. The black dots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the red ones are computed with the stochastic representation. The initial condition, at $t = 0$, is identical for all methods.

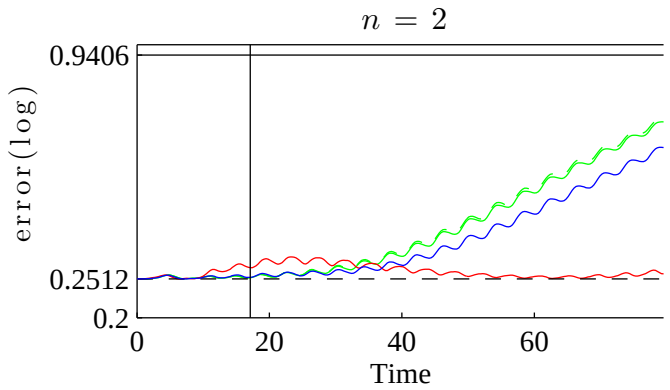


Figure 7.11: Normalized error for $n = 2$ modes, of a wake flow at Reynolds number $Re = 300$. The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue line corresponds to the standard POD Galerkin (without eddy viscosity). The red one stands for our model with a stationary variance tensor. The green colour curve denotes a modal eddy viscosity reduced order model. The modal eddy viscosity coefficients are estimated on the three first vortex shedding. The end of the learning period is indicated by a vertical bar at $t = 18$. The dashed line indicates the error associated with the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error when we only consider the temporal mean velocity.

which is greater than the minimal error associated to the modal truncation:

$$err(t) \quad \geqslant \quad \left(\frac{\sum_{i=n+1}^{N} \lambda_i}{\|\phi_0\|^2_{L^2(\Omega)} + \sum_{i=1}^{N} \lambda_i} \right)^{1/2} . \tag{7.5.15}$$

Equation (7.5.14) defines the criterion error plotted in figures 7.11 and 7.12, whereas (7.5.15) constitutes a lower bound of this error. In figures 7.11 and 7.12, we displayed successively the error plots, obtained for the standard POD Galerkin model without subgrid dissipative term, our model with stationary and nonstationary variance tensors, and finally for a deterministic modal eddy viscosity model. This subgrid model, proposed in Rempfer and Fasel (1994) consists in modifying the reduced order system by adding a strong isotropic diffusive term (Laplacian) to stabilize the system. This eddy viscosity is said to be modal since different viscosity coefficients are attached to each *chronos*. Those coefficients are estimated by a least squares fitting on the first quarter of the data. Modal eddy viscosity in its least squares form ressembles indeed to a data assimilation strategy in which the best stationary isotropic dissipative forcing is estimated from the discrepancy between the model and the data. The same isotropic dissipation is imposed on the whole fluid domain at every time step. As such this subgrid dissipation is much more difficult to interpret in terms of local signatures of the small-scale coherent structures.

In figures 7.11 and 7.12, the doted lines indicates the minimal error associated with the reduced subspace truncation error. The black solid line corresponds to the error level associated with the temporal mean velocity – i.e. setting all the *chronos* to 0. In this case:

$$err_{|b=0}(t) \quad \approx \quad \left(\frac{\sum_{i=1}^{N} \lambda_i}{\|\phi_0\|^2_{L^2(\Omega)} + \sum_{i=1}^{N} \lambda_i} \right)^{1/2} .$$

This term does not constitute an upper bound of the error. However, it provides the error level reached by the null model. In figures 7.11 and 7.12, fixing a log-scale for the y axis, we observe readily the exponential divergence of the standard POD reduced order (in blue).

We observe that at low Reynolds number the modal eddy viscosity model does not allow to capture accurately, on a long time period, the complex non-linear dynamics undergone by the non-resolved modes (figure 7.11). Just after the learning period (3 vortex shedding), the eddy viscosity models performs well. However, they start diverging after 30 time steps. The stationary model leads, in contrast, to a stable simulation in the long run.

For the case HR LES at Reynolds number $Re = 3900$ (figure 7.12), we compared the eddy viscosity approaches (modal and constant) with stationary and nonstationary models of the variance tensors. In this case the variance tensor as well as the eddy viscosity coefficients have been estimated on 3 vortex shedding as well. The performances of the modal and single characteristic times attached to the variance tensor have been evaluated and compared. The error plots are shown in figure 7.12. The introduction of different small-scale characteristic time steps associated with the different modes significantly improves the results that were obtained for a single common characteristic time. Both approaches are equivalent for short time period only. The introduction of modal characteristic times is clearly beneficial in the long run. The nonstationary representation performs only slightly better than the stationary one for this wake flow. However, the piece of information brought by the nonstationary diffusion modes enables a meaningful analysis of the small-scale contribution (see §7.5.2). Both eddy viscosity approaches perform well either in their modal or constant versions. Modal eddy viscosity appears to work better when a small number of modes is involved. The (stationary or nonstationary) variance tensor models that are associated with modal characteristic time scales exhibit nearly the same stabilizing skills as the eddy viscosity models. Both models lead to similar error levels. Nevertheless, it must be outlined that the two approaches are based on different assumptions. Eddy viscosity relies intrinsically on an homogeneous isotropic diffusion with no preferential direction of energy dissipation. The diffusion remains constant whatever the considered region: in the near or far wake regions, and even in the shear layers. However, as a fixed constant estimated through a mean squares procedure, it provides the optimal amount of missing energy dissipation (with respect to a spatio-temporal mean of the squared norm) that is required to stabilize the reduced dynamical system. Conversely, as shown in the previous section, the variance tensor and the associated diffusion modes provide a finer representation of the small-scales action in terms of energy dissipation but also in terms of energy redistribution. As for the simulation of the reduced system, both models lead to comparable error levels. They enable to stabilize the system in a similar way, but the variance tensor models unveil important hints on the small-scale flow structuration.

7.6 Conclusion

We investigated the study of reduced order modelling based on a stochastic representation of the small-scales proposed by Mémin (2014) and Resseguier et al. (2017a). This principle gives rise naturally to a drift correction generated by the inhomogeneity of the small scale velocity variance and to an inhomogeneous diffusion term. The diffusion term is closely related to eddy viscosity assumption. Indeed, for an isotropic divergence-free random field, the stochastic representation boils down to the classical eddy viscosity assumption. A POD Galerkin projection of the corresponding Navier-Stokes equations enables us to constitute a modified reduced-order dynamical system that includes a linear term gathering the effects of the effective advection and the diffusion exerted by the unresolved small-scale component. This function directly depends on the small-scale variance that must be specified to close the system. We proposed in this study a modelling based on the decomposition of this variance tensor on the *chronos* basis. The estimation has been performed on the residuals between the snapshots measurements and their resolved reconstruction on the *topos* basis. The coefficients of this decomposition quoted as the *diffusion modes* constitute meaningful features for the interpretation of the small-scale statistical organization. They allow for a detailed examination of the principal directions of the small-scale energy dissipation and also to extract advective structures generated by the small-scale velocity. For wake flow configurations the analysis of the anisotropy of the diffusion modes is appropriate to determine regions of the flow that are key players. The largest magnitudes of the anisotropy zero-mode occur in the vicinity of the pivotal zone of the shear layers rolling into vortices and where the drift correction is effective. Those diffusion modes, coupled with modal characteristic time scales, provide a subgrid model that shows the same (global) stabilizing skill as the isotropic eddy viscosity models.

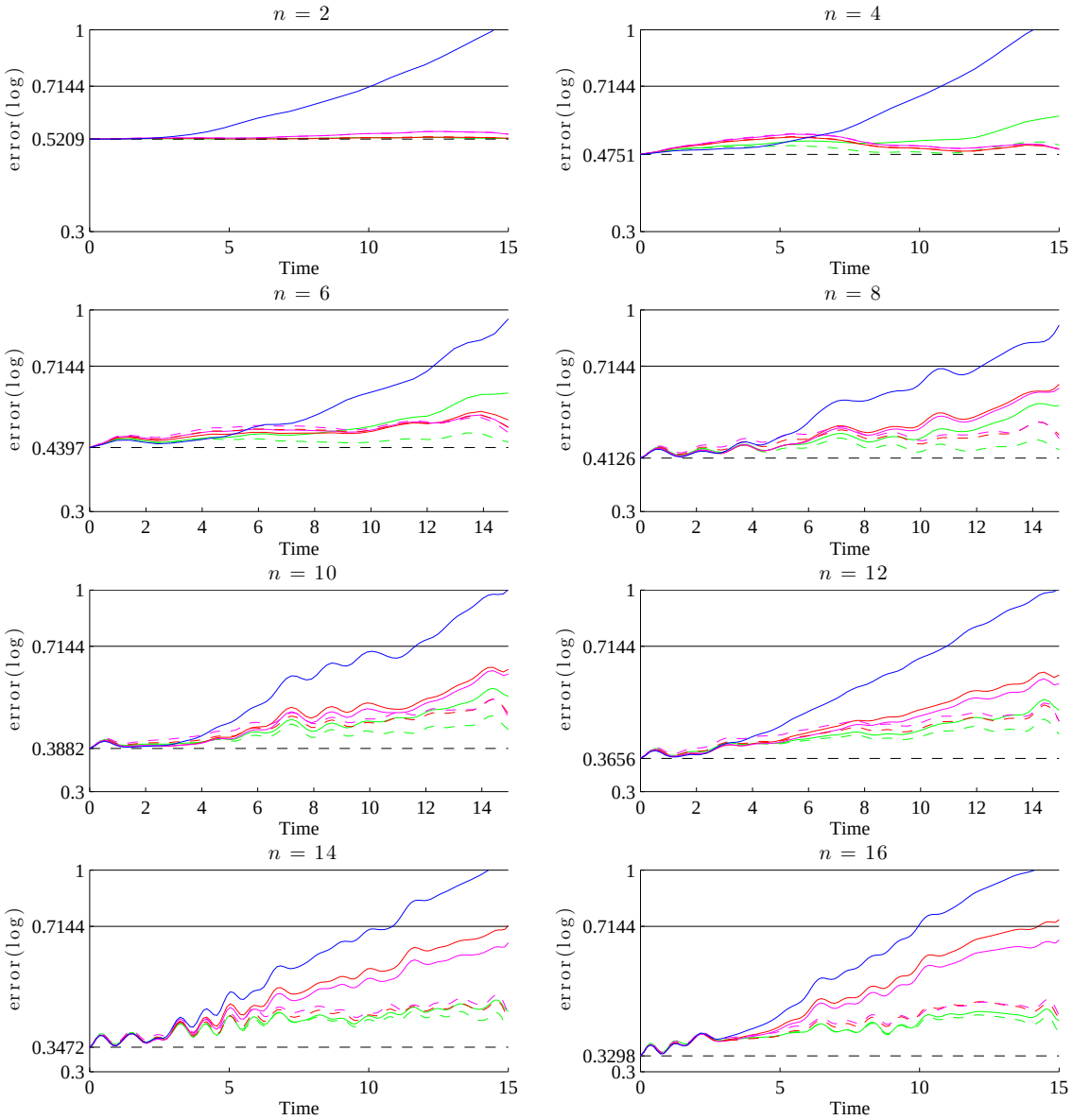


Figure 7.12: Normalized error for $n = 2$, 4, 6, 8, 10, 12, 14 and 16 modes at Reynolds number $Re = 3900$ (HR LES). The error is normalized by the energy of the solution: $\sum_{i=1}^{N} \lambda_i$. The blue colour line corresponds to the standard POD Galerkin (without eddy viscosity). The red curve stands for the stationary variance tensor. The magenta curve represents the nonstationary model of the variance tensor. The solid and dashed lines correspond to methods with single and modal characteristic time respectively. The green lines denote a reduced order model with eddy viscosity (solid line) and modal eddy viscosity (dashed line) models. The eddy viscosity coefficients are estimated through a least squares fit on the whole data sequence. The variance tensors are also estimated from a temporal mean on the whole sequence. The black doted line indicates the error associated with the mode truncation : $\sum_{i=n+1}^{N} \lambda_i$. The black solid line is the error when we only consider the temporal mean velocity.

Acknowledgements

The authors acknowledge the support of the SEACS project funded by the Brittany clusters of excellence (Cominslab, Lebesgue, and Mer).

Chapter 8

Conservative transport of fluid velocity under location uncertainty and dimensional reduction

Abstract

This chapter is devoted to a random Navier-Stokes representation and its application to ensemble forecasting in the context of reduced order models. Up to some forcings, the velocity is transported in a stochastic sense as with the buoyancy in chapters 2, 3, 4 and 5. Accordingly, our Navier-Stokes model involves an effective drift, an inhomogeneous and anisotropic diffusion but also an additive noise and a multiplicative skew-symmetric noise. Under some divergence-free conditions, the energy is conserved for each realization. Indeed, by the action of the unresolved time-uncorrelated component of the velocity, the energy diffused by the subgrid diffusion tensor is exactly compensated by the energy intake of the multiplicative noise. Thus, a turbulent diffusion of the mean velocity always generates a covariance inflation. When the resolved velocity component is expressed on a finite basis, the structure of the model restores energy fluxes between resolved modes and from resolved modes toward unresolved modes. A random Proper Orthogonal Decomposition (POD) dimensionally-reduced model is derived from this framework. The deterministic Reduced Order Model (ROM) coefficients are obtained by a Galerkin projection whereas the correlations of the multiplicative and additive noises are estimated from the residual velocity, the model structure, and the evolution of the resolved modes. The low computational cost of these rigorous estimators makes them applicable to study turbulent flows. The stochastic POD-ROM has been applied to wake flows at Reynolds 300 and 3900 for Uncertainty Quantification (UQ) purposes. Ensembles are forecast outside the learning interval. Our stochastic model stabilizes unstable modes, maintains the variability of stable modes and shows very good prediction skills.

8.1 Transport under location uncertainty of the velocity

8.1.1 Structure of the model

In Appendix 8.A, we derive a stochastic Navier-Stokes representation to describe the evolution of the large-scale velocity component $\boldsymbol{w}$:

$$\mathbb{D}_t\boldsymbol{w} + \boldsymbol{f} \times (\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \boldsymbol{g}\mathrm{d}t - \frac{1}{\rho}\boldsymbol{\nabla}p\mathrm{d}t + \nu\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{8.1.1}$$

where $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \boldsymbol{\sigma}\frac{\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$ is the unresolved small-scale velocity component assumed to be uncorrelated in time, $\mathbb{D}_t$ is the stochastic transport operator defined for any function Θ by:

$$\mathbb{D}_t\Theta \quad \triangleq \quad \underbrace{\mathrm{d}_t\Theta}_{\substack{\triangleq\ \Theta(\boldsymbol{x},t+\mathrm{d}t)-\Theta(\boldsymbol{x},t)\\ \text{Time increment}}} + \underbrace{(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\Theta}_{\text{Advection}} - \underbrace{\boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)}_{\text{Diffusion}}\mathrm{d}t, \tag{8.1.2}$$

with

$$\boldsymbol{w}^\star \triangleq \boldsymbol{w} - \frac{1}{2}\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}\right)^T. \tag{8.1.3}$$

This transport operator involves the terms of a deterministic material derivative as well as three new terms: an advection correction ($\boldsymbol{w}^\star$ instead of $\boldsymbol{w}$), an inhomogeneous and anisotropic diffusion and a multiplicative noise. This last term corresponds to the advection by the unresolved velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. We can formally rewrite the velocity evolution law (8.1.1) as follows:

$$\mathrm{d}_t\boldsymbol{w} = (\mathrm{d}\mathbb{M})(\boldsymbol{w}) \triangleq (I + L(\boldsymbol{w}) + F(\boldsymbol{w}) + C(\boldsymbol{w},\boldsymbol{w}))\,\mathrm{d}t + (\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t) + (\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w}), \tag{8.1.4}$$

where L, F and $(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)$ are linear differential operators and C is a bilinear differential operator. The first term of the right-hand side, I, includes pressure and gravity. The additive noise $(\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t)$ corresponds to the Coriolis accelaration and to the molecular viscous dissipation of the time-uncorrelated velocity component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. Under suitable boundary conditions, the algebraic structures of the different operators can be further detailed. The molecular viscosity represented by L and the inhomogeneous and anisotropic diffusion involved in F correspond to symmetric negative operators; the Coriolis force part of L, the advection correction part of F (under the incompressibility condition $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T = 0$), and the random advection involved in $(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)$ (under the incompressibility condition $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$) are skew-symmetric operators. Indeed, for every function f_1 and f_2, an integration by part gives:

$$\int_\Omega f_1\left(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t\right)f_2 \triangleq \int_\Omega f_1\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)f_2 = \int_\Omega f_1\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t f_2\right) = -\int_\Omega\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}f_1\right)f_2. \tag{8.1.5}$$

Moreover, the negative symmetric turbulent diffusion can be expressed with the random skew-symmetric operator $(\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)$:

$$\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla}\bullet\right) = \frac{1}{2}\sum_k\left(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla}\right)\left(\left(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla}\right)\bullet\right) = \frac{1}{2}\sum_k\boldsymbol{G}_{\bullet k}\boldsymbol{G}_{\bullet k} = -\frac{1}{2}\sum_k\boldsymbol{G}^\star_{\bullet k}\boldsymbol{G}_{\bullet k}, \tag{8.1.6}$$

where $G^\star$ denotes the adjoint of G. The resolved advection C (under the incompressibility condition $\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0$) is also skew-symmetric with respect to the second argument (*i.e.* $\boldsymbol{g}\mapsto C(\boldsymbol{f},\boldsymbol{g})$ is skew-symmetric). In the following, we do not consider the Coriolis force anymore.

8.1.2 Kinetic energy budget

Since the pressure does not influence the energy budget, neglecting the gravity, the molecular viscosity as well as the divergence of $\boldsymbol{w}^\star$ and $\boldsymbol{\sigma}$ lead by the Itō lemma to the kinetic energy budget:

$$\mathrm{d}\left(\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\mathbb{R}^d))^d}\right) = \int_\Omega\left(\mathrm{d}_t\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}\mathrm{d}_t < \boldsymbol{w}^T,\boldsymbol{w} >\right), \tag{8.1.7}$$

$$= \underbrace{\int_\Omega\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla}\boldsymbol{w}^T\right)\boldsymbol{w}\,\mathrm{d}t}_{\text{Loss by diffusion}} + \underbrace{\int_\Omega\frac{1}{2}\mathrm{tr}\left(\left(\boldsymbol{\nabla}\boldsymbol{w}^T\right)^T\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{w}^T\right)\mathrm{d}t}_{\text{Energy flux from the noise}}, \tag{8.1.8}$$

$$= 0. \tag{8.1.9}$$

The energy intake of the noise and the dissipation by turbulent diffusion exactly compensate each other. The energy is conserved for each realization. The diffusion decreases the energy of the velocity, especially the energy of the mean, $\|\mathbb{E}\{\boldsymbol{w}\}\|^2_{(\mathcal{L}^2(\Omega))^d}$, whereas the multiplicative noise

generates only random energy, $\|\boldsymbol{w}-\mathbb{E}\{\boldsymbol{w}\}\|^2_{(\mathcal{L}^2(\Omega))^d}$. Thus, the time-uncorrelated component of the velocity drains energy from the mean field to the random component of $\boldsymbol{w}$. We can also express this energy transfer with the expectation of equation (8.1.9):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega Var(\boldsymbol{w}) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\|\boldsymbol{w}-\mathbb{E}\{\boldsymbol{w}\}\|^2_{(\mathcal{L}(\Omega))^d} = -\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbb{E}\{\boldsymbol{w}\}\|^2_{(\mathcal{L}(\Omega))^d}. \tag{8.1.10}$$

Besides the physical relevance of the energy conservation, this variance inflation is of primary interest for ensemble forecasts. After its generation by the noise, this random energy may come back to the mean field by nonlinear interactions.

8.1.3 Incompressible Navier-Stokes under location uncertainty with and without noise

Note that the Navier-Stokes model (8.1.1) involves noises whereas the Navier-Stokes representation of Mémin (2014) used in chapter 2 (equation (2.3.7)), chapter 6 (equation (6.2.7)) and chapter 7 (equation (7.2.8)) does not. In these chapters, the large-scale velocity, $\boldsymbol{w}$, is assumed to be differentiable w.r.t. time. In other words, we neglect the irregular component of the large-scale velocity which is continuous but not differentiable. We only keep the very smooth component and the very irregular component, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. So, this statement may be understood as a gap-scale assumption. The ensuing Navier-Stokes model is deterministic even though it involves subgrid terms. As such, it is very useful for deterministic LES-like simulations. In contrast, this chapter does not rely on that assumption and proposes a Navier-Stokes model (8.1.1) with noises. Up to some forcings, the large-scale velocity is transported under location uncertainty as every other tracers. Unlike the Navier-Stokes model without noise, that model has to be simulated with MCMC and can be applied to uncertainty quantification, especially for data assimilation purposes.

8.2 Galerkin projection

8.2.1 A finite-dimensional system with correlated additive and multiplicative noise

Let $\boldsymbol{\phi}_0$ be a stationary spatial field that represents a possible stationary background velocity. Its precise definition will be given in the following. As in the previous chapters, we express the velocity anomaly, $\boldsymbol{w}-\boldsymbol{\phi}_0$, in a finite-dimensional set of orthogonal spatial modes $(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}$. Yet, the following analysis is not restricted to POD-ROMs. The spatial modes, $(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}$ can also be Fourier modes or any other orthogonal set. In the case of Fourier modes, we would probably choose $\boldsymbol{\phi}_0 = 0$ whereas in the POD-ROM framework, $\boldsymbol{\phi}_0 = \overline{\boldsymbol{w}}$ is the temporal mean of the velocity $\boldsymbol{w}$. The associated temporal modes, b_i, of the velocity are random and depend on a realization $\omega \in \breve{\Omega}$ whereas the spatial modes $\boldsymbol{\phi}_i$, are assumed to be deterministic and stationary:

$$\boldsymbol{w}(\boldsymbol{x},t,\omega) \quad = \quad \sum_{i=0}^{n} b_i(t,\omega)\boldsymbol{\phi}_i(\boldsymbol{x}), \tag{8.2.1}$$

where $b_0 = 1$.

A ROM is a system gathering SDEs of all the temporal modes. With ROMs, we aim at approximating the evolution of a true velocity field denoted $\boldsymbol{v}$. Usually, a ROM enables only a finite number of modes (8.2.1) to be solved. To compensate for the mode truncation, the unresolved velocity component $\boldsymbol{v}-\boldsymbol{w}$ is assumed to be uncorrelated in time and hence well modeled by $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. Accordingly, the resolved component $\boldsymbol{w}$ is the solution of the Navier-Stokes representation (8.1.4). To obtain the reduced model, this SPDE is projected onto each spatial modes $\boldsymbol{\phi}_i$ $(1 \leqslant i \leqslant n)$:

$$\mathrm{d}b_i = (\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b}) \triangleq \int_\Omega \boldsymbol{\phi}_i \cdot (\mathrm{d}\mathbb{M})(\boldsymbol{w}), \tag{8.2.2}$$

where $\boldsymbol{b} = (b_i)_{0\leqslant i\leqslant n}$. The coefficients of $\mathrm{d}\mathbb{M}^R$ can be evaluated with (8.1.4):

$$(\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b}) = \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot I\right)}_{\triangleq i_i} \mathrm{d}t + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot L(\boldsymbol{\phi}_p)\right)}_{\triangleq l_{pi}} b_p \mathrm{d}t + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot F(\boldsymbol{\phi}_p)\right)}_{\triangleq \breve{f}_{pi}} b_p \mathrm{d}t$$
$$+ \sum_{p,q=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot C(\boldsymbol{\phi}_p, \boldsymbol{\phi}_q)\right)}_{\triangleq c_{pqi}} b_p b_q \mathrm{d}t + \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot (\boldsymbol{H}\mathrm{d}\boldsymbol{B}_t)\right)}_{\triangleq (\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t)} + \sum_{p=0}^{n} \underbrace{\left(\int_\Omega \boldsymbol{\phi}_i \cdot (\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{\phi}_p)\right)}_{\triangleq (\boldsymbol{\alpha}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t)} b_p. \tag{8.2.3}$$

The coefficients $i_0, l_{\bullet 0}, \breve{f}_{\bullet 0}, c_{\bullet\bullet 0}, (\boldsymbol{\theta}_{0\bullet}\mathrm{d}\boldsymbol{B}_t)$, $(\boldsymbol{\alpha}_{\bullet 0\bullet}\mathrm{d}\boldsymbol{B}_t)$ and the operator $(\mathrm{d}\mathbb{M}_0^R)$ can be defined in a similar way. Nevertheless, they do not corresponds to the dynamics of $b_0 = 1$. Moreover, they are not necessary to simulate the ROM. The terms $(\boldsymbol{\alpha}\mathrm{d}\boldsymbol{B}_t) = (\boldsymbol{\alpha}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t)_{0\leqslant p,i\leqslant n}$ and $(\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t) = (\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t)_{0\leqslant i\leqslant n}$ correspond to a Gaussian skew-symmetric matrix and a Gaussian vector respectively with correlated coefficients. All the correlations can be expressed through the spatial modes and the small-scale velocity covariance tensor $\boldsymbol{a}(\boldsymbol{x}, \boldsymbol{y})$. If the modes $\boldsymbol{\phi}_i$ are spatial Fourier modes associated with small wave-numbers, the ROM (8.2.2) is a (random) LES-like simulation expressed in Fourier space and $\boldsymbol{b}$ is the set of Fourier coefficients of the solution.

8.2.2 Energy leak

The matrix $(\boldsymbol{\alpha}\mathrm{d}\boldsymbol{B}_t)$ is antisymmetric since it expresses the advection by the divergence-free velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$. However, due to its random structure, it increases the kinetic energy, $\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d}$. This created variance is dissipated by the symmetric part of matrix $\breve{\boldsymbol{f}} = (\breve{f}_{ij})_{0\leqslant i,j\leqslant n}$ (see equation (8.1.9)). Indeed, under the incompressibility condition $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a})^T = 0$, the symmetric part of $\breve{\boldsymbol{f}}$ corresponds to the dissipative sub-grid tensor, whereas its antisymmetric part encodes the advection correction. Up to molecular viscosity, gravity and boundary conditions effects, the global Navier-Stokes model (8.1.4) conserves the energy as demonstrated by equation (8.1.9). Nevertheless, it is important to outline that the reduced order model (8.2.2) does not exactly conserve the resolved kinetic energy $\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d}$. The mode truncation introduces a small leak of energy. This is due to the fact that the ROM (8.2.2) does not solves exactly the global Navier-Stokes model (8.1.4). It solves instead its Galerkin projection onto the reduced subspace spanned by the spatial modes $(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}$:

$$\mathrm{d}_t\boldsymbol{w} = \sum_{i=0}^{n} \mathrm{d}b_i\ \boldsymbol{\phi}_i = \sum_{i=1}^{n} (\mathrm{d}\mathbb{M}_i^R)(\boldsymbol{b})\ \boldsymbol{\phi}_i = \sum_{i=1}^{n} \left(\int_\Omega \boldsymbol{\phi}_i \cdot (\mathrm{d}\mathbb{M})(\boldsymbol{w})\right)\ \boldsymbol{\phi}_i = \Pi_\phi\left[(\mathrm{d}\mathbb{M})(\boldsymbol{w})\right], \tag{8.2.4}$$

where Π_ϕ is the projection onto the reduced subspace. Specifically, for any function $\boldsymbol{h}$, the projection, Π_ϕ, is defined as follows:

$$(\Pi_\phi[\boldsymbol{h}])\,(\boldsymbol{x}, t) \triangleq \sum_{p=1}^{n} \left(\int_\Omega \boldsymbol{\phi}_p \cdot \boldsymbol{h}(\bullet, t)\right) \boldsymbol{\phi}_p(\boldsymbol{x}). \tag{8.2.5}$$

If the spatial modes are Fourier modes associated with small wave-numbers, the projection $\Pi_\phi[\boldsymbol{h}]$ is the large-scale component of the field h. As proofed in Appendix 8.B, the kinetic energy diffused in the ROM (8.2.2)-(8.2.3) by the subgrid tensor by units of time is given by:

$$\int_\Omega \mathrm{d}_t\boldsymbol{w}^T\boldsymbol{w} = \int_\Omega \boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla}\boldsymbol{w}^T\right)\boldsymbol{w}\ \mathrm{d}t = -\frac{1}{2}\ \sum_k \|(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\|^2_{(\mathcal{L}^2(\Omega))^d}\ \mathrm{d}t. \tag{8.2.6}$$

The above expression is similar to the dissipation of the complete stochastic Navier-Stokes model (8.1.8). Nevertheless, the velocity $\boldsymbol{w}$ is different because it is the solution of another SPDE. The

dissipation (8.2.9) also corresponds to the energy diffused in the ROMs of chapters 6 and 7. In contrast, the Itō term encoding the energy flux from the noise is

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega < \boldsymbol{w}^T, \boldsymbol{w} >= \frac{1}{2}\sum_k \|\Pi_\phi\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\|^2_{(\mathcal{L}^2(\Omega))^d} = \boldsymbol{b}^T\left(\frac{1}{2}\sum_{p=1}^n \boldsymbol{\alpha}_{\bullet p\bullet}\boldsymbol{\alpha}^T_{\bullet p\bullet}\right)\boldsymbol{b}. \tag{8.2.7}$$

Finally, as for the full Navier-Stokes model, we can evaluate the variation of kinetic energy with the Itō formula:

$$\mathrm{d}\left(\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d}\right) = \int_\Omega\left(\mathrm{d}_t\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}\mathrm{d}_t < \boldsymbol{w}^T, \boldsymbol{w} >\right), \tag{8.2.8}$$

$$= -\frac{1}{2}\sum_k \left\|\Pi^\perp_\phi\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\right\|^2_{(\mathcal{L}^2(\Omega))^d} < 0, \tag{8.2.9}$$

where $\Pi^\perp_\phi = \mathbb{I}_d - \Pi_\phi$ is the projector onto the space orthogonal to the spatial modes. The subgrid diffusion extracts energy from some modes and the multiplicative noise gives energy to others. These stabilization and destabilization effects restore a large part of the energy transfers occurring between resolved modes. Yet, the uncorrelated velocity component also drains energy from the resolved modes to the unresolved modes. This energy flux is exactly the right-hand side of (8.2.9). In contrast, the transport under location uncertainty does not naturally transfer energy from unresolved modes to resolved modes. This explains the energy leak.

We could prevent this dissipation by using

$$-\frac{1}{2}\sum_k \left(\Pi_\phi\left[\boldsymbol{G}_{\bullet k}\right]\right)^\star \Pi_\phi\left[\boldsymbol{G}_{\bullet k}\right], \tag{8.2.10}$$

represented in the ROM by

$$-\frac{1}{2}\sum_{p=1}^n \boldsymbol{\alpha}_{\bullet p\bullet}\boldsymbol{\alpha}^T_{\bullet p\bullet}, \tag{8.2.11}$$

instead of the full turbulent diffusion operator:

$$-\frac{1}{2}\sum_k \boldsymbol{G}^\star_{\bullet k}\boldsymbol{G}_{\bullet k} = \boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla}\bullet\right). \tag{8.2.12}$$

Notwithstanding, the additional dissipation (8.2.9) should not be a problem. First, it enables a molecular viscous energy dissipation in the unresolved modes by a direct cascade from the resolved modes. So, for very turbulent flow described by few modes, the energy leak appears necessary. Moreover, our main concern consists in restoring energy fluxes between resolved modes. Indeed, the main energy exchanges appear between the energetic modes. As in POD, the most energetic modes of the ROM are often the resolved modes. Those energy fluxes are very difficult to model correctly in ROMs. As an example, Sapsis and Majda (2013b) observe that a basis encoding 50% of the energy can lead to a ROM missing more than 98% of energy transfers. Restoring the energy fluxes between resolved modes is a main issue and it is the issue we want to address with the models under location uncertainty.

Note also that due to the interactions with the stationary background $\boldsymbol{\phi}_0$ the energy variation of temporal modes $\sum_{i=1}^n b_i^2 = \|\sum_{i=1}^n b_i\boldsymbol{\phi}_i\|^2_{(\mathcal{L}^2(\Omega))^d}$ differs from the kinetic energy variation:

$$\mathrm{d}\left(\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d} - \frac{1}{2}\sum_{i=1}^n b_i^2\right) = \mathrm{d}\int_\Omega \boldsymbol{\phi}_0\cdot\left(\frac{1}{2}\boldsymbol{\phi}_0 + \sum_{i=1}^n b_i\boldsymbol{\phi}_i\right) = \int_\Omega \boldsymbol{\phi}_0\cdot\mathrm{d}\boldsymbol{w} = (\mathrm{d}\mathbb{M}^R_0)(\boldsymbol{b}) \neq 0. \tag{8.2.13}$$

8.2.3 Proper Orthogonal Decomposition

In the POD framework, the spatial modes, so-called *topos*, are obtained from a set of velocity snapshots $(\boldsymbol{v}_{obs}(\bullet, t_i))_{0\leqslant i\leqslant N}$. More precisely, the *topos* are the solution of the constrained optimization

problem:

$$\underset{(\boldsymbol{\phi}_i)_{1\leqslant i\leqslant n}}{\text{Maximize}} \sum_{i=1}^{n} \int_0^T \left(\int_\Omega \boldsymbol{\phi}_i \cdot (\boldsymbol{v}_{obs} - \overline{\boldsymbol{v}}) \right)^2 \text{ subject to } \int_\Omega \boldsymbol{\phi}_i \cdot \boldsymbol{\phi}_j = \delta_{ij}. \tag{8.2.14}$$

Intuitively, the *topos* are the n orthonormal functions which can best explain the snapshots' temporal variability. The mode $\boldsymbol{\phi}_0 = \overline{\boldsymbol{v}}$ is set to the time averaged velocity. The temporal modes $\boldsymbol{b}$ are called *chronos* and the temporal variances, $\frac{1}{T}\int_0^T b_i^2$, are denoted λ_i. Moreover, if the snapshots describe a divergence-free velocity field, the *topos* $\boldsymbol{\phi}_i$ are divergence-free as well.

8.3 Estimations of subgrid terms

The matrix $\breve{\boldsymbol{f}}$ is defined by the *topos* and the one-point quadratic cross-variation tensor $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{x})$. The correlations of the Gaussian noises $(\boldsymbol{\alpha}\mathrm{d}\boldsymbol{B}_t)$ and $(\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t)$ also involve the two-point quadratic cross-variation tensor of the small-scale velocity, $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})$. For instance, the coefficients of the covariance matrix of the Gaussian vector $(\boldsymbol{\theta}\mathrm{d}\boldsymbol{B}_t)$ are:

$$\begin{aligned}
\mathbb{E}\left\{(\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t)\,(\boldsymbol{\theta}_{j\bullet}\mathrm{d}\boldsymbol{B}_t)\right\} &= \mathbb{E}\left\{\left(\int_\Omega \mathrm{d}\boldsymbol{x}\ \boldsymbol{\phi}_i(\boldsymbol{x})^T \nu \boldsymbol{\nabla}_x^2 \boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t\right)\left(\int_\Omega \mathrm{d}\boldsymbol{y}\ \nu\boldsymbol{\nabla}_y^2\left(\boldsymbol{\sigma}(\boldsymbol{y})\mathrm{d}\boldsymbol{B}_t\right)^T \boldsymbol{\phi}_j(\boldsymbol{y})\right)\right\}, \\
&= \iint_{\Omega^2} \mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{y}\ \boldsymbol{\phi}_i(\boldsymbol{x})^T \nu^2 \boldsymbol{\nabla}_x^2 \boldsymbol{\nabla}_y^2 \boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})\mathrm{d}t\ \boldsymbol{\phi}_j(\boldsymbol{y}).
\end{aligned}$$

Therefore, the whole stochastic ROM is closed knowing the tensor $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})$. However, in practice, if the covariance tensor $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})$ is inhomogeneous, its size is so large that its estimation or even its storage becomes prohibitive. Nevertheless, there are at least three ways to overcome this difficulty. The first one is to assume a model structure for the covariance. For instance, it can be considered as isotropic or at least only slightly anisotropic and inhomogeneous. In this case, the covariance can be specified through a phenomenological model of the velocity spectrum, as in Kraichnan (1968) and chapters 3 and 5 of this thesis. Modeling the covariance is almost mandatory when the smallest-scale velocity is not observed. Yet if the small-scale velocity is observed, other techniques can be considered. Indeed, one can build a reduced representation of the small-scale velocity $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, learned from available snapshots or realizations. This efficient method often relies on a Karhunen-Loeve decomposition. In the particular framework of POD, a great number of *topos* have to be estimated and the number of coefficients involved in the reduced order system increases. Depending on the desired accuracy of the noise representation, the complexity of this method may increase quickly. Here, we will rely on a third method, specifically setup for ROM frameworks. Similarly to the second method, the noise structure is learned from observed small-scales snapshots, without using the covariance $\boldsymbol{a}(\boldsymbol{x},\boldsymbol{y})$. This method estimates directly the correlations of the random ROM's coefficients.

8.3.1 Estimation formulas

In order to close our new stochastic ROM system, we need to estimate the variance tensor $\boldsymbol{a}$ and the variance and correlations between the noises of the ROM. First, recalling that $b_0 = 1$, we note that:

$$(\boldsymbol{\alpha}_{\bullet i\bullet}\mathrm{d}\boldsymbol{B}_t)^T\,\boldsymbol{b} + (\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t) = \sum_{k=1}^{n} (\boldsymbol{\alpha}_{ki\bullet}\mathrm{d}\boldsymbol{B}_t)\,b_k + ((\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_t)\,.$$

Therefore, the multiplicative and additive noises of the ROM correspond to the first and second term of the right-hand side respectively. Then, we need to estimate the following correlations:

$$\left\{\begin{aligned}
\boldsymbol{a}(\boldsymbol{x},\boldsymbol{x})\mathrm{d}t &= \mathbb{E}\left\{(\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t)\,(\boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t)^T\right\}, \\
(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\cdot(\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}t &= \mathbb{E}\left\{((\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_t)\,((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_t)\right\}, \\
(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\cdot\boldsymbol{\alpha}_{qj\bullet}\mathrm{d}t &= \mathbb{E}\left\{((\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_t)\,(\boldsymbol{\alpha}_{qj\bullet}\mathrm{d}\boldsymbol{B}_t)\right\}, \\
\boldsymbol{\alpha}_{pi\bullet}\cdot\boldsymbol{\alpha}_{qj\bullet}\mathrm{d}t &= \mathbb{E}\left\{(\boldsymbol{\alpha}_{pi\bullet}\mathrm{d}\boldsymbol{B}_t)\,(\boldsymbol{\alpha}_{qj\bullet}\mathrm{d}\boldsymbol{B}_t)\right\},
\end{aligned}\right. \tag{8.3.1}$$

for $1 \leqslant i, j, p, q \leqslant n$. Assuming the variance tensor $\boldsymbol{a}$ to be stationary leads to rigorous estimators through time averaging. Besides the variance tensor, there are $O(n^4)$ correlations to estimate.

For any function $\boldsymbol{\xi}$ and $q > 0$, let us introduce the functionals:

$$K_{jq}[\boldsymbol{\xi}] \triangleq -\int_{\Omega} \boldsymbol{\phi}_j \cdot \left(\left(\boldsymbol{\xi} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_q \right), \tag{8.3.2}$$

and

$$K_{j0}[\boldsymbol{\xi}] \triangleq \int_{\Omega} \boldsymbol{\phi}_j \cdot \left(- \left(\boldsymbol{\xi} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_0 + \nu \Delta \boldsymbol{\xi} \right). \tag{8.3.3}$$

With these notations, the estimation formulas are:

$$(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet}) \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) \approx K_{j0} \left[\frac{\Delta t}{N+1} \sum_{k=0}^{N} \left(\frac{\mathrm{d}b_i^{obs}}{\mathrm{d}t} \right) (t_k) (\boldsymbol{v} - \boldsymbol{w})_{obs}(t_k, \bullet) \right], \tag{8.3.4}$$

$$\boldsymbol{\alpha}_{pi\bullet} \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) \approx K_{j0} \left[\frac{\Delta t}{\lambda_p (N+1)} \sum_{k=0}^{N} b_p^{obs}(t_k) \left(\frac{\mathrm{d}b_i^{obs}}{\mathrm{d}t} \right) (t_k) (\boldsymbol{v} - \boldsymbol{w})_{obs}(t_k, \bullet) \right], \tag{8.3.5}$$

$$\boldsymbol{\alpha}_{pi\bullet} \cdot \boldsymbol{\alpha}_{qj\bullet} \approx K_{jq} \left[\frac{\Delta t}{\lambda_p (N+1)} \sum_{k=0}^{N} b_p^{obs}(t_k) \left(\frac{\mathrm{d}b_i^{obs}}{\mathrm{d}t} \right) (t_k) (\boldsymbol{v} - \boldsymbol{w})_{obs}(t_k, \bullet) \right], \tag{8.3.6}$$

where $1 \leqslant i, j, p, q \leqslant n$, $(\boldsymbol{v} - \boldsymbol{w})_{obs}$ is the observed residual velocity,

$$b_i^{obs} = \int_{\Omega} \boldsymbol{\phi}_i \cdot \boldsymbol{v}_{obs}, \tag{8.3.7}$$

$$\lambda_i^{obs} = \frac{1}{T} \int_0^T (b_i^{obs})^2, \tag{8.3.8}$$

and $N+1$ the number of available snapshots. The derivation of these formulas is provided in the Appendix 8.C. Since the time averaging is performed before the application of the functionals K_{jq}, the differential operators involved in these functionals need to be computed only once for each couple (i, j, p, q). For moderate values of n, the low computational cost of these estimators is well adapted to the derivation of ROMs describing strongly turbulent flows. These estimations are direct due to the orthogonality of the observed *chronos* b_i^{obs}. In a more general (non-orthogonal) setting, the estimated correlations are solutions of a linear inverse problem. The variance tensor will be estimated as in chapters 6 and 7 by:

$$\boldsymbol{a}(\boldsymbol{x}, \boldsymbol{x}) \approx \frac{\Delta t}{N+1} \sum_{k=0}^{N} (\boldsymbol{v} - \boldsymbol{w})_{obs}(\boldsymbol{x}, t_k) \, (\boldsymbol{v} - \boldsymbol{w})_{obs}^{T}(\boldsymbol{x}, t_k). \tag{8.3.9}$$

With these estimations done, all the parameters of the stochastic ROM (8.2.2)-(8.2.3) are known. It is then possible to forecast an ensemble of realizations of the reduced model through Monte-Carlo simulations.

8.3.2 Modal time step

We perform the same modal time step estimation than in chapter 7 and modify the stochastic ROM accordingly:

$$\mathrm{d}b_i = \left(i_i + \left(\frac{\Delta t_i}{\Delta t} \breve{\boldsymbol{f}}_{\bullet i} + \boldsymbol{l}_{\bullet i} \right)^T \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i} \boldsymbol{b} \right) \mathrm{d}t + \sqrt{\frac{\Delta t_i}{\Delta t}} \left(\boldsymbol{\alpha}_{\bullet i\bullet} \mathrm{d}\boldsymbol{B}_t \right)^T \boldsymbol{b} + \sqrt{\frac{\Delta t_i}{\Delta t}} \left(\boldsymbol{\theta}_{i\bullet} \mathrm{d}\boldsymbol{B}_t \right). \tag{8.3.10}$$

8.4 Ensemble forecasting

We have numerically tested the UQ skills of the ROM (8.2.2)-(8.2.3) on two DNS simulations of wake flows at Reynolds 300 and 3900 respectively. The first one is quasi two dimensional whereas

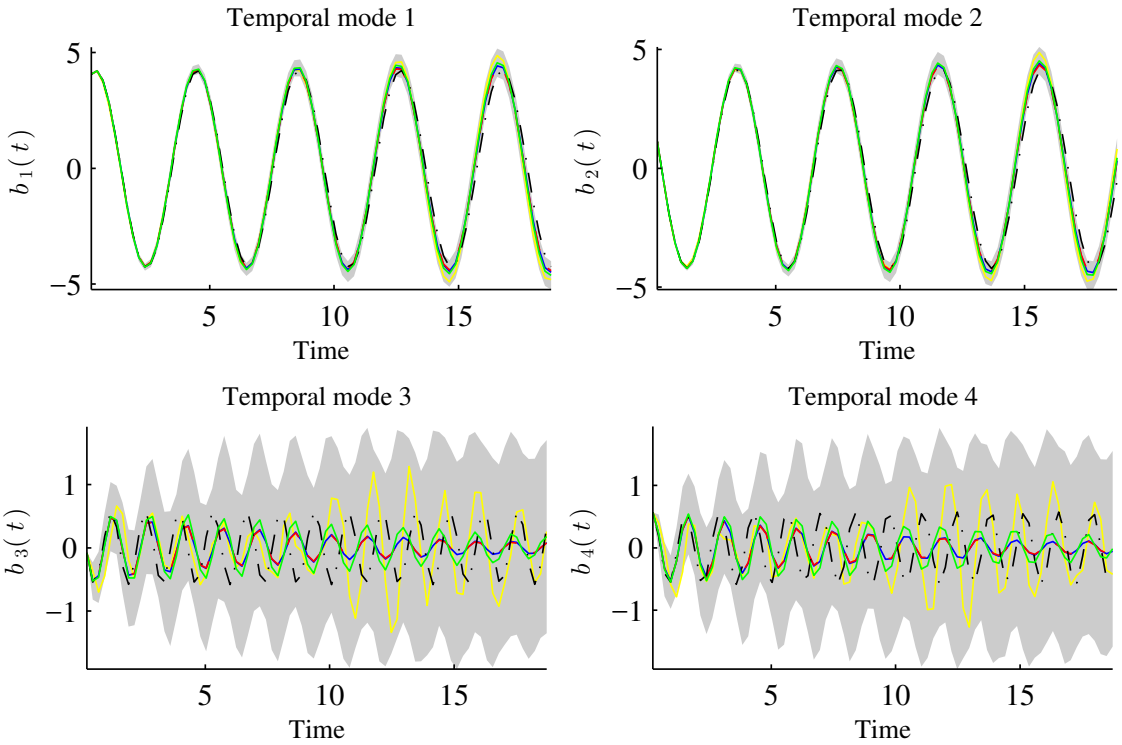


Figure 8.1: Forecast of the 4 modes for $n = 4$ modes of a wake flow at Reynolds 300. The dash-dot black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the dashed red ones are computed with the smooth stochastic ROM (see chapters 6 and 7). From the ensemble simulated according to the full-stochastic POD-ROM, we have plotted one realization (yellow line), the ensemble mean (green line) and $+/-1.96\times$ the standard-deviation (shaded gray). The initial condition at $t = 0$ is common to all methods.

the second one is three dimensional. The Modal time step method (8.3.10) is considered for the high Reynolds simulation since the associated dynamics involve a wide set of time scales. The reference velocity snapshots involved in the following tests have been used to estimate neither the *topos* nor the structure of the ROM. As such, our simulations correspond to forecasts rather than to reconstructions. Ensemble of simulations are simulated for both Reynolds numbers.

In Figure 8.1 an ensemble of 4 modes of the 300-Reynolds flow are compared to the reference *chronos* (dash-dot black line): $b_i^{ref} = \int_\Omega \boldsymbol{\phi}_i \cdot \boldsymbol{v}_{ref}$. We learned the ROM parameters on 10 vortex sheddings and simulate it on the 5 following ones. We also show the results given by the standard deterministic POD-Galerkin projection (blue) and the stochastic ROM without noise of chapters 6 and 7 (dotted red line). In the following, we will refer to this latter ROM as the smooth stochastic ROM whereas the focus of this chapter, the ROM (8.2.2)-(8.2.3), will be referred to as the full-stochastic ROM. The ensemble mean computed with this ROM (green) is similar to the smooth ROM and the deterministic POD for the first two modes. Yet, in modes 3 and 4 the ensemble mean is the only system which is not completely damped. This must be due to energy transfers from other modes restored by the multiplicative noise and to the large random energy ($+/-1.96\times$ the standard-deviation in shaded gray) of modes 3 and 4 which fed the mean by nonlinear interactions. This large random energy also explains why the ensemble mean of these modes is more damped than the reference. Indeed, the energy of these modes is contained both in the mean and in the variance. The ensemble mean has progressively transformed its energy into variance. But, this variance is coherent as illustrated by a realization of the ensemble (yellow) which correctly follows the reference of modes 3 and 4 during the first and the last 5 units of time.

For $n = 4$ modes, the energy transfers involving modes 3 and 4 have been maintained by randomizing them. A partially random mode is more desirable than a single erroneous deterministic

mode. Nevertheless, such a randomness is always less informative than an accurate deterministic mode. Accordingly, a ROM should randomly transfer energy only when it is not possible to deterministically do so. With $n = 8$ modes, the dynamics is better resolved and less spreading is necessary as revealed by Figure 8.2. Modes 3 and 4 are now well resolved and mainly deterministic. This suggests that modes from 5 to 8 determine a large amount of energy fluxes between the mode 3 (resp. 4) and the four first modes by triades interactions. Indeed, the source of randomness $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$ plays a key role on modes 3 and 4 only for $n = 4$ modes, *i.e.* only when this unresolved velocity mimics the modes 5 to 8. For $n = 8$ modes, this is now the modes 7 and 8 which depend on unresolved dynamics. They are thus partially random. The non-stationary spreading of the ensemble well represent the reference *chronos* even the under-resolved two last modes. All things considered, Figures 8.1 and Figure 8.2 have demonstrated that our stochastic model injects exactly the right amount of randomness in the system.

Although necessary when the number of modes is not large enough, random energy transfers strongly increase the modes' variance keeping the biases constant and hence increase the Root Mean Square Error (RMSE) of the ensemble:

$$RMSE \triangleq \hat{\mathbb{E}} \left\| \boldsymbol{w} - \boldsymbol{v}_{ref} \right\|^2_{(\mathcal{L}(\Omega))^d} = \left\| \hat{\mathbb{E}}\{\boldsymbol{w}\} - \boldsymbol{v}_{ref} \right\|^2_{(\mathcal{L}(\Omega))^d} + \int_\Omega \widehat{Var}(\boldsymbol{w}). \tag{8.4.1}$$

Figure 8.3 shows the RMSE (magenta crosses) from $n = 2$ to 16 modes. For $n = 4$ compared to $n = 2$, we notice an increasing of variance and RMSE even though the ensemble minimal error decreases. This confirms our previous analysis. The modes 3 and 4 must be strongly related to the dynamics of higher-order modes. As such, the dynamics of modes 3 and 4 involve important random energy fluxes which increases the variance. In contrast, the variance and the RMSE is much lower for $n = 8$. With more modes, a larger part of the energy can be deterministically transferred reducing the variance and hence the RMSE. In this case, a lower variance is better because it comes with a better resolution of the ROM. At a given RMSE – a given error amplitude of the system – the spreading has to be as large as possible to be as close as possible to the reference. Equation 8.4.1 also expresses this geometric constraint: a larger variance means a smaller bias and thus an ensemble centered closer to the reference. Furthermore, for data assimilation procedures a larger variance prevents the degeneracy of the filter. Beside the bias (green), the distance between the reference to the ensemble (magenta line) can be geometrically understood as the minimal distance *i.e.* the distance to the closest particle of the ensemble. In an assimilation with Particle Filter (PF) with a simple model of measurement errors, this particle is the Maximum *A Posteriori* (MAP). During the assimilation of a measurement, the RMSE decreases and becomes closer to this minimal error. Unlike Ensemble Kalman Filter (EnKF), there is no correction and this minimal error is a lower bound for the error. Figure 8.3 unveils a small minimal distance for every value of n and in this way confirms the very good UQ skills of the ROM. For $n = 16$ modes, the results are still good. Nevertheless, they do not overtake the skills of the 8-mode ROM. This may be due to the relatively small amount of data exploited to learn the ROM coefficients. Indeed, the multiplicative noise is described by $O(n^4)$ coefficients ($O(10^5)$ for $n = 16$). A larger learning set is hence necessary.

We have also applied our algorithm to snapshots derived from a DNS at Reynolds 3900. Two additional issues arose. First, the flow being much more turbulent, it is very difficult to simulate a ROM with few modes. Moreover, this data set has been generated with a spatially-unresolved simulation. For more details, chapter 7 describes this issue with more details and referred to this data set as Low-Resolution DNS (LR DNS). Accordingly, the data do not exactly follow the Navier-Stokes equation and deriving a ROM from physics by a Galerkin projection becomes less accurate. Even with these additional difficulties, Figure 8.4 shows that the ROM stabilizes the global system. Moreover, modes 3 and 4 of a full-stochastic POD-ROM with 16 modes are able to spread the ensemble during the 10 last units of time especially to compensate the phase shift of the mean.

Figure 8.5 displays small ensemble errors (RMSE and minimal error) for the first 15 units of time (3 vortex sheddings). This suggests that an assimilation procedure would lead to good results if a measurement were assimilated every 2 or 3 vortex sheddings.

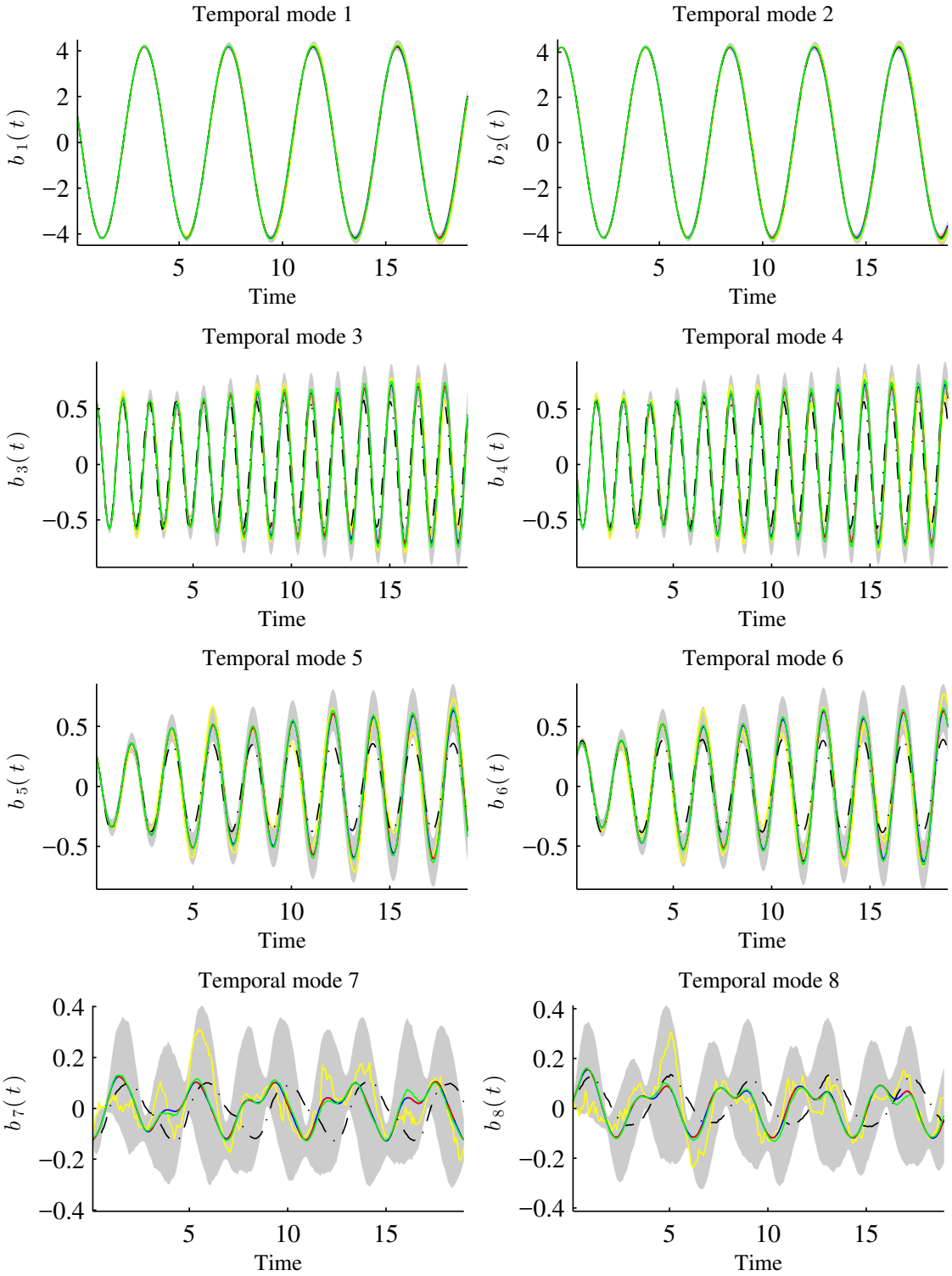


Figure 8.2: Forecast of the 8 modes for $n = 8$ modes of a wake flow at Reynolds 300. The dash-dot black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the dashed red ones are computed with the smooth stochastic ROM (see chapters 6 and 7). From the ensemble simulated according to the full-stochastic POD-ROM, we have plotted one realization (yellow line), the ensemble mean (green line) and $+/-1.96\times$ the standard-deviation (shaded gray). The initial condition at $t = 0$ is common to all methods.

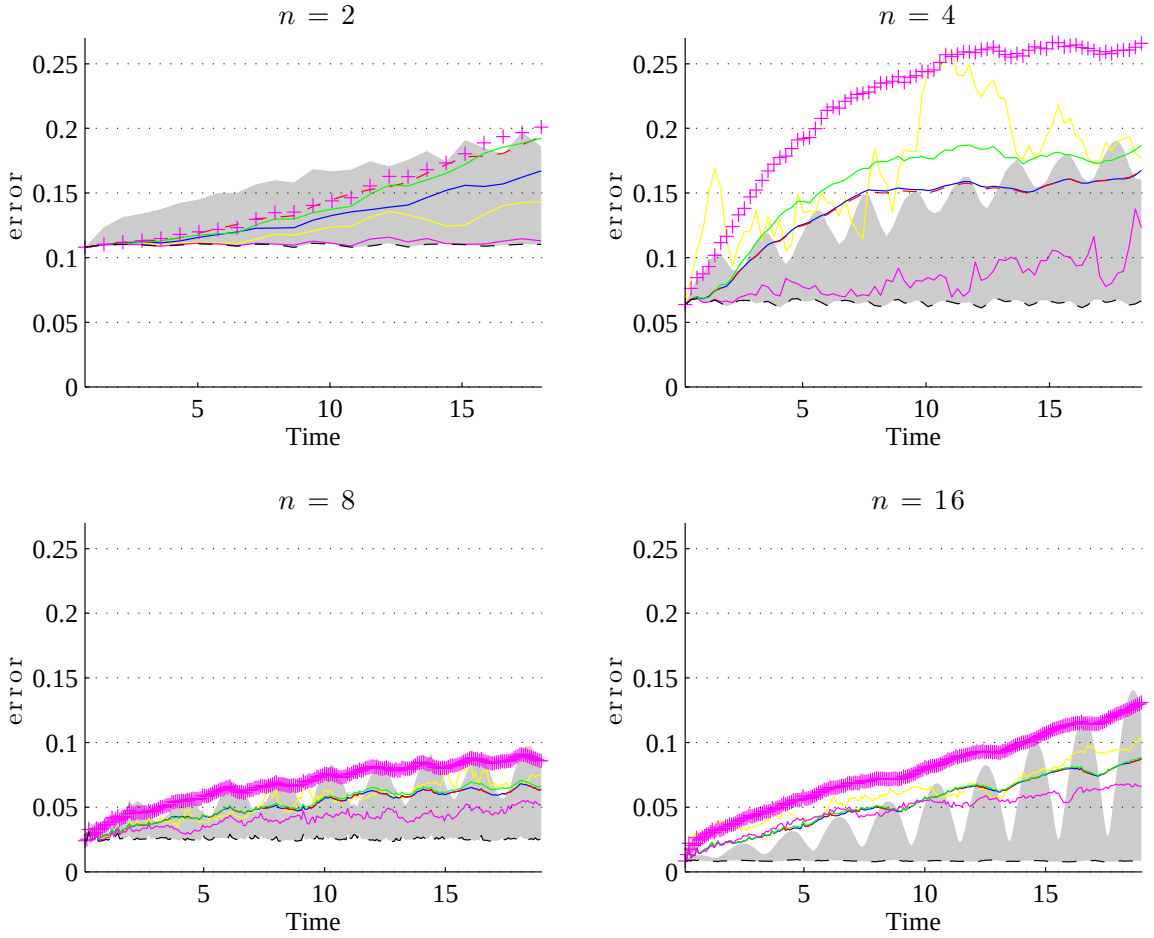


Figure 8.3: Normalized error for $n = 2$, 4, 8 and 16 modes of a wake flow at Reynolds 300. The error is normalized by an approximation of the square root of the solution energy: $\left(\|\overline{\boldsymbol{v}}\|^2_{(\mathcal{L}(\Omega))^d} + \sum_{i=1}^N \lambda_i\right)^{1/2}$. The dash black plots indicate the error associated with the mode truncation : $\left(\sum_{i=n+1}^N b_i^2\right)^{1/2}$. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the dashed red ones are computed with the smooth stochastic ROM (see chapters 6 and 7). From the ensemble simulated according to the full-stochastic POD-ROM, we have plotted the errors of one realization (yellow line); the ensemble bias (green line); 1.96 $\times$ the standard-deviation (shaded gray); the ensemble RMSE (magenta crosses); and the minimal distance to the reference (magenta line). The initial condition at $t = 0$ is common. The normalized error considering only the time mean velocity is about 0.93 and is thus not visible in the Figure.

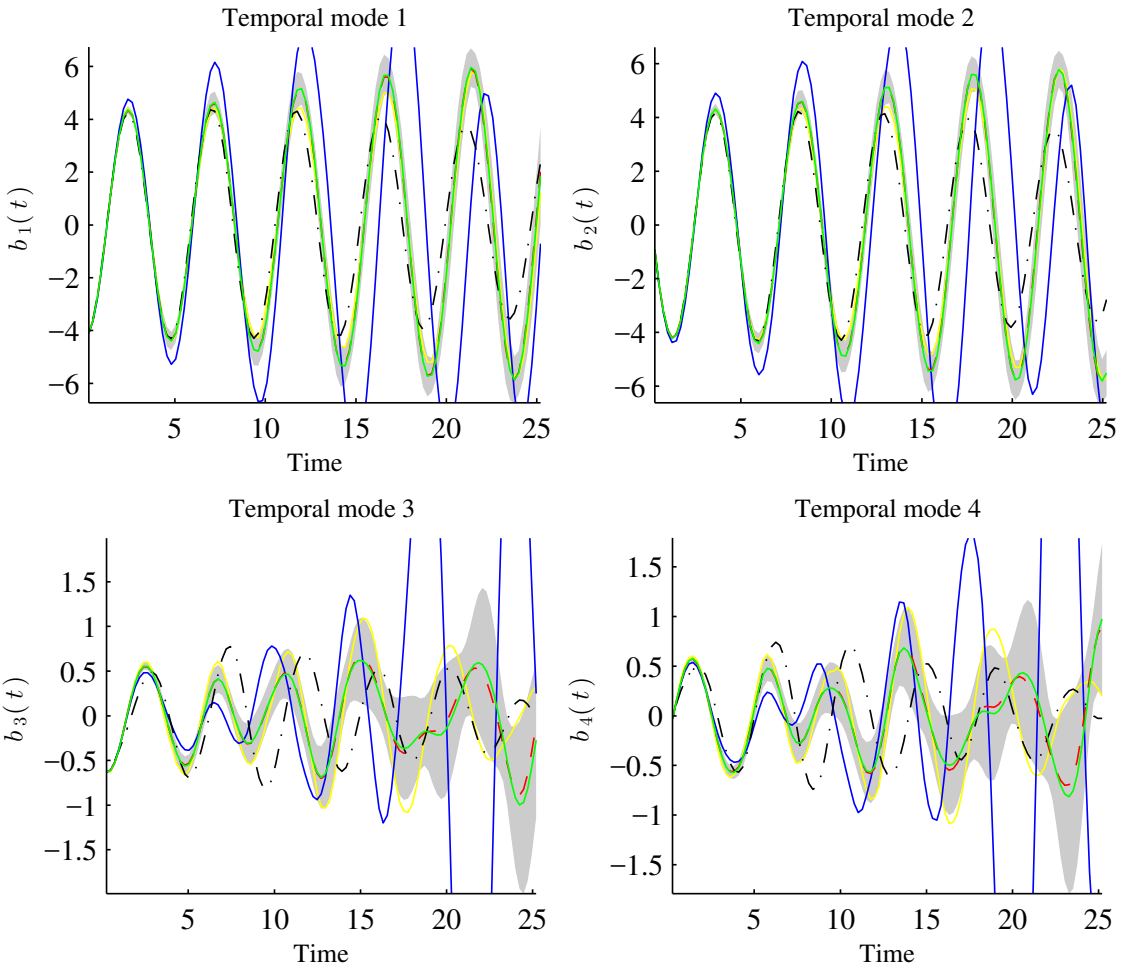


Figure 8.4: Forecast of the first 4 modes with for $n = 16$ modes of a wake flow at Reynolds 3900. The dash-dot black plots are the observed references. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the dashed red ones are computed with the smooth stochastic ROM (see chapters 6 and 7). From the ensemble simulated according to the full-stochastic POD-ROM, we have plotted one realization (yellow line), the ensemble mean (green line) and $+/-1.96 \times$ the standard-deviation (shaded gray). The initial condition at $t = 0$ is common to all methods.

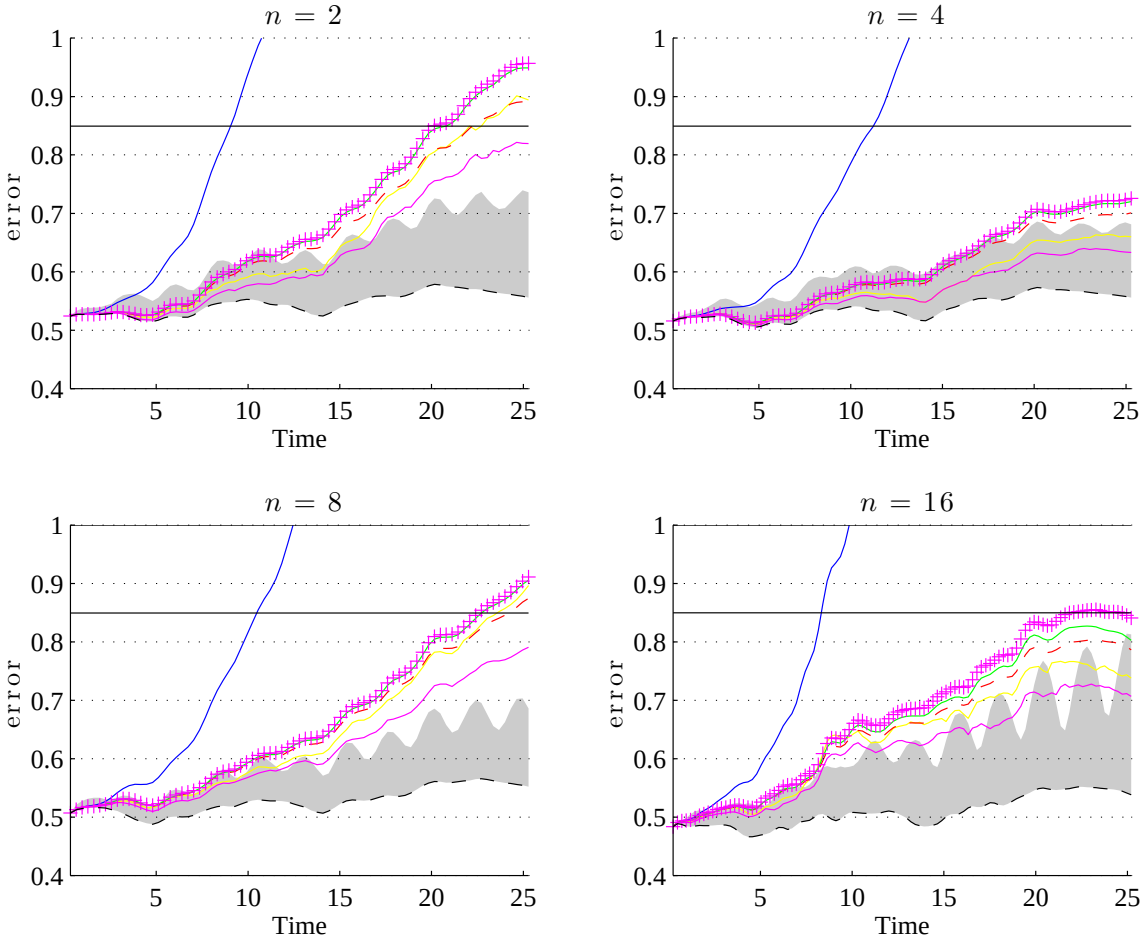


Figure 8.5: Normalized error for $n = 2$, 4, 8 and 16 modes of a wake flow at Reynolds 3900. The error is normalized by an approximation of the square root of the solution energy: $\left(\|\overline{\boldsymbol{v}}\|^2_{(\mathcal{L}(\Omega))^d} + \sum_{i=1}^N \lambda_i\right)^{1/2}$. The dash black plots indicate the error associated with the mode truncation : $\left(\sum_{i=n+1}^N b_i^2\right)^{1/2}$. The blue lines correspond to the solutions computed with a standard POD-Galerkin whereas the dashed red ones are computed with the smooth stochastic ROM (see chapters 6 and 7). From the ensemble simulated according to the full-stochastic POD-ROM, we have plotted the errors of one realization (yellow line); the ensemble bias (green line); 1.96 $\times$ the standard-deviation (shaded gray); the ensemble RMSE (magenta crosses); and the minimal distance to the reference (magenta line). The initial condition at $t = 0$ is common. The black solid line at the top is the error considering only the time mean velocity.

8.5 Conclusion

In this chapter, we have derived a random conservative Navier-Stokes representation based on the modeling under location uncertainty. After explaining its algebraic structure, we showed that this model conserves the energy of each realization. The energy dissipated by the inhomogeneous and anisotropic diffusion backscatters to the solution as variance through a skew-symmetric multiplicative noise. To derive a ROM, the solution is expressed in a finite dimensional basis. The energy fluxes between the resolved modes and toward the unresolved modes are maintained by the joint effect of noise and diffusion but the energy backscatterings from the unresolved modes toward the resolved modes are lost. This energy leak can be worked around. However, we preferred to keep this structure to enable the direct cascade toward the viscous scales. In the context of POD-ROM, a set of statistics have been derived to estimate the noises and diffusion parameters. These estimators rely on the physical structure of our stochastic Navier-Stokes model, the processing of the residual velocity – the velocity component not represented by the resolved modes – and, a partial fitting of one realization of the resolved modes' evolution. Despite their low computational costs, these statistics enable to estimate the influence of an unresolved inhomogeneous time-uncorrelated velocity. All these correlation parameters, the spatial modes and the coefficients of the ROM are estimated on a learning set derived from DNS simulations of wake flows at Reynolds 300 and 3900. Ensembles are then forecast from these stochastic ROMs. The UQ skills are excellent especially at Reynolds 300. The variability of stable modes is randomly maintained by the joint effect of diffusion and multiplicative noise. This keeps the ensemble very close to the reference. Furthermore, this random forcing naturally adapts the spreading to the amount of interactions that can be resolved deterministically. Indeed, the variance is larger, when energy can only be randomly transferred whereas the variance is lower when these fluxes can be resolved deterministically.

In the near future, we would like to compare the UQ skills of our POD-ROM on larger data sets with a deterministic POD-ROM with random initial conditions and other stochastic POD-ROMs. The needed ensemble size could also been discussed. Finally, a filtering procedure will probably be implemented.

Appendix

8.A Navier-Stokes model under location uncertainty

The following derivation is inspired of a proof of Mémin (2014) reproduced in chapter 2, originally introduced by Mikulevicius and Rozovskii (2004). Most conservation laws of mechanics state that a variation of an extensive quantity, $\int_{\mathcal{V}(t)} \rho\big(\Theta+\boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}\big)$, is due to some external actions. We explicitly make the associated specific property, $\Theta + \boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}$, appear to simplify the following development. Moreover, the component $\boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}$ expresses a possible time-uncorrelated component of this intensive quantity. For instance, the property $\Theta+\boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}$ could express the specific internal energy in the first principle or the velocity components in the conservation of linear momentum. All these balances can be expressed as:

$$\mathrm{d}\int_{\mathcal{V}(t)}\rho\big(\Theta + \boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}\big) = \int_{\mathcal{V}(t)}\mathrm{d}_t F_\Theta. \tag{8.A.1}$$

The left-hand term must be interpreted in a distribution sense, the small-scale velocity, $\boldsymbol{\sigma}_\Theta\dot{\boldsymbol{B}}$, being non-continuous. For every test function h, we have:

$$\int_{\mathbb{R}_+} h(t)\mathrm{d}\int_{\mathcal{V}(t)}\rho\Theta - \int_{\mathbb{R}_+}\frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)}\rho\boldsymbol{\sigma}_\Theta\mathrm{d}\boldsymbol{B}_t = \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}\mathrm{d}_t F_\Theta. \tag{8.A.2}$$

Both sides of this equation must have the same structure and the forces can be written as:

$$\int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}\mathrm{d}_t F_\Theta = -\int_{\mathbb{R}_+}\frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)}\rho\boldsymbol{\sigma}_\Theta\mathrm{d}\boldsymbol{B}_t + \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)}\left(\eta_\Theta\mathrm{d}t + \boldsymbol{\gamma}_\Theta\mathrm{d}\boldsymbol{B}_t\right). \tag{8.A.3}$$

The right-hand first term must compensate for the white-noise distributional differentiation of (8.A.2), whereas the last term of (8.A.3) provides the structure of the forces under location uncertainty.
The transport equation applied to the $\rho\Theta$ and the continuity equation $D_t\rho + \rho\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star \mathrm{d}t = 0$ give:

$$\begin{aligned}
\mathrm{d}\int_{\mathcal{V}(t)} \rho\Theta &= \int_{\mathcal{V}(t)} \left(D_t\left(\rho\Theta\right) + \rho\Theta\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star \mathrm{d}t\right), && (8.A.4)\\
&= \int_{\mathcal{V}(t)} \left(\rho D_t\Theta + D_t\rho\Theta + \mathrm{d}\left\langle \int_0^t D_{t'}\rho, \int_0^t D_{t'}\Theta \right\rangle + \rho\Theta\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star \mathrm{d}t\right), && (8.A.5)\\
&= \int_{\mathcal{V}(t)} \rho D_t\Theta, && (8.A.6)
\end{aligned}$$

where all quadratic cross-variation evaluated in Lagrangian coordinates is zero. This result being true for any volume $\mathcal{V}(t)$, we can remove the integral and use the previous results (8.A.2) and (8.A.3) :

$$D_t\Theta = \frac{1}{\rho}\eta_\Theta \mathrm{d}t + \frac{1}{\rho}\boldsymbol{\gamma}_\Theta \mathrm{d}\boldsymbol{B}_t. \qquad (8.A.7)$$

Then the formula (2.2.10) of chapter 2 yields:

$$\mathbb{D}_t\Theta = -\mathrm{tr}\left(\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)\left(\frac{1}{\rho}\boldsymbol{\gamma}_\Theta\right)\right)\mathrm{d}t + \frac{1}{\rho}\eta_\Theta \mathrm{d}t + \frac{1}{\rho}\boldsymbol{\gamma}_\Theta \mathrm{d}\boldsymbol{B}_t. \qquad (8.A.8)$$

Then, we apply this general relation to each component of the velocity $\Theta + \boldsymbol{\sigma}_\Theta \mathrm{d}\boldsymbol{B}_t = w_i + \boldsymbol{\sigma}_{i\bullet}\mathrm{d}\boldsymbol{B}_t$. The pressure is decomposed into a component continuous in time p and time-uncorrelated component $\dot{p}_\sigma = \frac{\mathrm{d}_t p_\sigma}{\mathrm{d}t}$. Accordingly the forcings read:

$$\left\{\begin{array}{lll}
\eta_{w_i} &=& -\rho(\boldsymbol{f}\times\boldsymbol{w})_i + \rho g\delta_{i3} - \partial_i p + \rho\nu\Delta w_i,\\
\boldsymbol{\gamma}_{w_i} &=& -\rho(\boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_i - \partial_i \mathrm{d}_t p_\sigma + \rho\nu\Delta(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_i.
\end{array}\right. \qquad (8.A.9)$$

Then, from (8.A.8) the stochastic Navier-Stokes model under location uncertainty is:

$$\begin{aligned}
\mathbb{D}_t\boldsymbol{w} + \boldsymbol{f}\times(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \boldsymbol{g}\mathrm{d}t - \frac{1}{\rho}\boldsymbol{\nabla}(p\mathrm{d}t + \mathrm{d}_t p_\sigma) + \nu\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\\
- \mathrm{tr}\left(\left(\boldsymbol{\sigma}^T\boldsymbol{\nabla}\right)\left(-\boldsymbol{f}\times\boldsymbol{\sigma} + \nu\Delta\boldsymbol{\sigma}\right)\right)\mathrm{d}t + \mathrm{d}_t\left\langle \int_0^t \left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\cdot\boldsymbol{\nabla}\right), \int_0^t \frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}_{t'}p_\sigma \right\rangle. \qquad (8.A.10)
\end{aligned}$$

In this paper, we neglect the time-uncorrelated pressure $\dot{p}_\sigma$ and all the correlations of the second line. After this, the Navier-Stokes representation reads:

$$\mathbb{D}_t\boldsymbol{w} + \boldsymbol{f}\times(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \boldsymbol{g}\mathrm{d}t - \frac{1}{\rho}\boldsymbol{\nabla}p\mathrm{d}t + \nu\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t). \qquad (8.A.11)$$

8.B Energy dissipation

Hereafter, we neglect the boundary conditions. This assumption which is valid for wake flows, if the spatial domain is sufficiently large, enables us to proceed easily to integrations by parts. Moreover, we assume that the incompressibility conditions $\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star = 0$ and $\boldsymbol{\nabla}\cdot\boldsymbol{\sigma} = 0$ hold. We first apply the Itō formula to the local kinetic energy to obtain

$$\mathrm{d}\left(\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d}\right) = \int_\Omega \left(\mathrm{d}_t\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}\mathrm{d}_t < \boldsymbol{w}^T, \boldsymbol{w} >\right). \qquad (8.B.1)$$

Using the projected Navier-Stokes model (8.2.4) without viscosity, weight and pressure

$$\mathrm{d}_t\boldsymbol{w} = \Pi_\phi\left[(\mathrm{d}\mathbb{M})(\boldsymbol{w})\right] = \Pi_\phi\left[\left(F(\boldsymbol{w}) + C(\boldsymbol{w},\boldsymbol{w})\right)\mathrm{d}t + (\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w})\right], \qquad (8.B.2)$$

the Itō term is straightforward:

$$\int_\Omega \mathrm{d}_t < \boldsymbol{w}^T, \boldsymbol{w} > \;=\; \sum_k \left\|\Pi_\phi\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\right\|^2_{(\mathcal{L}^2(\Omega))^d} \mathrm{d}t. \tag{8.B.3}$$

From the definition (8.2.5) of the projection operator Π_ϕ, we can express the above quantity with the ROM coefficients:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega < \boldsymbol{w}^T, \boldsymbol{w} > &= \sum_k \sum_{p=1}^n \left(\int_\Omega \boldsymbol{\phi}_p \cdot (\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right)^2, && (8.B.4)\\
&= \sum_{p=1}^n \sum_k \left(\sum_{q=1}^n b_q \int_\Omega \boldsymbol{\phi}_p \cdot (\boldsymbol{G}_{\bullet k})(\boldsymbol{\phi}_q)\right)^2, && (8.B.5)\\
&= \boldsymbol{b}^T \left(\frac{1}{2}\sum_{p=1}^n \boldsymbol{\alpha}_{\bullet p \bullet}\boldsymbol{\alpha}^T_{\bullet p \bullet}\right)\boldsymbol{b}. && (8.B.6)
\end{aligned}$$

By (8.B.1) again, the first term of the energy budget (8.B.1) reads

$$\int_\Omega \mathrm{d}_t \boldsymbol{w}^T \boldsymbol{w} \;=\; \int_\Omega \Pi_\phi\left[(\mathrm{d}\mathbb{M})(\boldsymbol{w})\right] \cdot \boldsymbol{w}. \tag{8.B.7}$$

To remove the projection operator, Π_ϕ, we first exploit the symmetry of the projection operator and then the fact that $\boldsymbol{w}$ is already in the reduced subspace:

$$\begin{aligned}
\int_\Omega \mathrm{d}_t \boldsymbol{w}^T \boldsymbol{w} &= \int_\Omega (\mathrm{d}\mathbb{M})(\boldsymbol{w}) \cdot \Pi_\phi\left[\boldsymbol{w}\right], && (8.B.8)\\
&= \int_\Omega (\mathrm{d}\mathbb{M})(\boldsymbol{w}) \cdot \boldsymbol{w}, && (8.B.9)\\
&= \int_\Omega \left(F(\boldsymbol{w})\mathrm{d}t + C(\boldsymbol{w},\boldsymbol{w})\mathrm{d}t + (\boldsymbol{G}\mathrm{d}\boldsymbol{B}_t)(\boldsymbol{w})\right)^T \boldsymbol{w}, && (8.B.10)\\
&= \int_\Omega \left(\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{a}}{2}\boldsymbol{\nabla}\boldsymbol{w}^T\right)\mathrm{d}t + \left((\boldsymbol{w}^\star + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla}\right)\boldsymbol{w}^T\right)\boldsymbol{w}, && (8.B.11)\\
&= -\frac{1}{2}\int_\Omega \sum_{q=1}^d \left(\boldsymbol{\nabla} w_q\right)^T \boldsymbol{a}\boldsymbol{\nabla} w_q \mathrm{d}t + \frac{1}{2}\int_\Omega \underbrace{\boldsymbol{\nabla}\cdot(\boldsymbol{w}^\star\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)}_{=0} \|\boldsymbol{w}\|_2^2, && (8.B.12)\\
&= -\frac{1}{2}\int_\Omega \sum_k \left\|\left(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla}\right)\boldsymbol{w}\right\|^2 \mathrm{d}t, && (8.B.13)\\
&= -\frac{1}{2}\sum_k \left\|\left(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla}\right)\boldsymbol{w}\right\|^2_{(\mathcal{L}^2(\Omega))^d} \mathrm{d}t. && (8.B.14)
\end{aligned}$$

Then, we decompose the multiplicative noise into two components: one in the reduced subspace and one in the orthogonal of the reduced subspace.

$$\left(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla}\right)\boldsymbol{w} = (\boldsymbol{G}_{\bullet k})(\boldsymbol{w}) = \Pi_\phi\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right] + \Pi_\phi^\perp\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right], \tag{8.B.15}$$

where $\Pi_\phi^\perp = \mathbb{I}_d - \Pi_\phi$ is the projector on the orthogonal of the reduced subspace. By orthogonality, the norm (8.B.14) can be split as follows:

$$\begin{aligned}
\int_\Omega \mathrm{d}_t \boldsymbol{w}^T \boldsymbol{w} \;=\; & -\frac{1}{2}\sum_k \left\|\Pi_\phi\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\right\|^2_{(\mathcal{L}^2(\Omega))^d} \mathrm{d}t\\
& -\frac{1}{2}\sum_k \left\|\Pi_\phi^\perp\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\right\|^2_{(\mathcal{L}^2(\Omega))^d} \mathrm{d}t. && (8.B.16)
\end{aligned}$$

Finally, using the expression of the Itō term (8.B.3), the kinetic energy budget (8.B.1) simplifies to:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\boldsymbol{w}\|^2_{(\mathcal{L}^2(\Omega))^d}\right) \;=\; -\frac{1}{2}\sum_k \left\|\Pi_\phi^\perp\left[(\boldsymbol{G}_{\bullet k})(\boldsymbol{w})\right]\right\|^2_{(\mathcal{L}^2(\Omega))^d} < 0. \tag{8.B.17}$$

For a thorough mathematical analysis, we can precise the meaning of the norm appearing in the right-hand sides of (8.B.3), (8.B.6), (8.B.14), (8.B.16) and (8.B.17). The functions $\boldsymbol{B} \mapsto \boldsymbol{\sigma}\boldsymbol{B}$, $\boldsymbol{B} \mapsto (\boldsymbol{G}\boldsymbol{B})(\boldsymbol{h})$ and $\boldsymbol{B} \mapsto (\boldsymbol{H}\boldsymbol{B})$ are infinite-dimensional operators from $\left(\mathcal{L}^2(\Omega)\right)^d$ to $\left(\mathcal{L}^2(\Omega)\right)^d$ (for a function $\boldsymbol{h}$ of $\left(\mathcal{L}^2(\Omega)\right)^d$ smooth enough). Accordingly, the right-hand sides of (8.B.3), (8.B.14), (8.B.16) and (8.B.17) are Hilbert-Schmidt norms. The infinite-dimensional operators $\boldsymbol{B} \mapsto (\boldsymbol{\alpha}\boldsymbol{B})\boldsymbol{V}$, $\boldsymbol{B} \mapsto (\boldsymbol{\theta}\boldsymbol{B})$ map $\left(\mathcal{L}^2(\Omega)\right)^d$ to $\mathbb{R}^n$ (for any vector $\boldsymbol{V} \in \mathbb{R}^{n+1}$). So, the right-hand sides of (8.B.6) are canonical norms of $(\mathcal{L}^2(\Omega))^{n+1}$.

8.C Estimation formulas

We recall that:

$$\mathrm{d}b_i = \left(i_i + \left(\breve{\boldsymbol{f}}_{\bullet i} + \boldsymbol{l}_{\bullet i}\right)^T \boldsymbol{b} + \boldsymbol{b}^T \boldsymbol{c}_{\bullet\bullet i}\boldsymbol{b}\right)\mathrm{d}t + \left(\boldsymbol{\alpha}_{\bullet i\bullet}\mathrm{d}\boldsymbol{B}_t\right)^T \boldsymbol{b} + \left(\boldsymbol{\theta}_{i\bullet}\mathrm{d}\boldsymbol{B}_t\right). \tag{8.C.1}$$

Therefore, the constraints $\int_0^T b_k = \delta_{k0}$ lead to:

$$\begin{aligned}
< b_i, \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) > &= \sum_{k=1}^n < \int_0^t (\boldsymbol{\alpha}_{ki\bullet}\mathrm{d}\boldsymbol{B}_s)\, b_k, \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) > \\
&\quad + < \int_0^t ((\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_s), \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) >, && (8.C.2)\\
&= \sum_{k=1}^n \left(\int_0^T b_k\right) \boldsymbol{\alpha}_{ki\bullet} \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) + T(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet}) \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}), && (8.C.3)\\
&= T(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet}) \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}). && (8.C.4)
\end{aligned}$$

We denote $\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \triangleq \boldsymbol{\sigma}(\boldsymbol{x})\mathrm{d}\boldsymbol{B}_t$. Its realizations are the Eulerian residual velocities up to the time step $(\boldsymbol{v}(\boldsymbol{x},t) - \boldsymbol{w}(\boldsymbol{x},t))\Delta t$. Then the definition of $((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s)$ and the properties of quadratic covariations give the first type of estimator:

$$\begin{aligned}
(\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})(\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) &= \frac{1}{T}\int_0^T \mathrm{d} < b_i, \int_0^t (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s >, && (8.C.5)\\
&= \frac{1}{T}\int_0^T \mathrm{d} < b_i, \left[\int_\Omega \boldsymbol{\phi}_{j\bullet} \cdot \left(-\left(\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla}\right)\boldsymbol{\phi}_0 + \nu\Delta\tilde{\boldsymbol{X}}^{\boldsymbol{x}}\right)\right] >, && (8.C.6)\\
&= \frac{1}{T}\mathbb{P} - \lim_{\Delta t\to 0}\sum_{t_k=0}^T (\mathrm{d}b_i)(t_k)\left[\int_\Omega \boldsymbol{\phi}_j \cdot \left(-\left(\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla}\right)\boldsymbol{\phi}_0 + \nu\Delta\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}}\right)\right](t_k) && (8.C.7)\\
&= K_{j0}\left[\frac{1}{T}\mathbb{P} - \lim_{\Delta t\to 0}\sum_{t_k=0}^T (\mathrm{d}b_i)(t_k)\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}}(t_k)\right], && (8.C.8)
\end{aligned}$$

where we used the linearity of the operator:

$$K_{j0}[\boldsymbol{\xi}] \triangleq \int_\Omega \boldsymbol{\phi}_j \cdot \left(-\left(\boldsymbol{\xi} \cdot \boldsymbol{\nabla}\right)\boldsymbol{\phi}_0 + \nu\Delta\boldsymbol{\xi}\right). \tag{8.C.9}$$

Again by the ROM (8.C.1), the orthogonality of the *chronos* yields:

$$\begin{aligned}
\int_0^T b_p \mathrm{d} < b_i, \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) > &= \int_0^T \sum_{k=1}^n b_p \mathrm{d} < \int_0^t (\boldsymbol{\alpha}_{ki\bullet}\mathrm{d}\boldsymbol{B}_s)\, b_k, \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) > \\
&\quad + \int_0^T b_p \mathrm{d} < \int_0^t ((\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet})\mathrm{d}\boldsymbol{B}_s)\,, \int_0^t ((\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})\mathrm{d}\boldsymbol{B}_s) >, \qquad (8.C.10) \\
&= \sum_{k=1}^n \left(\int_0^T b_p b_k \right) \boldsymbol{\alpha}_{ki\bullet} \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) \\
&\quad + \left(\int_0^T b_p \right) (\boldsymbol{\theta}_{i\bullet} + \boldsymbol{\alpha}_{0i\bullet}) \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}), \qquad (8.C.11) \\
&= T\lambda_p \boldsymbol{\alpha}_{pi\bullet} \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}). \qquad (8.C.12)
\end{aligned}$$

This results leads to the second type of estimator:

$$\begin{aligned}
\boldsymbol{\alpha}_{pi\bullet} \cdot (\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet}) &= \frac{1}{\lambda_p T} \int_0^T b_p \mathrm{d} < b_i, \left[\int_\Omega \boldsymbol{\phi}_j \cdot \left(- \left(\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_0 + \nu \Delta \tilde{\boldsymbol{X}}^{\boldsymbol{x}} \right) \right] >, \qquad (8.C.13) \\
&= \frac{1}{\lambda_p T} \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_k=0}^T b_p(t_k)(\mathrm{d}b_i)(t_k) \left[\int_\Omega \boldsymbol{\phi}_j \cdot \left(- \left(\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_0 + \nu \Delta \mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \right) \right] (t_k), \qquad (8.C.14) \\
&= K_{j0} \left[\frac{1}{\lambda_p T} \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_k=0}^T b_p(t_k)(\mathrm{d}b_i)(t_k) \mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \right]. \qquad (8.C.15)
\end{aligned}$$

The last estimator derivation is similar replacing $(\boldsymbol{\theta}_{j\bullet} + \boldsymbol{\alpha}_{0j\bullet})$ above by $\boldsymbol{\alpha}_{qj\bullet}$:

$$\begin{aligned}
\boldsymbol{\alpha}_{pi\bullet} \cdot \boldsymbol{\alpha}_{qj\bullet} &= -\frac{1}{\lambda_p T} \int_0^T b_p \mathrm{d} < b_i, \left[\int_\Omega \boldsymbol{\phi}_j \cdot \left(\left(\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_q \right) \right] >, \qquad (8.C.16) \\
&= -\frac{1}{\lambda_p T} \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_k=0}^T b_p(t_k)(\mathrm{d}b_i)(t_k) \left[\int_\Omega \boldsymbol{\phi}_j \cdot \left(\left(\mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \cdot \boldsymbol{\nabla} \right) \boldsymbol{\phi}_q \right) \right] (t_k), \qquad (8.C.17) \\
&= K_{jp} \left[-\frac{1}{\lambda_p T} \mathbb{P} - \lim_{\Delta t \to 0} \sum_{t_k=0}^T b_p(t_k)(\mathrm{d}b_i)(t_k) \mathrm{d}\tilde{\boldsymbol{X}}^{\boldsymbol{x}} \right], \qquad (8.C.18)
\end{aligned}$$

where

$$K_{jq}[\boldsymbol{\xi}] \triangleq -\int_\Omega \boldsymbol{\phi}_j \cdot \left((\boldsymbol{\xi} \cdot \boldsymbol{\nabla}) \boldsymbol{\phi}_q \right). \qquad (8.C.19)$$

Chapter 9

Effects of smooth flows on tracer gradients and tracer spectra

Abstract

During a finite-time advection, the norm of tracer gradients can increase or decrease, depending on the angle between flow gradients and initial tracer gradients. When the correlation between these two fields is weak – either because the tracer has a negligible back effect on the flow or because a strong spatial smoothing has been performed – the averaged squared norm of tracer gradients can only increase. Moreover, the local growth rate of the tracer gradients is independent of the initial tracer distribution. As presented, this growth rate is directly related to FTLE and mesochronic velocity. A simple model is then proposed to locally and globally describe the time evolution of the growth rate in the case of a stationary or slowly time evolving Eulerian velocity field. The key processes are locally uniform shears and foldings around stationary vortices. Finally, the squared norm mean of tracer gradients controls and specifies the time evolution of the spectral tail and its slope. Accordingly, the advection time and low-pass filter width can be determined for practical applications using a Lagrangian advection method. This analysis suggests a practical eddy diffusivity parameterization. Numerical experiments on a toy model and using satellite data illustrate these developments.

9.1 Introduction

Since the first images from space, the attention of both theoreticians and remote sensing scientists has been triggered by the abundance of various ocean tracer patterns and signatures in the mesoscale and sub-mesoscale (1-50 km) ranges (e.g. Gower et al., 1980; Lesieur and Sardouny, 1981). To date, global direct quantification of horizontal dispersion and mixing at such scales is still not available. Yet, from precise measurements of the ocean topography and its related dynamics, significant progress has been made. Nowadays, combined satellite altimeter measurements satisfactorily detail the large scale ocean dynamics. The ocean's mesoscale (10-50 km) and submesoscale ($<$ 10 km) variability and energy are still more challenging to map with conventional radar altimeters, mainly because of the narrow illuminated swath of each instrument, regardless of the orbital configuration (Dufau et al., 2016).

Nonetheless, a now-common strategy is to derive small-scale tracer structures and so-called Lagrangian coherent structures, from the available smooth altimetry-derived velocities (e.g. Price et al., 2006; Lehahn et al., 2007). Indeed, using a Lagrangian-dynamical framework, an initial larger-scale tracer field can be advected on higher-resolution grids, generating much smaller-scale patterns (Aref, 1984; Pierrehumbert and Yang, 1993). Typical moderate- to large-scale ocean cyclonic and anti-cyclonic eddies trap and advect fluid parcels over weeks to months. As pictured, with time, these fluid parcels with different origins, temperature, salinity and possibly different biogeochemical properties and/or contaminant loadings, come closer, to sharpen fronts but also to possibly dilute their properties, and promote transformative chemical reactions. To help trigger

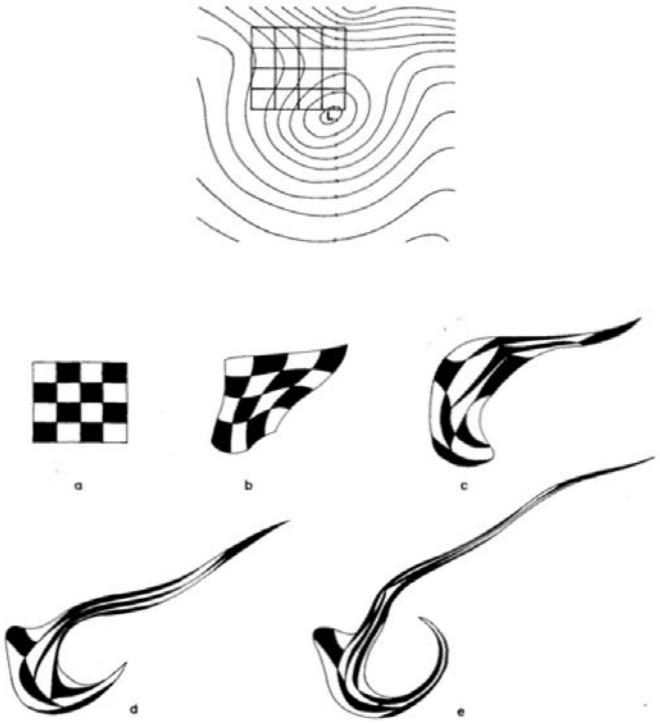


Figure 9.1: Deformation of an air layer in the atmosphere after 6 hours, 12 hours, 24 hours and 36 hours respectively. A simple barotropic model is used to simulate the flow (Welander, 1955).

these processes, stirring effect first characterizes the development of elongated structures, well illustrated by Welander (1955), using a simple velocity field to produce spectacular distortions, Figure 9.1. Initial square elements, small compared to the length scale of the deforming flow field, become subject to translation, rotation and shearing. With time, deformation is significant. Increasingly long and thin filaments wrap around the eddy, and possibly fold. Folds then appear where the velocity gradient is perpendicular to the stream direction. Accordingly, at a given scale of observations, mixing can be associated with processes that act to minimize filament thinning and dilute sharp differences (gradients).

After a long advection by a smooth and slowly-varying velocity, the expected growth of passive tracer gradients can be theoretically obtained, as well as its related spectral tail evolution. From this analysis, local and global diagnoses of stretching and folding can further be developed. From a practical point of view, using estimated velocities from altimeter-derived sea surface height (SSH) measurements, the initial tracer field, sea surface temperature (SST) or salinity (SSS), must thus be low-pass filtered prior to Lagrangian-advection operations. It can then be expected that the time of advection and the low-pass filter bandwidth are directly linked. Following the proposed development, an exact relation can indeed be determined. Such an analysis then provides a rigorous explanation to the heuristic choices used in Dencausse et al. (2014). This can further be compared to values that can be inferred from the knowledge of the Rossby deformation radius or the mean squared vorticty (Berti and Lapeyre, 2014).

The section 2 describes the tracer gradients after a finite-time advection, and the associated stretching and folding diagnoses. In section 3, the averaged tracer gradient norm is shown to control the tracer spectrum tail. Tools are provided to monitor the aforementioned Lagrangian advection method. Based on the proposed developments, section 4 rapidly discusses subgrid parametrizations of large-scale flow simulations.

9.2 Time evolution for the tracer gradient norm

In this section, we propose exact and approximate results to describe the mean of the gradient squared norm of an advected tracer T:

$$\overline{\|\boldsymbol{\nabla} T\|^2}, \tag{9.2.1}$$

where the averaging operator $\overline{\bullet}$ is defined for every function f as follows:

$$\overline{f} \;=\; \frac{1}{S}\,\mathbb{E}\int_{\Omega} f, \tag{9.2.2}$$

$\Omega \subset \mathbb{R}^2$ is the spatial domain and $S < \infty$ its area. Even though we will mainly focus on deterministic dynamics, the expectation, $\mathbb{E}$, of the above formula enables us to directly generalize our results to random cases.

9.2.1 Exact flow properties

Given a possibly random velocity field $\boldsymbol{v}$, the flow $\boldsymbol{x}_0 \mapsto \boldsymbol{\phi}(\boldsymbol{x}_0)$ is defined as:

$$\boldsymbol{\phi}(\boldsymbol{x}_0) = \boldsymbol{\phi}(\boldsymbol{x}_0, t) = \boldsymbol{x}_0 + \int_0^t \mathrm{d}t' \boldsymbol{v}(\boldsymbol{\phi}(\boldsymbol{x}_0, t'), t'). \tag{9.2.3}$$

For a divergence-free velocity, $\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0$, we have $\det(\boldsymbol{\nabla}\boldsymbol{\phi}^T) = 1$, where $\boldsymbol{\nabla}\boldsymbol{\phi}^T$ is the spatial gradient tensor of the flow. Subsequently, the right Cauchy-Green deformation tensor, $\boldsymbol{\nabla}\boldsymbol{\phi}^T\left(\boldsymbol{\nabla}\boldsymbol{\phi}^T\right)^T$, and its inverse, shall have two real and identical strictly positive eigenvalues. Only the stable direction, corresponding to the eigenvector associated with the eigenvalue smaller than 1, and the unstable direction, corresponding to the eigenvector associated with the eigenvalue larger than 1, are switched. Along the stable (resp. unstable) direction, the distance between two points decreases (resp. increases). More precisely, a matrix diagonalization leads to:

$$\boldsymbol{\nabla}\boldsymbol{\phi}^T\left(\boldsymbol{\nabla}\boldsymbol{\phi}^T\right)^T \;=\; \boldsymbol{P}^T\boldsymbol{D}^{-1}\boldsymbol{P} \;\text{ with }\; D_{ii}^{-1} = 1 + \alpha^2\left(1 - (-1)^i\frac{\beta}{\alpha}\right), \tag{9.2.4}$$

where $\boldsymbol{P}$ is an orthogonal matrix, $\alpha^2(\boldsymbol{x}_0) = \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{\phi}^T\|^2 - 1 \geqslant 0$ and $\beta^2(\boldsymbol{x}_0) = \alpha^2 + 2$, using the Frobenius matrix norm. The eigenvalues D_{ii} define the Finite Time Lyapunov Exponents (FTLE) (Pierrehumbert and Yang, 1993; Haller and Yuan, 2000; Thiffeault and Boozer, 2001; Haller, 2005; Haller and Sapsis, 2011). The largest and the smallest FTLEs are:

$$\Lambda = \frac{1}{2t}\log(D_{11}^{-1}) \;\text{ and }\; -\Lambda = \frac{1}{2t}\log(D_{22}^{-1}). \tag{9.2.5}$$

In particular, when both the largest FTLE, Λ, and the time, t, are large, the term α^2 is large and $\beta/\alpha = \sqrt{1 + \frac{2}{\alpha^2}}$ is small. This leads to the approximation:

$$\Lambda \approx \frac{1}{t}\log(\alpha), \tag{9.2.6}$$

Therefore, the FTLE ridges – often considered as proxies of mixing barriers – coincide with the α ridges. The Cauchy-Green tensor encodes insightful information on the flow. For instance, it rules the tracer gradient norm evolution. Noting that at a given time, the transported tracer gradient, $\boldsymbol{\nabla}T$, can be written from the initial gradient field, $\boldsymbol{\nabla}T_0$, as:

$$\boldsymbol{\nabla}T(\boldsymbol{x}) = \boldsymbol{\nabla}(T_0(\boldsymbol{\phi}^{-1}(\boldsymbol{x}))) = [\boldsymbol{\nabla}\boldsymbol{\phi}^T]^{-1}(\boldsymbol{\phi}^{-1}(\boldsymbol{x}))\boldsymbol{\nabla}T_0(\boldsymbol{\phi}^{-1}(\boldsymbol{x})). \tag{9.2.7}$$

Then, using incompressiblity in the variable change and the matrix diagonalization (9.2.4), we get an exact expression of the averaged squared norm of tracer gradients:

$$\begin{aligned}
\overline{\|\boldsymbol{\nabla} T\|^2} - \overline{\|\boldsymbol{\nabla} T_0\|^2} &= \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\,\|\boldsymbol{\nabla} T(\boldsymbol{x})\|^2 - \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}_0\,\|\boldsymbol{\nabla} T_0(\boldsymbol{x}_0)\|^2\,, && (9.2.8)\\
&= \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}_0\,\left\|[\boldsymbol{\nabla}\boldsymbol{\phi}^T]^{-1}(\boldsymbol{x}_0)\boldsymbol{\nabla} T_0(\boldsymbol{x}_0)\right\|^2 - \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}_0\,\|\boldsymbol{\nabla} T_0(\boldsymbol{x}_0)\|^2\,, && (9.2.9)\\
&= \overline{(\boldsymbol{\nabla} T_0)^T\left(\left[\boldsymbol{\nabla}\boldsymbol{\phi}^T\,(\boldsymbol{\nabla}\boldsymbol{\phi}^T)^T\right]^{-1} - \mathbb{I}_d\right)\boldsymbol{\nabla} T_0}, && (9.2.10)\\
&= \overline{\alpha^2\left(\underbrace{\left(1+\frac{\beta}{\alpha}\right)}_{>0}\left(\boldsymbol{P}^T\boldsymbol{\nabla} T_0\right)_2^2 + \underbrace{\left(1-\frac{\beta}{\alpha}\right)}_{<0}\left(\boldsymbol{P}^T\boldsymbol{\nabla} T_0\right)_1^2\right)}, && (9.2.11)
\end{aligned}$$

where $\left(\boldsymbol{P}^T\boldsymbol{\nabla} T_0\right)_i$ is the i-th component of the vector $\boldsymbol{P}^T\boldsymbol{\nabla} T_0$. Thus, the Cauchy-Green tensor and the initial tracer gradient completely determine the averaged squared norm of advected tracer gradients (equation (9.2.10)). To simplify the above expression, we define the angle between the tracer gradient and the compressive (stable) direction of the direct flow as:

$$\cos(\theta^{\boldsymbol{\phi}}_{T_0}) = \frac{\left(\boldsymbol{P}^T\boldsymbol{\nabla} T_0\right)_2}{\|\boldsymbol{\nabla} T_0\|_2}. \qquad (9.2.12)$$

Finally, we infer the following compact expression:

$$\overline{\|\boldsymbol{\nabla} T\|^2} - \overline{\|\boldsymbol{\nabla} T_0\|^2} = \overline{\|\boldsymbol{\nabla} T_0\|^2\,\alpha^2\left(1+\frac{\beta}{\alpha}\cos\left(2\theta^{\boldsymbol{\phi}}_{T_0}\right)\right)}. \qquad (9.2.13)$$

The advection globally increases (decreases) the tracer gradient norm if the initial tracer gradient is locally close enough to the stable (unstable) direction of the direct flow. This corresponds to $\theta^{\boldsymbol{\phi}}_{T_0}$ close to $0[\pi]$ or $\frac{\pi}{2}[\pi]$ respectively. This is modulated by the initial amplitude of the tracer gradients, α^2 and $\frac{\beta}{\alpha} = \sqrt{1+\frac{2}{\alpha^2}} \geqslant 1$. It should be noticed that α and β do not explicitly depend on the tracer.

9.2.2 Decorrelation approximations

Over the space, the angle $\theta^{\boldsymbol{\phi}}_{T_0}$ takes different values. If the flow gradients and the initial tracer gradients are not correlated and are oriented along various angles over the space, the variance of $\theta^{\boldsymbol{\phi}}_{T_0}$ will likely be large (*i.e.* close to 2π). Then, due to the overlapping, the distribution of $2\theta^{\boldsymbol{\phi}}_{T_0}[2\pi]$ over the space becomes close to an uniform law on $[0, 2\pi]$. The average over the space of the term $\cos\left(2\theta^{\boldsymbol{\phi}}_{T_0}\right)$ will become close to zero, and

$$\overline{\|\boldsymbol{\nabla} T\|^2} - \overline{\|\boldsymbol{\nabla} T_0\|^2} \approx \overline{\|\boldsymbol{\nabla} T_0\|^2\,\alpha^2} > 0. \qquad (9.2.14)$$

On average, the tracer gradients will thus always increase by stretching. Figure 9.1 illustrates well the process. The tracer is completely passive. It can be a dye or an oil spill introduced at time $t = 0$. In this case, the tracer and the flow are completely decorrelated and the initial structure of the tracer is quickly stretched and folded to fill a broad range of scales. In contrast, geophysical tracers are generally correlated to the flow. This correlation ensues from the previous advection (from time $t = -\infty$ until $t = 0$) which had led to the initial tracer T_0 and from the non-linear effects. It is expressed by the angle $\theta^{\boldsymbol{\phi}}_{T_0}$ in the right-hand-side integrand of equation (9.2.13) which can be locally positive or negative. Accordingly, that correlation restricts or reduces locally enhanced strong gradients. In the Lagrangian advection method (Berti and Lapeyre, 2014; Dencausse et al., 2014), the tracers – Sea Surface Temperature (SST) or Sea Surface Salinity (SSS) – are not passive.

The flow and the tracers are correlated. This is the reason of the preservation of the tracers' very-large-scale structures such as the background meridional gradient. However, at mesoscales and submesoscales, the initial tracer and the flow have been decorrelated by the filtering. So, our result (9.2.14) predicts a strengthening of mesoscale and submesoscale tracer gradients. This is exactly what is happening in the works of Berti and Lapeyre (2014) and Dencausse et al. (2014).

By assuming again no correlation between the initial tracer and the flow, the formula (9.2.14) can be further approximated by:

$$\frac{\overline{\|\boldsymbol{\nabla} T\|^2}}{\overline{\|\boldsymbol{\nabla} T_0\|^2}} \approx 1 + \overline{\alpha^2}. \tag{9.2.15}$$

Before modeling the time dependance of the averaged growth rate, $\overline{\alpha^2}$, we describe the links between the mixing diagnostic introduced by Mezić et al. (2010) and our exact result (9.2.13).

9.2.3 Link with the mixing criterion of Mezić et al. (2010)

As in Mezić et al. (2010), the mesochronic velocity is defined as the velocity time-averaged along a trajectory:

$$\breve{\boldsymbol{v}}(\boldsymbol{x}_0, t) \triangleq \frac{1}{t}\int_0^t \mathrm{d}t' \boldsymbol{V}(\boldsymbol{x}_0, t') = \frac{\boldsymbol{\phi}(\boldsymbol{x}_0) - \boldsymbol{x}_0}{t}, \tag{9.2.16}$$

where $\boldsymbol{V}$ is the Lagrangian velocity. As derived by the authors, the incompressibility of the flow yields:

$$1 = \det(\boldsymbol{\nabla}\boldsymbol{\phi}^T) = \det(\mathbb{I}_d + t\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T) = 1 + t\ \mathrm{tr}\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right) + t^2 \det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right). \tag{9.2.17}$$

The mesochronic velocity is thus an incompressible flow with the following structure:

$$t\ \det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right) = -\mathrm{tr}(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T) = -\boldsymbol{\nabla}\cdot\breve{\boldsymbol{v}} \neq 0. \tag{9.2.18}$$

The definition of the local growth rate, α^2 thus reads:

$$\begin{aligned}
\alpha^2 &\triangleq \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{\phi}^T\|^2 - 1, && (9.2.19)\\
&= \frac{1}{2}\|\mathbb{I}_d + t\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\|^2 - 1, && (9.2.20)\\
&= -t\left(t\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)\right) + \frac{t^2}{2}\|\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\|^2, && (9.2.21)\\
&= \frac{t^2}{2}\left((\partial_x\breve{u} - \partial_y\breve{v})^2 + (\partial_y\breve{u} + \partial_x\breve{v})^2\right). && (9.2.22)
\end{aligned}$$

It expresses the strain of the mesochronic velocity. Then, let us introduce the mesochronic vorticity $\breve{\omega} \triangleq \boldsymbol{\nabla}^\perp \cdot \breve{\boldsymbol{v}}$. Note that the mesochronic vorticity is not the vorticity time-averaged along a trajectory. With this notation, the incompressibility constraint (9.2.18) enables us to rewrite equation (9.2.22) as a function of the determinant $\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)$:

$$\begin{aligned}
\alpha^2 &= \frac{t^2}{2}\left((\boldsymbol{\nabla}\cdot\breve{\boldsymbol{v}})^2 - 4\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right) + \breve{\omega}^2\right), && (9.2.23)\\
&= \frac{t^2}{2}\left(t^2\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)\left(\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right) - \frac{4}{t^2}\right) + \breve{\omega}^2\right). && (9.2.24)
\end{aligned}$$

Mezić et al. (2010) then call mesoelliptic areas, areas over which the tracer gradients turn while keeping their norm unchanged, and mesohyperbolic areas, areas overwhich the gradients increase or decrease. According to (9.2.13), mesoelliptic regions are thus associated with zero growth rate α^2, and expression (9.2.24) leads to the equality

$$\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)\left(\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right) - \frac{4}{t^2}\right) = -\left(\frac{\breve{\omega}}{t}\right)^2 \leqslant 0. \tag{9.2.25}$$

We thus retrieve, following a different approach, the criterion introduced by Mezić et al. (2010). Indeed, the authors separate mesohyperbolic areas from mesoelliptic areas depending on the sign of the following criterion:

$$\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)\left(\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)-\frac{4}{t^2}\right). \tag{9.2.26}$$

Here, our developments provide further understandings. For the mesoellipticity case, equation (9.2.25) relates the criterion (9.2.26) to the mesochronic vorticity $\breve{\omega}$. Conversely, making use of (9.2.24) and (9.2.22), the separation between mesoelliptic and mesohyperbolic behaviors is given by the sign of:

$$\begin{aligned} t^2\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)\left(\det\left(\boldsymbol{\nabla}\breve{\boldsymbol{v}}^T\right)-\frac{4}{t^2}\right) &= 2\left(\frac{\alpha}{t}\right)^2-\breve{\omega}^2, & (9.2.27)\\ &= \left(\partial_x\breve{u}-\partial_y\breve{v}\right)^2+\left(\partial_y\breve{u}+\partial_x\breve{v}\right)^2-\breve{\omega}^2. & (9.2.28)\end{aligned}$$

This interpretation then becomes reminiscent to the Okubo-Weiss criterion (Okubo, 1970; Weiss, 1991; Shivamoggi and van Heijst, 2011). Indeed, this separation explicits the competition between the strain and the rotation of the mesochronic velocity, encoded by α and $\breve{\omega}$, respectively.

The mesohyperbolicity corresponds to a stretching. It occurs when two points become closer or diverge. This property is naturally encoded in:

$$\|\boldsymbol{\phi}^{-1}(\boldsymbol{x}+\boldsymbol{\delta x})-\boldsymbol{\phi}^{-1}(\boldsymbol{x})\|^2\approx\|[\boldsymbol{\nabla}(\boldsymbol{\phi}^{-1})^T(\boldsymbol{x})]\boldsymbol{\delta x}\|^2=\boldsymbol{\delta x}^T\left([\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{y})]\,[\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{y})]^T\right)^{-1}\boldsymbol{\delta x}, \tag{9.2.29}$$

where $\boldsymbol{y}=\boldsymbol{\phi}^{-1}(\boldsymbol{x})$. Mixing can occur when folding is associated with stretching. Folding is then obviously associated with a three-points kinematic property. First, the three points are separated by stretching, creating a filament. Then, the filaments folds bringing the two opposite points closer again. This folding can trap an area having a distinct tracer value creating strong tracer gradients. To identify mixing zones, Mezić et al. (2010) thus separate two types of mesohyperbolicity: the couples of points which have turned ($\boldsymbol{\delta x}^T(\boldsymbol{\phi}(\boldsymbol{x}+\boldsymbol{\delta x})-\boldsymbol{\phi}(\boldsymbol{x}))<0$) and the others ($\boldsymbol{\delta x}^T(\boldsymbol{\phi}(\boldsymbol{x}+\boldsymbol{\delta x})-\boldsymbol{\phi}(\boldsymbol{x}))>0$). An area where both types of mesohyperbolicity are present and adjacent must have been folded and hence corresponds to a mixing zone. To separate these two types of mesohyperbolicity, the authors study the eigenvalues of the evolution matrix $(\boldsymbol{\nabla}\boldsymbol{\phi}^T)^T$ instead of its singular values. Note however that folding could also be studied directly relying on the Cauchy-Green tensor, $[\boldsymbol{\nabla}\boldsymbol{\phi}^T]\,[\boldsymbol{\nabla}\boldsymbol{\phi}^T]^T$ Indeed,

$$\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x}+\boldsymbol{\delta x}_1)-\boldsymbol{\phi}^{-1}(\boldsymbol{x})\right)^T\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x}+\boldsymbol{\delta x}_2)-\boldsymbol{\phi}^{-1}(\boldsymbol{x})\right)\approx\boldsymbol{\delta x}_1^T\left([\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{y})]\,[\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{y})]^T\right)^{-1}\boldsymbol{\delta x}_2, \tag{9.2.30}$$

where $\boldsymbol{y}=\boldsymbol{\phi}^{-1}(\boldsymbol{x})$. Adequately normalized, this scalar product defines an angle between a pair of points. If the angle is smaller than the initial angle, defined by $\boldsymbol{\delta x}_1^T\boldsymbol{\delta x}_2$, folding occurs. Such criterion can be analytically expressed with eigenvectors and eigenvalues of the Cauchy-Green tensor.

To note, few other methods exist in the literature to diagnose folding and its relation to stretching. For instance, Budišić and Thiffeault (2015) study wrapping of advected points with braids. This mathematical tool provides simple visualizations of the number of winding of given sets of advected points relatively to other sets of points. As introduced, the Finite Time Braids Exponents (FTBE) then quantify the increasing rate of the number of windings. Following an other approach, Ma et al. (2016) directly measure folding of material lines through an alanlysis of their curvature variations. In the following, we relate folding, stretching and gradient of the curfacture of streamlines in the case of a slowly varying Eulerian velocity field.

9.2.4 Time dependance

In the work of Dencausse et al. (2014), the time resolution of Eulerian velocity is relatively low. This is a common feature of velocity estimates from space. Today, their typical temporal resolution is about 10 days. Moreover, being associated with geostrophic dynamics, the velocity correlation

time is generally about 1 month. Therefore, the velocity field does not vary much during an advection of one or two weeks. Hence, the velocity will be assumed stationary. This assumption determines a specific form for the flow. In particular, the flow is not chaotic (Thiffeault, 2004). The analyis will first separate between two typical cases: open straight streamlines and closed curved streamlines.

Locally uniform shear

Let us first focus on locally straight streamlines (*i.e.* streamlines with zero curvature). In such a case, the strengthening of tracer gradients results from a velocity shear, similarly to usual infinitesimal-time stretching. We denote by x the local axis of the straight streamline and by $v_x = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \cdot \boldsymbol{v} = \|\boldsymbol{v}\|$, the velocity component on this direction. The incompressibility imposes

$$\left(\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \cdot \boldsymbol{\nabla}\right) v_x = \partial_x v_x = \boldsymbol{\nabla} \cdot \boldsymbol{v} = 0. \tag{9.2.31}$$

Thus, since the Eulerian velocity is stationary, the Lagrangian velocity is stationary as well:

$$\frac{\mathrm{d}\boldsymbol{V}}{\mathrm{d}t}(\boldsymbol{x}_0, t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{v}\left(\boldsymbol{\phi}(\boldsymbol{x}_0, t)\right)\right) = \left(\left(\boldsymbol{v} \cdot \boldsymbol{\nabla}\right)\boldsymbol{v}\right)\left(\boldsymbol{\phi}(\boldsymbol{x}_0, t)\right) = 0, \tag{9.2.32}$$

and the flow simplifies to

$$\boldsymbol{\phi}(\boldsymbol{x}_0, t) = \boldsymbol{x}_0 + \int_0^t \mathrm{d}t' V(\boldsymbol{x}_0, t') = \boldsymbol{x}_0 + v(\boldsymbol{x}_0)t. \tag{9.2.33}$$

This so-called ballistic regime is superdiffusive (Vallis, 2006; Falkovich et al., 2001). Taking the gradient of the above expression and, using the incompressibility again, the stretching rate reads

$$\alpha^2 \triangleq \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{\phi}^T\|^2 - 1 = \frac{1}{2}\left\|\mathbb{I}_d + \boldsymbol{\nabla}\boldsymbol{v}^T t\right\|^2 - 1 = \frac{t^2}{2}\left\|\boldsymbol{\nabla}\boldsymbol{v}^T\right\|^2 = t^2 \underbrace{\frac{1}{2}\left(\frac{\boldsymbol{v}^\perp}{\|\boldsymbol{v}^\perp\|} \cdot \boldsymbol{\nabla}\|\boldsymbol{v}\|\right)^2}_{=1/\tau_s^2}, \tag{9.2.34}$$

where the last equality comes from the orientation of velocity gradients imposed by (9.2.31). The time τ_s will be referred to as the shearing time.

Stationary convective cells

Close to vortices, streamlines are often closed or at least curved, and the previous development cannot be applied. For this purpose, we now focus on closed streamlines. Since the flow is incompressible, fluid parcels cannot accumulate. Therefore, those streamlines define loops, called stationary convective cells (Falkovich et al., 2001), where fluid parcels rotate periodically. Accordingly, the flow and thus the Lagrangian velocity are periodic and the flow is called subdiffusive (Vallis, 2006; Falkovich et al., 2001). This geometry can nevertheless create strong stretching in finite-time. Indeed, two concentric closed streamlines can define Lagrangian loops associated with different rotation periods. Rotations after rotations, a fluid parcel on the fastest loop will deviate from its initial neighboring parcel on the slowest loop. So, this differential rotation creates stretching. Moreover, it also induces folding. A filament distributed perpendicular to streamlines will be deformed by the continuous differential rotation. After a finite time, the filament will wrap around the convective cell creating spirals. Lehahn et al. (2007) illustrate a similar process with the action of stable and unstable manifolds on phytoplankton patches. To mathematically express the stretching induced by those convective cells, we propose to write the flow as follows:

$$\boldsymbol{\phi}(\boldsymbol{x}_0) = \boldsymbol{\phi}(\boldsymbol{x}_0, t) \approx \boldsymbol{x}_0 + \boldsymbol{g}(\boldsymbol{x}_0, f(\boldsymbol{x}_0)t), \tag{9.2.35}$$

where $\boldsymbol{g}$ is $1-$periodic with respect to its second variable and $f(\boldsymbol{x}_0)$ is the local temporal frequency. Accordingly, for a point initially on $\boldsymbol{x}_0$ in a closed streamline $\mathcal{C}$, the trajectory $t \mapsto \boldsymbol{\phi}(\boldsymbol{x}_0, t)$ runs from $\boldsymbol{x}_0$ to $\boldsymbol{x}_0$ through a path $\mathcal{P}$ embedded in $\mathcal{C}$ with a temporal period $1/f(\boldsymbol{x}_0)$ defined by:

$$\frac{1}{f(\boldsymbol{x}_0)} = \int_0^{1/f(\boldsymbol{x}_0)} \mathrm{d}t = \int_{\mathcal{P}} \frac{\mathrm{d}l}{\|\boldsymbol{v}\|} = \oint_{\mathcal{C}} \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|^2} \cdot \mathrm{d}\boldsymbol{l} = \iint_{\mathcal{A}} \boldsymbol{\nabla}^\perp \cdot \left(\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|^2}\right) \mathrm{d}\boldsymbol{A}, \tag{9.2.36}$$

where $\boldsymbol{\nabla}^{\perp}\cdot$ denotes the two dimensional-curl and $\mathcal{A}$ the surface delimited by $\mathcal{C}$. As the two last integrals only depend on the streamline $\mathcal{C}$ and not on the precise initial condition $\boldsymbol{x}_0$, the local frequency inherits from the same invariance. Since the Eulerian velocity is stationary, the points $\boldsymbol{x}_0$ and $\boldsymbol{\phi}(\boldsymbol{x}_0, t)$ are on the same streamline, and thus:

$$f(\boldsymbol{\phi}(\boldsymbol{x}_0, t)) = f(\boldsymbol{x}_0). \tag{9.2.37}$$

This frequency can be approximated by a local angular velocity $\dot{\theta}$, estimated using the streamline curvature, denoted $1/R$, as:

$$f \approx \frac{\dot{\theta}}{2\pi} \approx \frac{\|\boldsymbol{v}\|}{2\pi R} = \frac{1}{2\pi}\left[(\boldsymbol{v}\cdot\boldsymbol{\nabla})\,\frac{\boldsymbol{v}}{\|\boldsymbol{v}\|}\right]\cdot\frac{\boldsymbol{v}^{\perp}}{\|\boldsymbol{v}^{\perp}\|}. \tag{9.2.38}$$

In practice, since the exact formula (9.2.36) can be difficult to evaluate numerically, we will instead use the above approximation. In the following derivation, we however keep the exact definition (9.2.36). In particular, we still assume the frequency invariance along the streamline (9.2.37). The first coordinate of $\boldsymbol{g}$ encodes the spatial dependency of the loop (vectorial) amplitudes. Note that the model (9.2.35) is very general as it only assumes periodicity of Lagrangian trajectories. It enables us to partially decouple flow variations associated with different streamlines (*i.e.* different local frequencies f) and flow variations associated with different temporal phase shift along the streamline (*i.e.* different times t). To some extent, this second type of variation can be understood as different initial conditions in the same streamline, due to the periodicity assumption. Similar decomposition ideas were proposed by Thiffeault (2004) for chaotic (non-periodic) flows. Denotes:

$$(\partial_1\boldsymbol{g}^T)(\boldsymbol{z}_1, z_2) = \boldsymbol{\nabla}_{\boldsymbol{z}_1}\left(\boldsymbol{g}^T(\boldsymbol{z}_1, z_2)\right) \text{ and } (\partial_2\boldsymbol{g})(\boldsymbol{z}_1, z_2) = \partial_{z_2}\left(\boldsymbol{g}(\boldsymbol{z}_1, z_2)\right). \tag{9.2.39}$$

Note that both terms are $1-$periodic with respect to its second variable. By the frequency invariance (9.2.37), we can replace $f(\boldsymbol{x}_0)$ by $f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))$ in the model (9.2.35):

$$\boldsymbol{\phi}(\boldsymbol{x}_0, t) = \boldsymbol{x}_0 + \boldsymbol{g}(\boldsymbol{x}_0, f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))t). \tag{9.2.40}$$

Then, replacing back $f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))$ by $f(\boldsymbol{x}_0)$ after evaluating the gradient, the stretching of the flow reads:

$$\begin{aligned}
\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t) &= \mathbb{I}_d + (\partial_1\boldsymbol{g}^T)(\boldsymbol{x}_0, f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))t) \\
&\quad + t\ \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t)\ \boldsymbol{\nabla} f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))\ (\partial_2\boldsymbol{g}^T)(\boldsymbol{x}_0, f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))t), && (9.2.41)\\
&= \mathbb{I}_d + (\partial_1\boldsymbol{g}^T)(\boldsymbol{x}_0, f(\boldsymbol{x}_0)t) \\
&\quad + t\ \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t)\ \boldsymbol{\nabla} f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))\ (\partial_2\boldsymbol{g}^T)(\boldsymbol{x}_0, f(\boldsymbol{x}_0)t). && (9.2.42)
\end{aligned}$$

In the last equality, the second right-hand term is time-periodic and thus bounded. If we neglect its time variation, it writes

$$(\partial_1\boldsymbol{g}^T)(\boldsymbol{x}_0, f(\boldsymbol{x}_0)t) \approx (\partial_1\boldsymbol{g}^T)(\boldsymbol{x}_0, 0) = \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, 0) - \mathbb{I}_d = 0. \tag{9.2.43}$$

Introducing the original periodic model (9.2.35) into its definition, the Lagrangian velocity $\boldsymbol{V}$ reads:

$$\boldsymbol{V}(\boldsymbol{x}_0, t) = \frac{\mathrm{d}\boldsymbol{\phi}(\boldsymbol{x}_0, t)}{\mathrm{d}t} = f(\boldsymbol{x}_0)(\partial_2\boldsymbol{g})(\boldsymbol{x}_0, f(\boldsymbol{x}_0)t). \tag{9.2.44}$$

Finally, the flow gradient expression (9.2.42) can be rewritten using equations (9.2.43) and (9.2.44):

$$\begin{aligned}
\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t) &= \mathbb{I}_d + t\ \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t)\frac{\boldsymbol{\nabla} f(\boldsymbol{\phi}(\boldsymbol{x}_0, t))}{f(\boldsymbol{x}_0)}\boldsymbol{V}^T(\boldsymbol{x}_0, t), && (9.2.45)\\
&= \mathbb{I}_d + t\ \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t)\left(\frac{1}{f}\boldsymbol{\nabla} f\ \boldsymbol{v}^T\right)(\boldsymbol{\phi}(\boldsymbol{x}_0, t)), && (9.2.46)
\end{aligned}$$

where the frequency invariance (9.2.37) was used in the last equality. After this, we factorize terms in $\boldsymbol{\nabla}\boldsymbol{\phi}^T$,

$$\mathbb{I}_d = \boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0, t)\left(\mathbb{I}_d - t\ \left(\frac{1}{f}\boldsymbol{\nabla} f\ \boldsymbol{v}^T\right)(\boldsymbol{\phi}(\boldsymbol{x}_0, t))\right). \tag{9.2.47}$$

Then, we inverse the matrix equation and remap with the inverse flow, $\boldsymbol{\phi}^{-1}$:

$$\left(\boldsymbol{\nabla}\boldsymbol{\phi}^T\right)^{-1}\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x},t),t\right) \quad = \quad \mathbb{I}_d - t\,\left(\frac{1}{f}\boldsymbol{\nabla} f\;\boldsymbol{v}^T\right)(\boldsymbol{x}). \tag{9.2.48}$$

Note that since the frequency is a function of the streamline (9.2.37),

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(f(\boldsymbol{x}_0)\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(f\left(\boldsymbol{\phi}(\boldsymbol{x}_0,t)\right)\right) = \left(\boldsymbol{v}\cdot\boldsymbol{\nabla} f\right)\left(\boldsymbol{\phi}(\boldsymbol{x}_0,t),t\right). \tag{9.2.49}$$

Therefore, the frequency gradient is orthogonal to the velocity, and

$$\|\boldsymbol{\nabla} f\| \approx \left|\frac{\boldsymbol{v}^\perp}{\|\boldsymbol{v}^\perp\|}\cdot\boldsymbol{\nabla} f\right|. \tag{9.2.50}$$

As the eigenvalues of the flow gradients $\boldsymbol{\nabla}\boldsymbol{\phi}^T$ are the inverse of one another, this matrix and its inverse have the same Frobenius norm. So, the time dependance of the growth rate in the final grid (points $\boldsymbol{x}$) follows from its definition and from (9.2.48):

$$\begin{aligned}
\alpha^2\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x},t),t\right) &\triangleq \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{\phi}^T\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x},t),t\right)\|^2 - 1, && (9.2.51)\\
&= \frac{1}{2}\|\left(\boldsymbol{\nabla}\boldsymbol{\phi}^T\right)^{-1}\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x},t),t\right)\|^2 - 1, && (9.2.52)\\
&= -\underbrace{\left(\frac{t}{f}\boldsymbol{\nabla} f\cdot\boldsymbol{v}\right)}_{=0 \text{ by } (9.2.49)}(\boldsymbol{x}) + \underbrace{\left(\frac{t^2}{2f^2}\|\boldsymbol{\nabla} f\|^2\|\boldsymbol{v}\|^2\right)}_{=1/\tau_f^2}(\boldsymbol{x}), && (9.2.53)
\end{aligned}$$

where τ_f will be referred to as the folding time. As (9.2.15) only involves the spatial average of α^2, we can further simplify the model by spatial integration. By integrating equation (9.2.53) over a specific domain Ω_f, we can conclude with the variable change defined by the incompressible flow:

$$\int_{\Omega_f}\mathrm{d}\boldsymbol{x}_0\;\alpha^2(\boldsymbol{x}_0,t) = \int_{\boldsymbol{\phi}(\Omega_f)}\mathrm{d}\boldsymbol{x}_0\;\alpha^2\left(\boldsymbol{\phi}^{-1}(\boldsymbol{x},t),t\right) = t^2\int_{\boldsymbol{\phi}(\Omega_f)}\frac{\mathrm{d}\boldsymbol{x}}{\tau_f^2(\boldsymbol{x})}. \tag{9.2.54}$$

The subspace Ω_f is a subset of Ω where the concept of wrapping convective cells is relevant. Since we considered closed streamlines, we assume that $\boldsymbol{\phi}(\Omega_f) = \Omega_f$. This subspace will be properly defined in the following.

Global time dependence

In order to combine the folding time, τ_f, and the shearing time, τ_s, we locally define a local stretching time τ:

$$\tau(\boldsymbol{x}_0) \triangleq \begin{cases} \tau_f(\boldsymbol{x}_0) \text{ if } R(\boldsymbol{x}_0) \leqslant \frac{L}{2} \\ \tau_s(\boldsymbol{x}_0) \text{ if } R(\boldsymbol{x}_0) > \frac{L}{2} \end{cases}, \tag{9.2.55}$$

where $1/R(\boldsymbol{x}_0)$ is the streamline curvature on $\boldsymbol{x}_0$, L the average diameter of a vortex. Following the previous models of shearing and folding, we model the stretching rate by:

$$\alpha = \frac{t}{\tau}. \tag{9.2.56}$$

Where gradients are created by a uniform shear, streamlines are straight, the curvature $1/R$ is weak and $\tau = \tau_s$, whereas, where the gradients strengthen by wrapping, the curvature is larger and $\tau = \tau_f$. To estimate the average vortices diameter L, we make uses of a toy model:

$$\boldsymbol{v} \quad = \quad U\begin{pmatrix}\cos\left(\frac{2\pi}{\lambda}x\right)\sin\left(\frac{2\pi}{\lambda}y\right)\\ \sin\left(\frac{2\pi}{\lambda}x\right)\cos\left(\frac{2\pi}{\lambda}y\right)\end{pmatrix}. \tag{9.2.57}$$

The vortices diameter is identified to the size of the convective cell:

$$L = \frac{\lambda}{2} = \left(\frac{6\pi^2 \overline{\|\boldsymbol{v}\|^2}}{\overline{\|\boldsymbol{\nabla} \boldsymbol{v}^T\|^2}} \right)^{1/2}. \tag{9.2.58}$$

We use this diameter estimator in the general case.

The condition (9.2.55) well defines a space partition $\Omega = \Omega_f \cup \Omega_s$ to integrate the growth rate α^2:

$$\overline{\alpha^2} = \left(\frac{t}{\tau_G} \right)^2 \text{ with } \frac{1}{\tau_G^2} \triangleq \overline{\left(\frac{1}{\tau^2} \right)} = \frac{1}{S} \, \mathbb{E} \left(\int_{\Omega_f} \frac{\mathrm{d}\boldsymbol{x}}{\tau_f^2(\boldsymbol{x})} + \int_{\Omega_s} \frac{\mathrm{d}\boldsymbol{x}_0}{\tau_f^2(\boldsymbol{x}_0)} \right), \tag{9.2.59}$$

where $\Omega_f \triangleq \{\boldsymbol{x} \in \Omega | R(\boldsymbol{x}) \leqslant \frac{L}{2}\}$ and $\Omega_s \triangleq \{\boldsymbol{x}_0 \in \Omega | R(\boldsymbol{x}_0) > \frac{L}{2}\}$. Again, we assumed that $\boldsymbol{\phi}(\Omega_f) = \Omega_f$ because the flow maps closed streamlines onto themselves. In the following, we will refer to τ_G as the global stretching time. The model (9.2.59) together with the folding and shearing time definitions (9.2.53)-(9.2.34) specify a global Eulerian estimate of finite-time stretching. Unlike usual diagnosis, such as FTLE and Finite Size Lyapunov Exponents (FSLE) (d'Ovidio et al., 2009), the proposed global model does not require any integration of the flow.

According to (9.2.15), the evolution law (9.2.59) determines the tracer gradient norm:

$$\frac{\overline{\|\boldsymbol{\nabla} T\|^2}}{\overline{\|\boldsymbol{\nabla} T_0\|^2}} \approx 1 + \left(\frac{t}{\tau_G} \right)^2. \tag{9.2.60}$$

9.2.5 Numerical illustrations

Let us first exemplify these analytical developments with an instructive toy example, before dealing with satellite observations. We advected a large tracer filament by a stationary velocity field using a backward Lagrangian advection (Figure 9.2). For technical details on this method, we referred to Berti and Lapeyre (2014) and Dencausse et al. (2014). The tracer progressively wraps, eventually creating infinitely long filaments. With the formula (9.2.59), we estimate a global stretching time of 13.36 days. This roughly corresponds to half a rotation.

Figure 9.3 represents the spatial distribution of the stretching rate α^2, the factor α/β and the mesochronic vorticity $\tilde{\omega}$ at several times for our toy model. The spatial distribution of α stabilizes after one week only. As demonstrated, this non-dimentionalized number is significant on the folding area (the border of the vortex). The ratio $\alpha/\beta = \sqrt{1 + \frac{2}{\alpha^2}}$ – which quantifies the significance of the orientation of tracer gradient – decreases along time in the mixing area. It stabilizes to its minimum value, say 1, at $t \approx \tau_G$. The mesochronic vorticity is first concentrated in the center of the cylinder. Then, after each global stretching time, a new ring of mesochronic vorticity is added to the mixing area.

Figure (9.4) displays the spatial distribution of the folding time, τ_f, the shearing time, τ_s, and the stretching time, τ, for this toy model. Folding and stretching time are represented both in the initial grid ($\boldsymbol{x}_0$) and in the advected grid ($\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{x}_0, t)$). This remapping on the initial grid is needed as the folding time is locally defined in the advected grid (see (9.2.53) and (9.2.54)). For this remapping, we integrated the forward flow $\boldsymbol{x}_0 \mapsto \boldsymbol{\phi}(\boldsymbol{x}_0, t)$. Note that the remapping enables a better visualization of the stretching spatial distribution, but this mapping is not necessary for the global stretching time computation (9.2.59). The folding effects are dominant in this toy model and the folding time well captures the spatial structure of α^2. Yet, the inverse folding time diverges outside of the vortex as the streamline curvatures tend to zero. Indeed, according to (9.2.38) zero curvature implies zero local frequency f and thus infinite folding time (see (9.2.53)). For such a weak curvature, the relevant model is the uniform shear. Following the space partition (9.2.55), the stretching time is chosen as a shearing time in these areas. The global models of time evolution of the averaged stretching rate (9.2.59) and of the tracer gradients (9.2.60) are also successfully tested in Figure 9.5.

We now perform a similar analysis using satellite data. Velocities are estimated from altimeter-derived SSH fields. The velocity field, on which the Eulerian estimates will be applied, corresponds

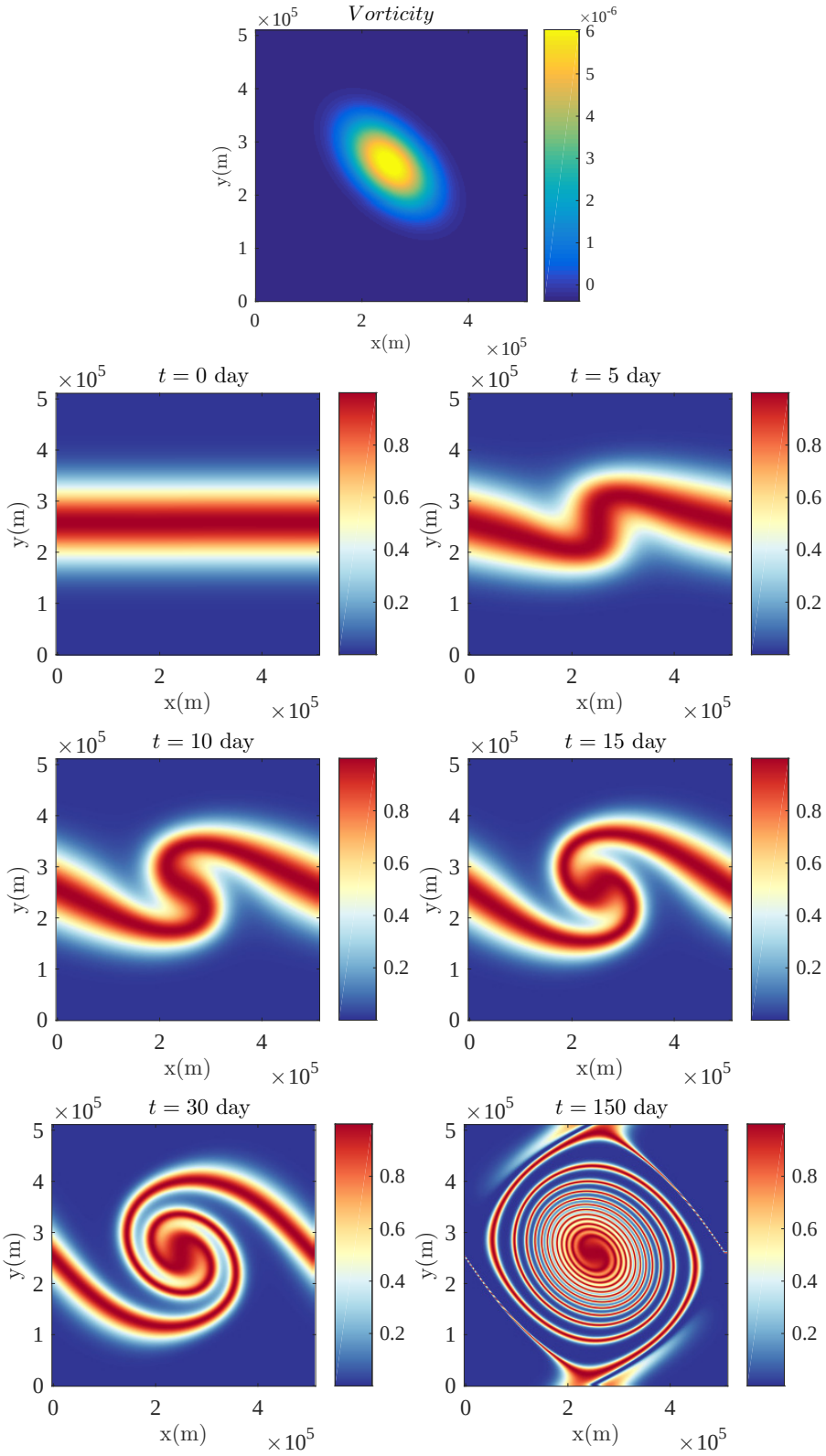


Figure 9.2: Advecting vorticity of the toy model (top in s^{-1}) and tracer (dimensionless) advected using a backward Lagrangian method at time $t = 0, 5, 10, 15, 30$ and 150 days.

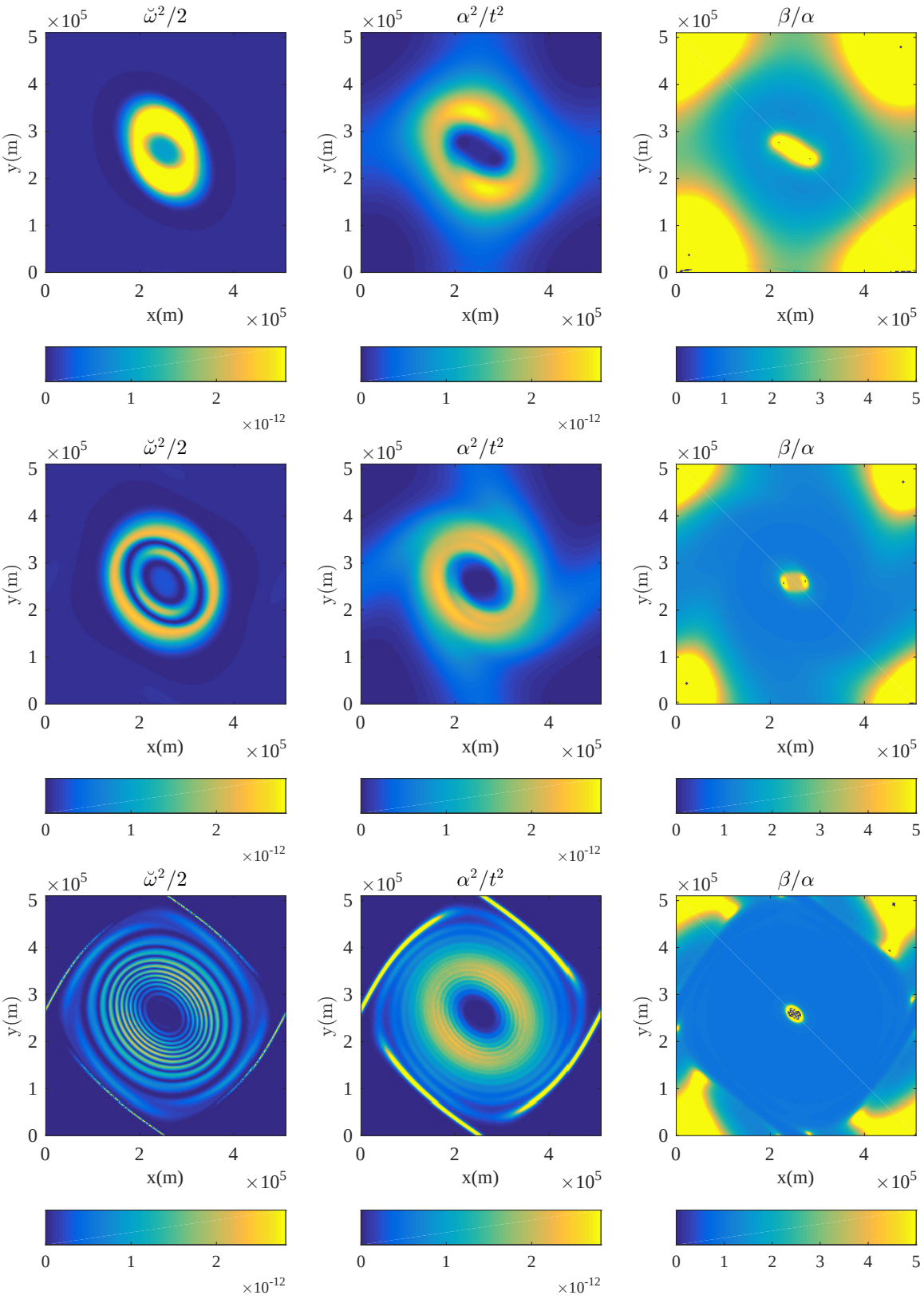


Figure 9.3: Values of $\tilde{\omega}^2/2$ (s^{-2}) (left), $(\alpha/t)^2$ (s^{-2}) (middle) and the ratio of α/β (dimensionless) (right) in the initial grid (points $\boldsymbol{x}_0$) at time (from top to bottom) $t = 15, 30$ and 150 days for the toy model.

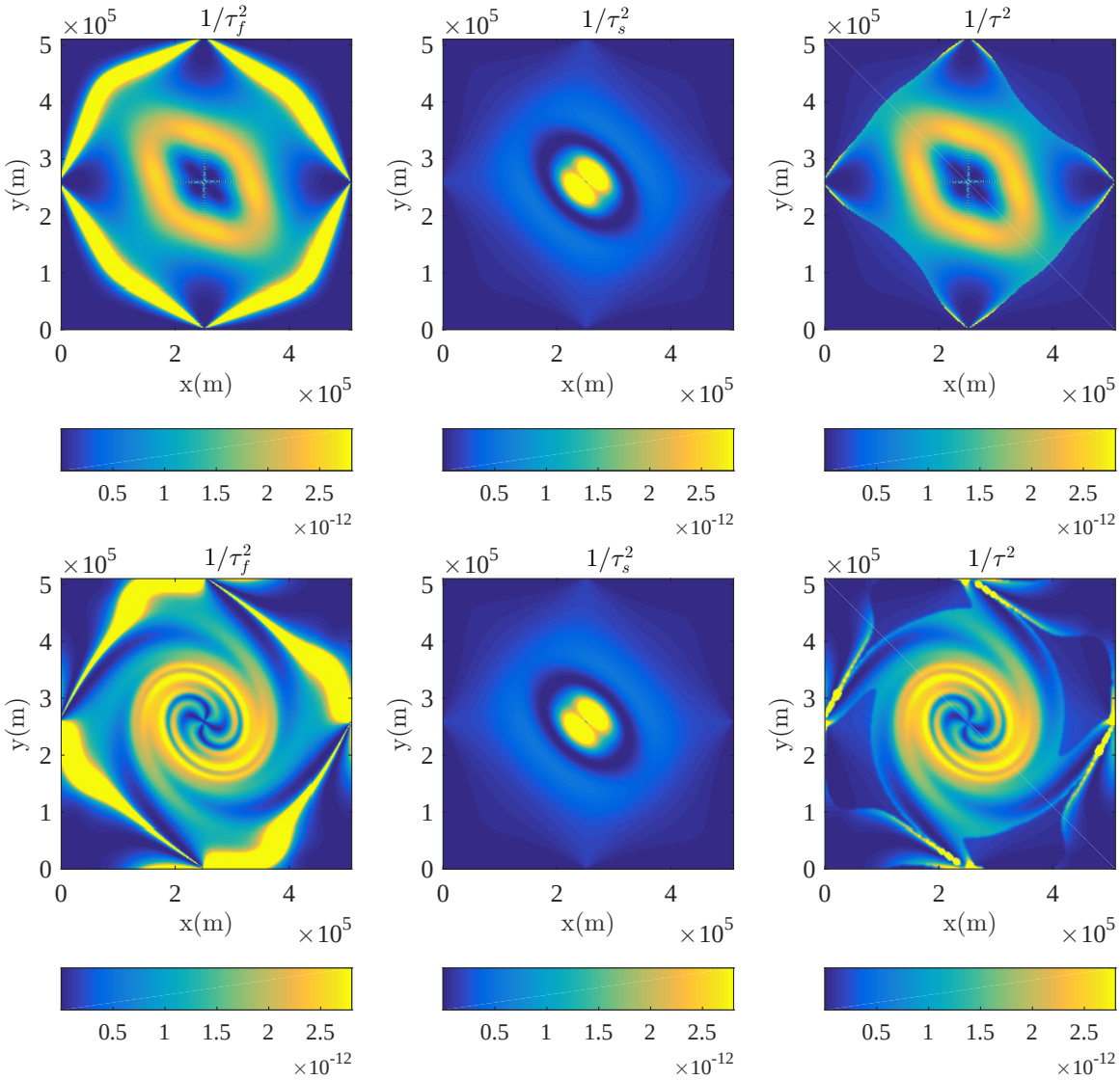


Figure 9.4: Squared inverse of the folding time in the final grid (points $\boldsymbol{x}$) (top left) and initial grid (points $\boldsymbol{x}_0$) (bottom left), the shearing in the initial grid (points $\boldsymbol{x}_0$) (top and bottom middle) and the stretching time in the final grid (points $\boldsymbol{x}$) (top right) and initial grid (points $\boldsymbol{x}_0$) (bottom right) for the toy model. All plots are in s^{-2}. In order to represent folding and stretching time in the initial grid, these fields were advected during 30 days.

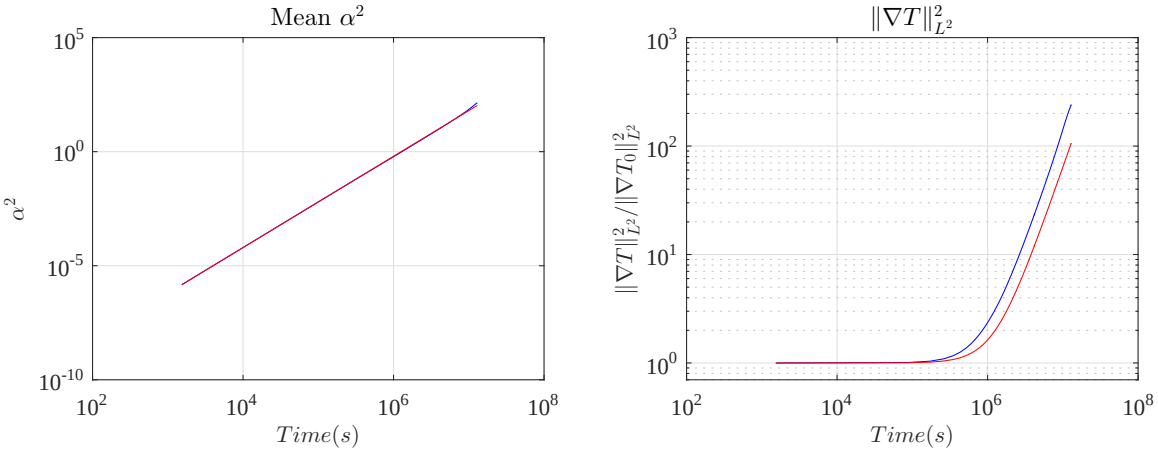


Figure 9.5: The averaged growth rate, $\overline{\alpha^2}$, (left) and the averaged squared norm of tracer gradients for the toy model, both in log-log plot along time. The blue line is the real value and the red line our model.

to January 1st, 2011, in the Antarctic Circumpolar Current (ACC) region, south of Australia. We begin by a small spatial window of $10° \times 10°$. A larger window will be considered subsequently. Figure 9.6 delineates the Kinetic energy (KE) and the vorticity fields. The KE shows the ACC eastward jet between latitudes $-50°$ and $-48°$. Two (warm) anticyclones and a (cold) depression are also visible both in the vorticity and in initial SST fields at $(129°, -51°)$, $(131°, -54°)$ and $(130.5°, -49.5°)$, respectively. We then advect the January 1st, 2011, SST field. Similarly to Dencausse et al. (2014), we linearly interpolate in time the daily velocity data to perform the advection operation. The vortice dipole closed to the jet creates a mushroom-like structure in the advected tracer. Each vortice wraps the tracer, creating spirals. The small southern anticyclone $(131°, -54°)$ seems weaker than the other anticyclone $(129°, -51°)$. Yet, it faster wraps the tracer, as velocities are certainly larger than over the dipole area.

Figure 9.7 shows the stretching rate, α, the mesochronic vorticity, $\tilde{\omega}$, and the estimate (weighting) of the tracer/flow correlation, β/α. The folding, shearing and stretching time are presented in the same Figure. A slight low-pass spatial filtering (2-km filter width) is applied to the stretching time to help distinguish the filamentous structures. The spatial distributions of stretching rate and stretching time are very similar. The amplitude of the stretching time is slightly underestimated (ratio of about 2). The sides of the aforementioned vortices exhibit intense mixing, whereas the inverse shearing time is weak. As for the toy model, folding effects due to differential rotations near the boundaries of vortices are the leading mixing processes.

Now, we consider a larger space window to visualize a broader variety of structures and dynamical processes. The spatial location and the date remain the same. Figure 9.8 displays the KE and the vorticity. The jet and many vortices are visible. In the same manner, the SST is advected (Figure 9.9). The advection creates small-scale structures which turn to unphysical spirals when the advection time is too long. At 48 days of advection the advected domain is strongly deformed, especially by the eastward jet. Figure 9.10 compares the stretching ratio, α^2, and the estimated stretching time in this larger spatial window. As found, most stretching structures are well predicted by the proposed model.

Finally, Figure 9.11 presents the time evolution of the averaged stretching rate (9.2.59) and of the averaged tracer gradients norm (9.2.60). The reference plots clearly exhibit the structures prescribed by models (9.11) and (9.2.59):

$$\overline{\alpha^2} = \left(\frac{t}{\tau_G}\right)^2 \quad \text{and} \quad \frac{\overline{\|\boldsymbol{\nabla} T\|^2}}{\|\boldsymbol{\nabla} T_0\|^2} = \begin{cases} 1 & \text{if } t \ll \tau_G \\ \left(\frac{t}{\tau_G}\right)^2 & \text{if } t \gg \tau_G \end{cases} . \tag{9.2.61}$$

We estimate a global stretching time of 1.67 days and the plots reveal a good match even though the stretching time seems slightly underestimated (by a factor of about ~ 1.7). The small shift

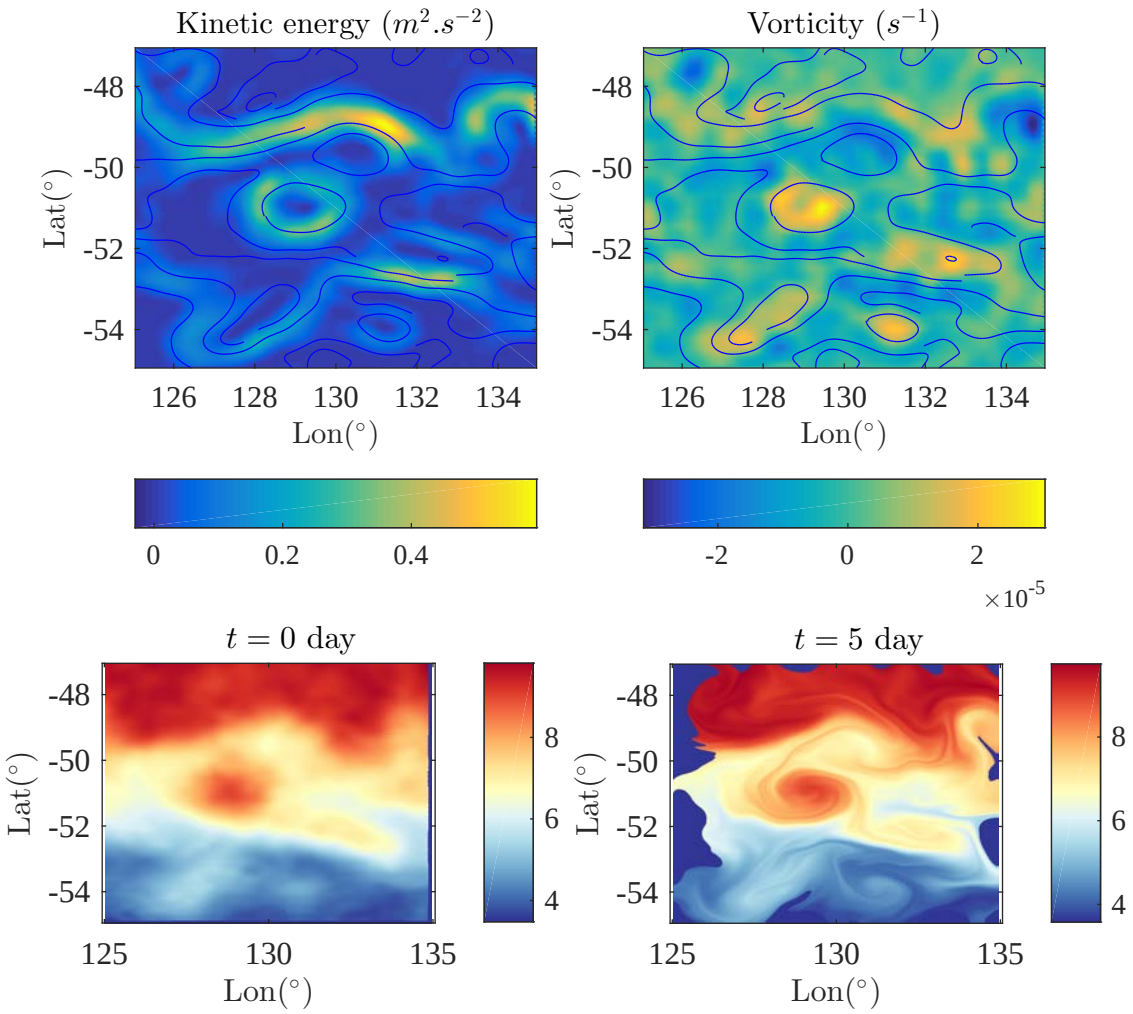


Figure 9.6: Kinetic energy (KE) (top left in $m^2.s^{-2}$), vorticity (top right in s^{-1}), SST (bottom left in °C), all measured by satellite the 1st of January 2011, and SST (in °C) after a 5-day advection (bottom right). On the top images, streamlines are superimposed. The streamlines, the KE, the vorticity and the advection are defined by SSH-derived velocity fields.

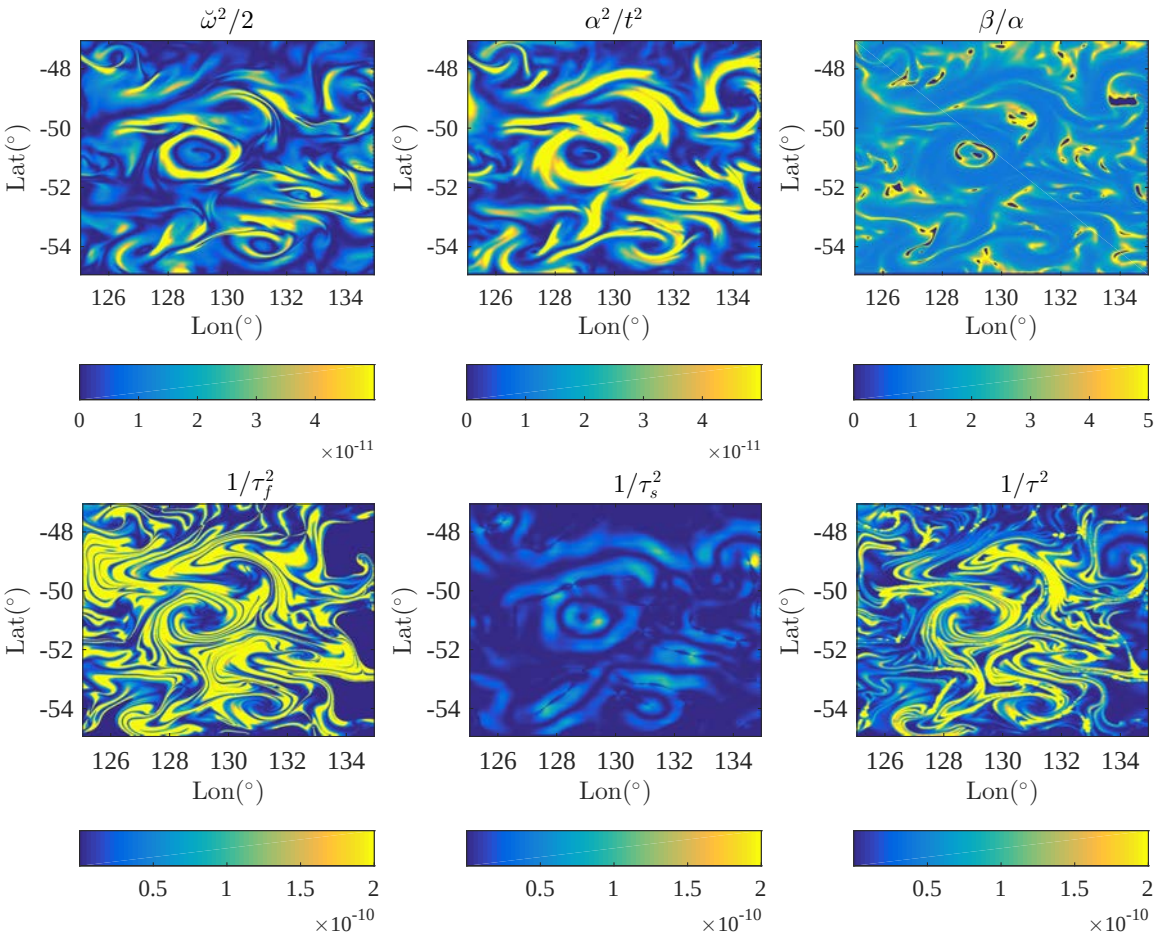


Figure 9.7: Values of the mesochronic vorticity, $\tilde{\omega}^2/2$, (s^{-2}) (top left), the stretching growth rate, $(\alpha/t)^2$, (s^{-2}) (top middle), the ratio α/β (dimensionless) (top right), the squared inverse of the folding time (s^{-2}) (bottom left), the shearing time (s^{-2}) (bottom middle) and the stretching time (s^{-2}) (bottom right), in the initial grid (points $\boldsymbol{x}_0$) at time $t = 5$ days for the SSH-derived velocity fields. We can observe the good match between the stretching rate and our Eulerian estimation of the stretching time.

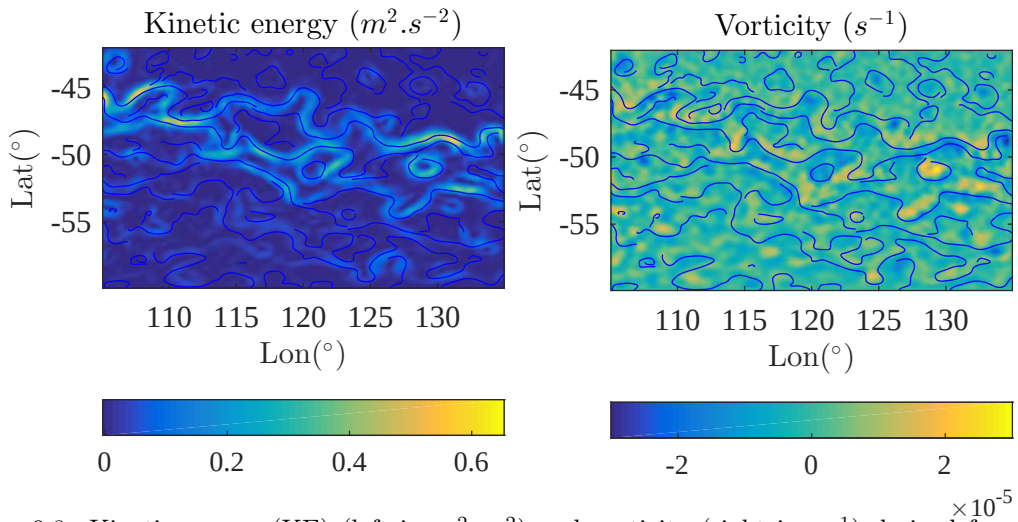


Figure 9.8: Kinetic energy (KE) (left in $m^2.s^{-2}$) and vorticity (right in s^{-1}) derived from SSH measured the 1st of January 2011.

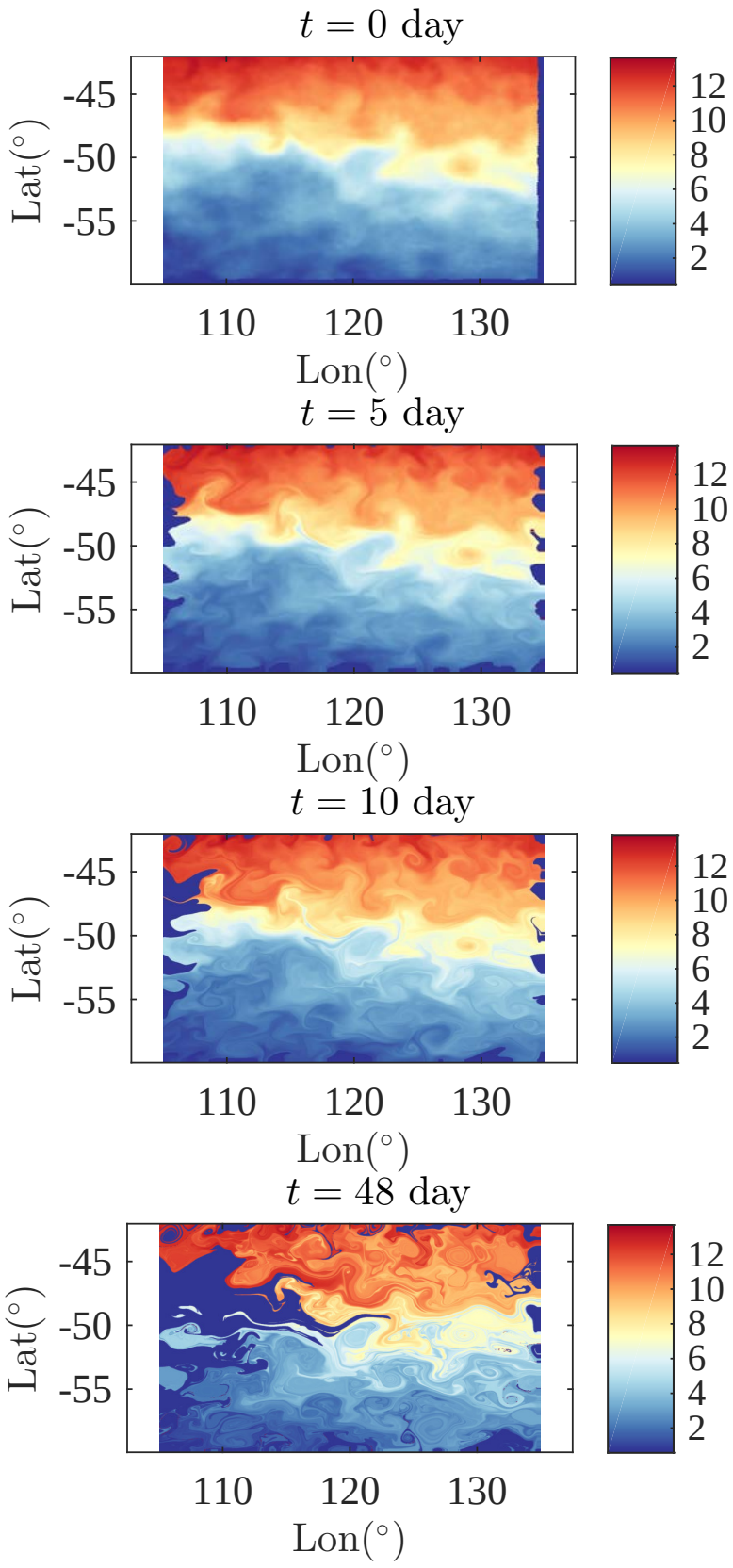


Figure 9.9: SST (in °C) (from top to bottom) measured by satellite the 1st of January 2011, after 5-, 10-, and 48-day advection.

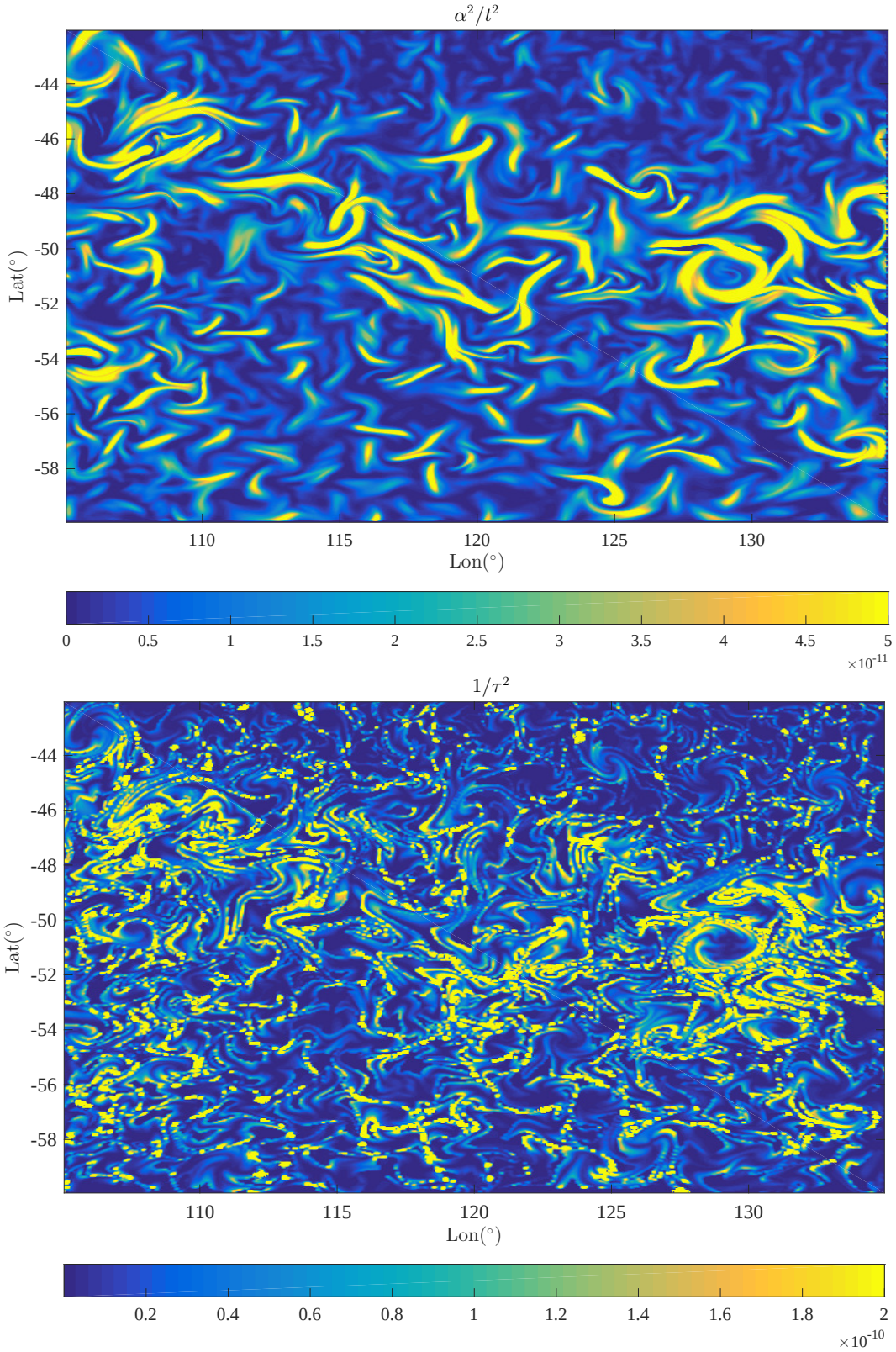


Figure 9.10: The stretching growth rate $(\alpha/t)^2$ (s^{-2}) (top) and the squared inverse of the stretching time, $1/\tau^2$, (s^{-2}) (bottom), in the initial grid (points $\boldsymbol{x}_0$) at time $t = 5$ days.

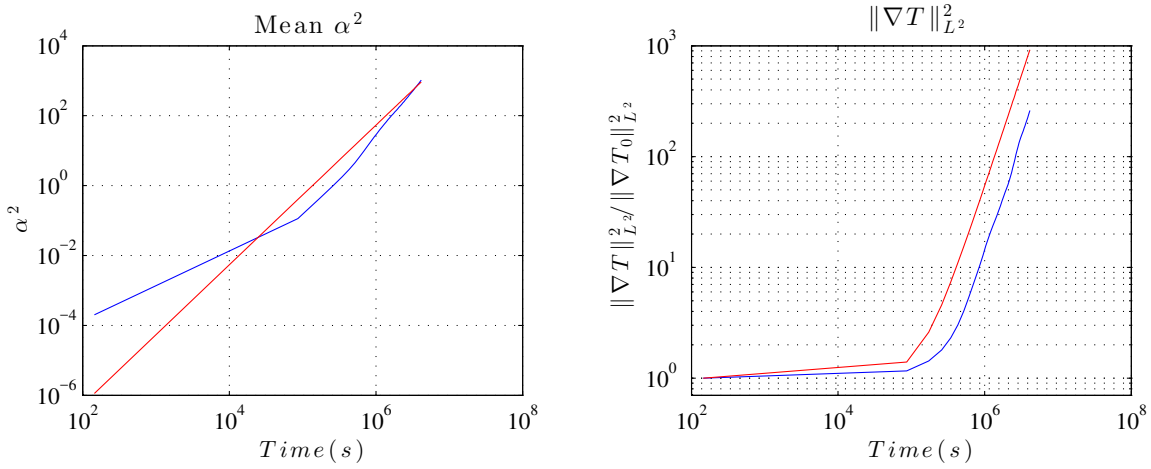


Figure 9.11: The averaged stretching rate, $\overline{\alpha^2}$, (left) and the ratio of squared-norm mean of tracer gradients, $\frac{\overline{\|\boldsymbol{\nabla} T\|^2}}{\|\boldsymbol{\nabla} T_0\|^2}$ for the satellite data, both in log-log plot along time. The blue line is the real value and the red line our model.

between predicted and reference averaged tracer gradients norm may also be explained by a residual correlation between the tracer and the underlying flow as explained previously with equation (9.2.13).

9.3 Tracer spectral tail

In the following, we study the distortion of scales of the tracer T during the advection, *i.e.* the evolution of its spectral tail. The main results of this section does not necessitate to consider the decorrelation assumption of 9.2.2, nor the stationary assumption used in 9.2.4. Developments are thus more general. After preliminary results related to tracer moments, we first derive a Gaussian approximation for the evolution of the spectrum tail, assuming spatial smoothness. This approximation is then applied to initial and advected tracers. Finally, we generalize the development and discussion to self-similar spectra.

9.3.1 Moment conservation

In the following, $\hat{T}_0$, $\hat{T}$, γ_{T_0}, γ_T and Γ_{T_0}, Γ_T denote the Fourier transforms, the covariances (associated with the averaging $\overline{\bullet}$) and the spectra of initial and final tracer fields, respectively. Furthermore, the Hessian of a function f will be denoted by $\boldsymbol{H}_f$. It is interesting to remark that the tracer mean, $\overline{T}$, does not change during the advection. Indeed, for the advected tracer $T(\boldsymbol{x}) = T_0(\boldsymbol{\phi}^{-1}(\boldsymbol{x}))$, the variable change defined by the inverse flow $\boldsymbol{x} \to \boldsymbol{\phi}^{-1}(\boldsymbol{x}) = \boldsymbol{x}_0$ and the incompressiblity constraint ($\det(\boldsymbol{\nabla}\boldsymbol{\phi}^T) = 1$) yields:

$$\overline{T} = \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; T(\boldsymbol{x}) = \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}_0\; \underbrace{\frac{1}{\det(\boldsymbol{\nabla}\boldsymbol{\phi}^T(\boldsymbol{x}_0))}}_{=1} T_0(\boldsymbol{x}_0) = \overline{T_0}. \tag{9.3.1}$$

As a consequence, we will assume without loss of generality that the tracer is centered. Similarly, the (spatial) variance conservation can be deduced:

$$\gamma_T(0) = \overline{T^2} = \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; T^2(\boldsymbol{x}) = \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}_0\; T_0^2(\boldsymbol{x}_0) = \overline{T_0^2}. \tag{9.3.2}$$

Since all scales are assumed to be resolved and the molecular diffusion is ineffective on the length and time scales of interest, there are no overlays of fluid parcels and no dilution of their properties.

Each fluid parcel conserves its tracer value during its advection. More generally, for all functions f, $\overline{f(T)}$ is conserved.

9.3.2 Smooth scalar approximation

First, let us consider the covariance of a smooth scalar f, for small spatial distance $\|\boldsymbol{\delta x}\|$. More specifically, we assume this covariance to be three times differentiable, which is a strong assumption on the scalar initial regularity. Yet, for tracers measured at mesoscales and re-interpolated on a sub-mesoscale spatial grid, this assumption is still relevant. The covariance can then be approximated near 0 by its Taylor expansion:

$$\gamma_f(\boldsymbol{\delta x}) \triangleq \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; f(\boldsymbol{x})f(\boldsymbol{x}+\boldsymbol{\delta x}), \tag{9.3.3}$$

$$= \overline{\|f\|^2} + \frac{1}{2}\boldsymbol{\delta x}^T\boldsymbol{H}_{\gamma_f}(0)\boldsymbol{\delta x} + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}(\|\boldsymbol{\delta x}\|^3), \tag{9.3.4}$$

$$= \overline{\|f\|^2}\exp\left(-\frac{1}{2}\boldsymbol{\delta x}^T\left(\frac{-\boldsymbol{H}_{\gamma_f}(0)}{\overline{\|f\|^2}}\right)\boldsymbol{\delta x}\right) + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}(\|\boldsymbol{\delta x}\|^3), \tag{9.3.5}$$

where, by integration by parts:

$$-\boldsymbol{H}_{\gamma_f}(0) = -\left(\boldsymbol{\nabla}_{\boldsymbol{\delta x}}\boldsymbol{\nabla}_{\boldsymbol{\delta x}}^T\gamma_f(\boldsymbol{\delta x})\right)_{|\delta_x=0}, \tag{9.3.6}$$

$$= -\frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; f(\boldsymbol{x})\left(\boldsymbol{\nabla}_{\boldsymbol{\delta x}}\boldsymbol{\nabla}_{\boldsymbol{\delta x}}^T f(\boldsymbol{x}+\boldsymbol{\delta x})\right)_{|\delta_x=0}, \tag{9.3.7}$$

$$= -\frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; f(\boldsymbol{x})\boldsymbol{H}_f(\boldsymbol{x}), \tag{9.3.8}$$

$$= \frac{1}{S}\,\mathbb{E}\int_\Omega \mathrm{d}\boldsymbol{x}\; \boldsymbol{\nabla}f(\boldsymbol{x})(\boldsymbol{\nabla}f(\boldsymbol{x}))^T > 0 \text{ (in the Lowner sense).} \tag{9.3.9}$$

Hence, $-\overline{\|f\|^2}\boldsymbol{H}_{\gamma_f}^{-1}(0)$ encodes the square of the correlation lengths of the tracer, and the covariance can be approximated by a Gaussian function near 0. The Fourier transform of the covariance expression (9.3.5) provides the approximation if the spectrum tail, as:

$$\Gamma_f(\boldsymbol{k}) \triangleq \frac{1}{S}\,\mathbb{E}|\hat{T}_0(\boldsymbol{k})|^2, \tag{9.3.10}$$

$$= \frac{1}{S}\,\mathbb{E}\left|\int_\Omega \mathrm{d}\boldsymbol{x}\; f(\boldsymbol{x})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right|^2, \tag{9.3.11}$$

$$= \hat{\gamma}_f(\boldsymbol{k}), \tag{9.3.12}$$

$$\underset{\|\boldsymbol{k}\|\to\infty}{\sim} \frac{2\pi\left(\overline{\|f\|^2}\right)^2}{\det\left(\boldsymbol{H}_{\gamma_f}(0)\right)^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{k}^T\left(-\overline{\|f\|^2}\boldsymbol{H}_{\gamma_f}^{-1}(0)\right)\boldsymbol{k}\right). \tag{9.3.13}$$

Note that the evolution of the opposite of the Hessian $-\boldsymbol{H}_{\gamma_f}(0) = \overline{\boldsymbol{\nabla}f(\boldsymbol{\nabla}f)^T}$ is much more difficult to describe than its trace. Indeed, equation (9.3.9) yields:

$$\mathrm{tr}\left(-\boldsymbol{H}_{\gamma_f}(0)\right) = \overline{\mathrm{tr}\left(\boldsymbol{\nabla}f(\boldsymbol{\nabla}f)^T\right)} = \overline{\|\boldsymbol{\nabla}f\|^2}, \tag{9.3.14}$$

To let this gradient norm to appear, instead of the covariance Hessian, we consider the omnidirectional spectrum. The omnidirectional spectrum of the scalar f is defined as follows:

$$\overline{\Gamma}_f(\kappa) \triangleq \kappa\oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}}\;\Gamma_f(\boldsymbol{k}), \tag{9.3.15}$$

where $\boldsymbol{k} = \kappa\begin{pmatrix}\cos(\theta_{\boldsymbol{k}})\\ \sin(\theta_{\boldsymbol{k}})\end{pmatrix}$ is the wavenumber-vector and κ the wavenumber. A shown in the Appendix 9.A, the Taylor expansion (9.3.5) then leads to:

$$\overline{\Gamma}_f(\kappa) \underset{\|\boldsymbol{k}\|\to\infty}{\sim} C_f\exp\left(-\frac{1}{2}L_f^2\kappa^2\right), \tag{9.3.16}$$

where

$$L_f^2 = \frac{\overline{\|f\|^2}}{\overline{\|\boldsymbol{\nabla} f\|^2}} \text{ and } C_f = 2\left(\frac{(2\pi)^3\left(\overline{\|f\|^2}\right)^3}{\overline{\|\boldsymbol{\nabla} f\|^2}}\right)^{1/2}. \tag{9.3.17}$$

9.3.3 Tail change

Applied, to both the initial tracer, T_0, and the advected tracer, T, the approximation (9.3.16) yields:

$$\overline{\Gamma}_T(\kappa) \underset{\|\boldsymbol{k}\|\to\infty}{\sim} \overline{\Gamma}_{T_0}(\kappa)\,\frac{C_T}{C_{T_0}}\exp\left(-\frac{1}{2}\left(L_T^2 - L_{T_0}^2\right)\kappa^2\right), \tag{9.3.18}$$

$$\underset{\|\boldsymbol{k}\|\to\infty}{\sim} \overline{\Gamma}_{T_0}(\kappa)\left(\frac{\overline{\|\boldsymbol{\nabla} T_0\|^2}}{\overline{\|\boldsymbol{\nabla} T\|^2}}\right)^{1/2}\exp\left(\frac{1}{2}\overline{\|T_0\|^2}\left(\frac{1}{\overline{\|\boldsymbol{\nabla} T_0\|^2}} - \frac{1}{\overline{\|\boldsymbol{\nabla} T\|^2}}\right)\kappa^2\right), \tag{9.3.19}$$

where the simplification in the last asymptotic equivalence is due to the variance conservation (9.3.2). As explained in section 9.2, if the initial tracer T_0 and the flow are decorrelated, the tracer gradients strengthen: $\overline{\|\boldsymbol{\nabla} T\|^2} > \overline{\|\boldsymbol{\nabla} T_0\|^2}$ (*i.e.* $L_T < L_{T_0}$) and by (9.3.19) the tracer spectral tail raises. More precisely, using the estimate (9.2.60), we get for final expression:

$$\overline{\Gamma}_T(\kappa) \underset{\|\boldsymbol{k}\|\to\infty}{\sim} \overline{\Gamma}_{T_0}(\kappa)\left(1+\left(\frac{t}{\tau_G}\right)^2\right)^{-1/2}\exp\left(\frac{1}{2}\frac{\overline{\|T_0\|^2}}{\overline{\|\boldsymbol{\nabla} T_0\|^2}}\frac{\kappa^2}{1+\left(\frac{\tau_G}{t}\right)^2}\right). \tag{9.3.20}$$

This is illustrated for the toy flow in Figure (9.3.20). The Gaussian approximation successfully captures the spectrum tail shift towards small scales. The associated spatial fields have been presented in Figure (9.2).

As simplified, the initial spectrum tail can then be recovered using an isotropic Gaussian filtering of the tracer with a degree of smoothness controlled by the variance coefficient:

$$\frac{\overline{\|T_0\|^2}}{\overline{\|\boldsymbol{\nabla} T_0\|^2}}\frac{1}{1+\left(\frac{\tau_G}{t}\right)^2}. \tag{9.3.21}$$

This adaptive filtering is exemplified in Figure 9.13. The tracer is advected during a time t, and then smoothens by a Gaussian filter with the width (9.3.21). The combine effect of advection and filtering moves the large-scale structures, but keeps the global amount of small-scale structures stationary.

9.3.4 Self-similar approximation

Even though the previous Gaussian approximation is useful to link advection and filtering, spectra of geophysical tracer field are more likely self-similar, and possibly non-smooth. Moreover, it can be useful to target specific spectral slopes using the Lagrangian advection method. For these reasons, we consider the following form for the spectrum of a scalar f:

$$\overline{\Gamma}_f(\kappa) = \begin{cases} A\left(1+\frac{\kappa}{\kappa_m}\right)^{-\zeta} & \text{if } \kappa \leqslant \kappa_\infty \\ 0 & \text{otherwise} \end{cases}, \tag{9.3.22}$$

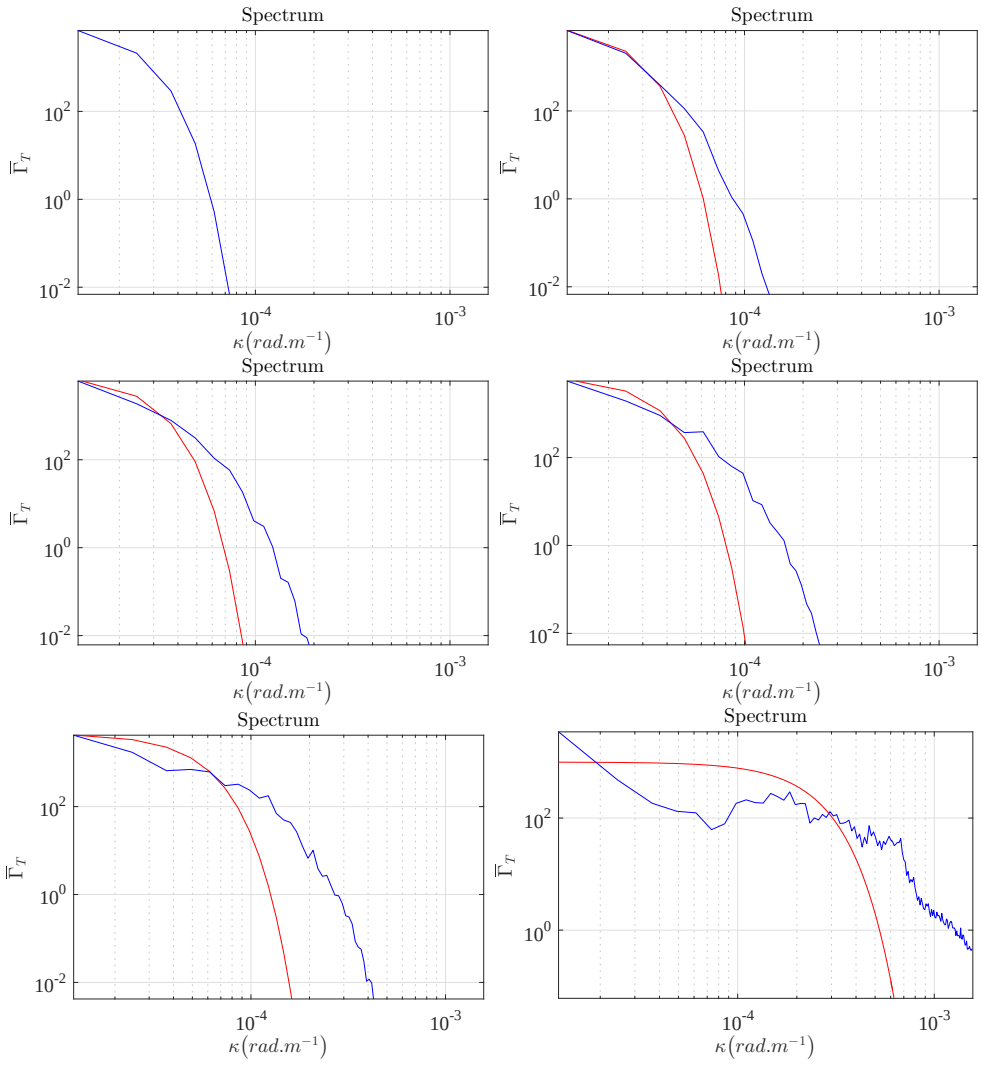


Figure 9.12: Omnidirectional spectra (blue) and its prediction using the Gaussian approximation (9.3.20) (red) for the toy model at $t = 0, 5, 10, 15, 30$ and 150 days. The associated spatial fields are displayed in Figure 9.2.

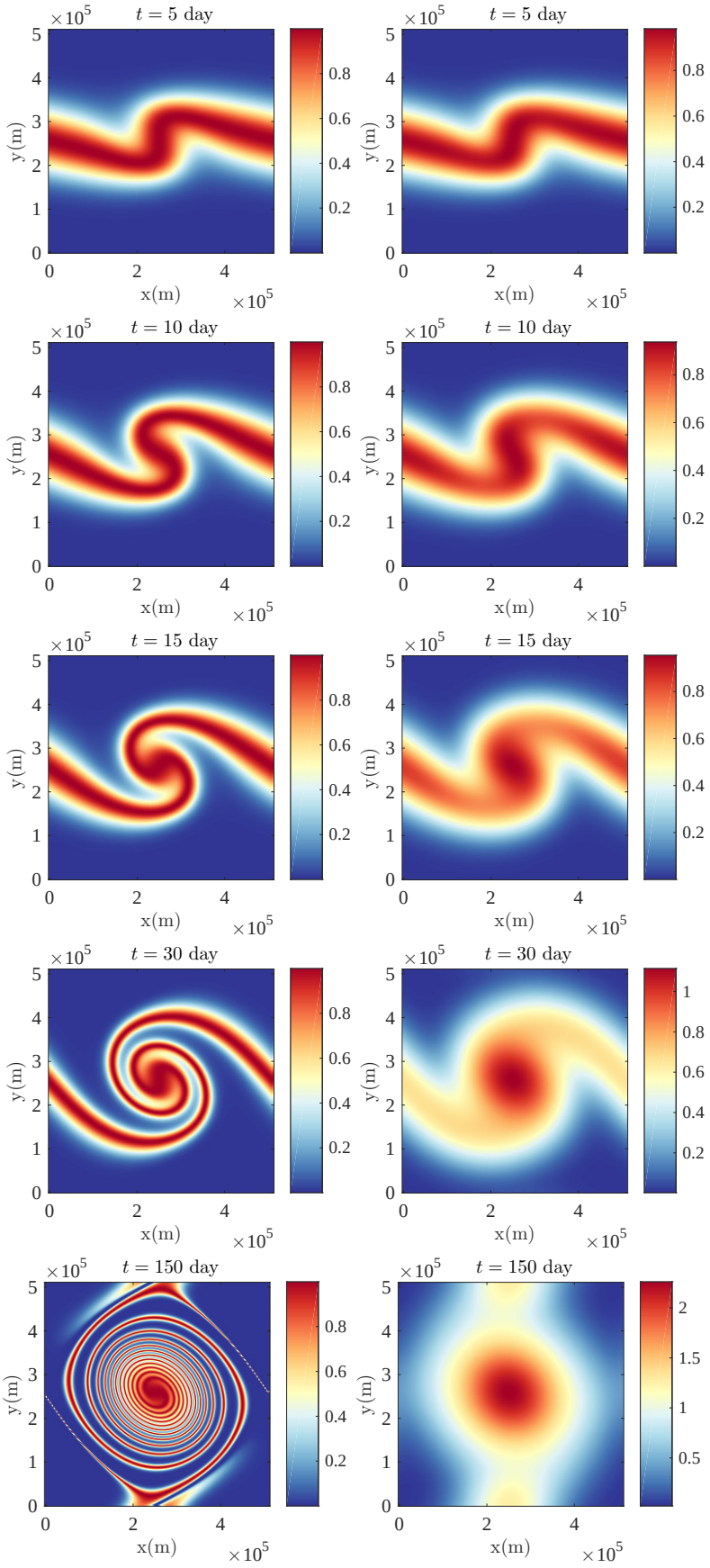


Figure 9.13: Tracer advected (left) and tracer advected and then smoothed by our adapted Gaussian filter (right) for the toy model at (from top to bottom) $t = 5, 10, 15, 30$ and 150 days.

Then,

$$L_f^2 = \frac{\overline{\|f\|^2}}{\overline{\|\boldsymbol{\nabla} f\|^2}}, \tag{9.3.23}$$

$$= \frac{\int_0^{\kappa_\infty} \left(1+\frac{\kappa}{\kappa_m}\right)^{-\zeta}}{\int_0^{\kappa_\infty} \kappa^2 \left(1+\frac{\kappa}{\kappa_m}\right)^{-\zeta}}, \tag{9.3.24}$$

$$= \frac{(\zeta-2)(\zeta-3)}{\kappa_m^2\left(2-\left(1+\frac{\kappa_\infty}{\kappa_m}\right)^{1-\zeta}\left((\zeta-1)(\zeta-2)\left(\frac{\kappa_\infty}{\kappa_m}\right)^2+2(\zeta-1)\frac{\kappa_\infty}{\kappa_m}+2\right)\right)}. \tag{9.3.25}$$

where κ_∞ is set by the numerical resolution. As long as the width of the inertial scale, $\kappa_\infty - \kappa_m$, is large enough, the above function is strictly positive and continuous w.r.t. the spectral slope, ζ, for all $\zeta > 1$.

For $f = T$ being the advected tracer, the wavenumber κ_m can encompass planetary length scales which does not vary much during the process of advection. This weak variation is due to the spatial correlation between tracer and flow (see (9.2.13)), but also to the relatively small advection time (mesoscale or submesocale time scales). The resolution, κ_∞, is constant as well. So, a targeted spectral slope, ζ, provides a length scale L_f to be reached on a given advection time, making use of (9.3.21).

Figure 9.14 illustrates this estimate, applied to the ACC SST field presented in the last section. As prescribed by the model, the spectrum tail slope of the advected SST reaches the value -3 after 5 days of advection, 2.5 after 10 days and -2 after 48 days. Figure 9.9 displays the spatial SST fields before and after advection. These data correspond to the summer (January in the southern hemisphere).

Interestingly, the seasonality variation can be studied. Figure 9.15 displays, for each day of the year 2011, the global stretching time, τ_G, the spectrum slope of the measured SST and the prescribed advection time to reach a -2.5 spectrum slope. As found, the stretching is faster during the winter, likely associated with a more intense eddy activity. The measured SST spectral slopes are relatively stationary (close to -4), whereas the small-scale velocity is expected to be stronger during winter. Accordingly, the prescribed advection time is smaller in wintertime.

Berti and Lapeyre (2014) proposed other Eulerian means to prescribe the advection time: the inverse of the vorticity Root Mean Square (RMS), $\left(\overline{(\boldsymbol{\nabla}^\perp \cdot \boldsymbol{v})^2}\right)^{-1/2}$, and of the velocity gradient RMS, $\left(\overline{\|\boldsymbol{\nabla}\boldsymbol{v}\|^2}\right)^{-1/2}$. The latter is directly linked to the shearing time (9.2.34). Yet, these estimates can encode shearing but not folding. Indeed, folding involves 2nd order derivatives of the velocity, such as to describe the curvature variation of adjacent streamlines (9.2.53). Moreover, these criteria do not depend on the initial nor on the resulting spectral slope. So, these criteria cannot fully control the necessary advection time, and shall likely lead to wrong resulting tracer spectral slopes. For instance, Figure 9.15 demonstrates that these criteria strongly underestimate the advection time needed to reach a -2.5 spectrum slope. Accordingly, these criteria mostly apply to very short advection time, with a resulting advected tracer already close to the true SST.

9.4 Estimation of eddy diffusivity

Considering satellite observations, geophysical tracers generally exhibit relatively stable spectra. The predicted raise of the tracer spectral tail, under multiple advection operations, shall thus be compensated. This can possibly be resulting from the combined effects of the well resolved, slow-varying and large-scale, velocity, and of an unresolved, fast-varying and likely small-scale, velocity. As the first velocity component will tend to raise the spectrum (9.3.20), the second component shall act to the mean tracer. As often suggested, this last process can be accounted for using a uniform eddy diffusivity, ν. As such, after an advection of Δt, it multiplies the spectrum by $\exp\left(-\nu\Delta t\kappa^2\right)$. To exactly compensate the spectrum increase (9.3.20) during Δt, and thus to keep the resulting

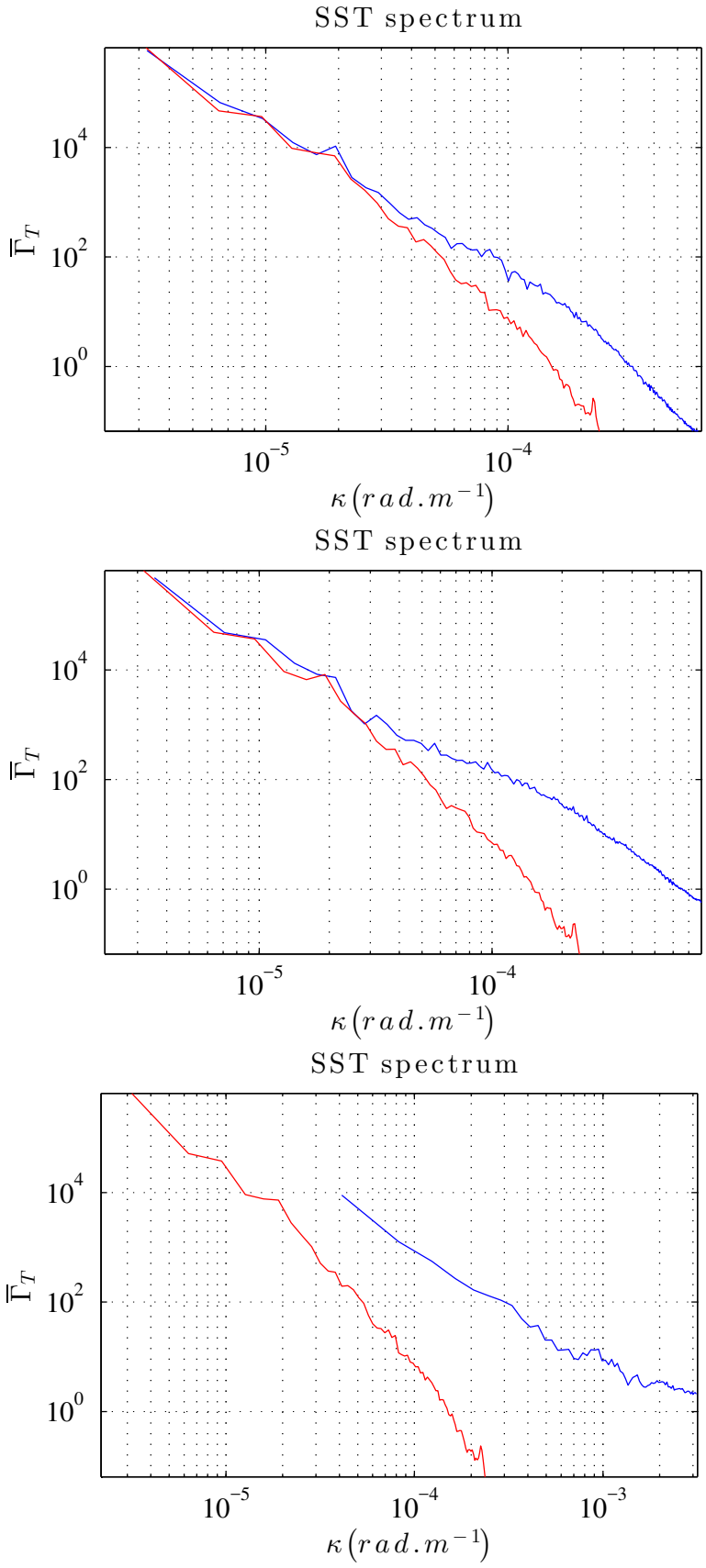


Figure 9.14: SST spectrum of the satellite data (red) and after prescribed advection (blue) with 5 days for a -3 spectrum slope (top), 10 days for a -2.5 spectrum slope (middle) and 48 days for a -2 spectrum slope (bottom). The associated spatial fields are presented in Figure 9.9. On the blue bottom spectrum, missing large-scale values are due to a strong deformation of the advected spatial domain by the jet (see Figure 9.9).

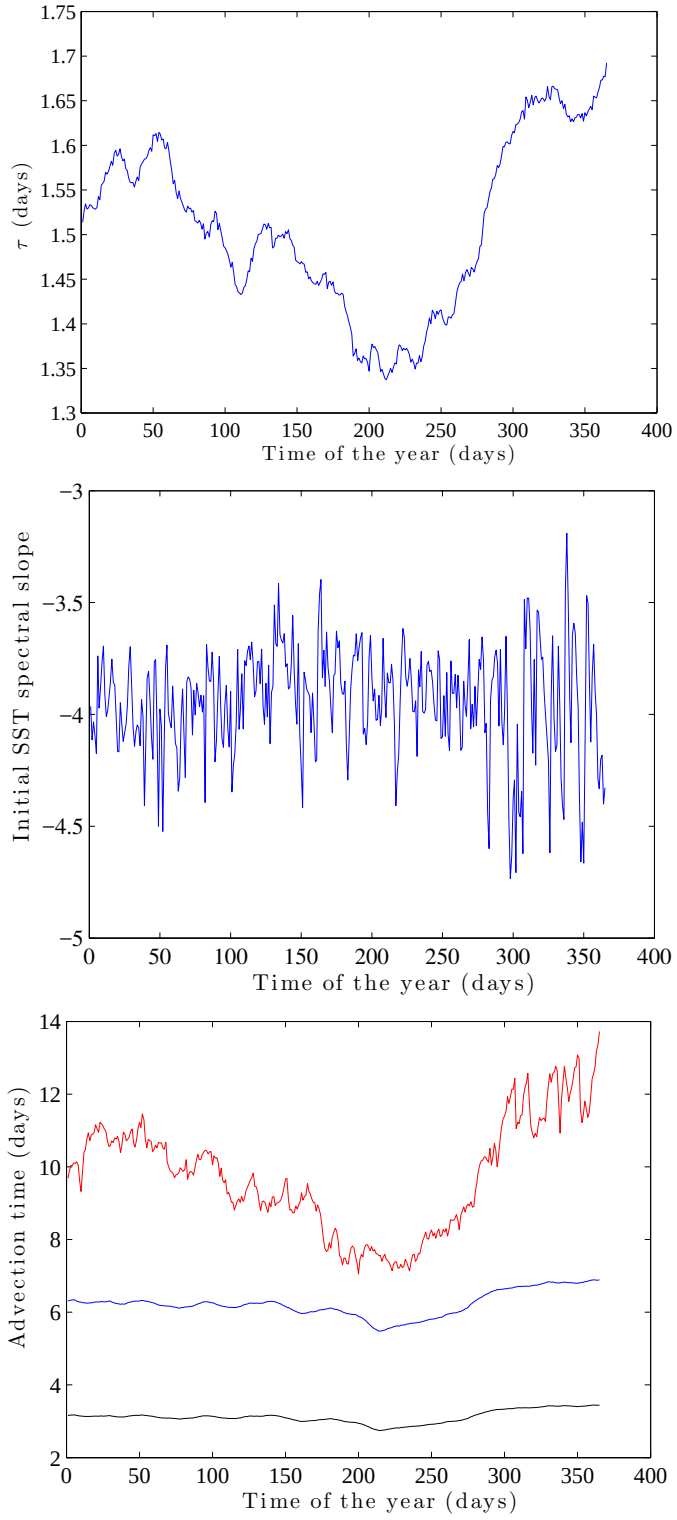


Figure 9.15: Global stretching time, τ_G (top in days), SST spectrum slope of Globcurrent data (middle) and prescribed advection time to reach a -2.5 spectrum slope (bottom in days) with the velocity gradient RMS (blue line), vorticity RMS (black line) and our model (red line).

advected tracer closer to its initial variance distribution over scales, the eddy diffusivity can then be set as:

$$\nu = \frac{1}{2\tau_G} \frac{\overline{\|T_0\|^2}}{\overline{\|\boldsymbol{\nabla} T_0\|^2}}\, c(t) \quad \text{with} \quad c(t) = \frac{\frac{\Delta t}{\tau_G}}{\left(1 + \left(\frac{\Delta t}{\tau_G}\right)^2\right)} = \begin{cases} \frac{\Delta t}{\tau_G} & \text{if } \Delta t \ll \tau_G \\ \frac{\tau_G}{\Delta t} & \text{if } \Delta t \gg \tau_G \end{cases} . \tag{9.4.1}$$

For small time time step, Δt, the eddy diffusivity is superdiffusive, whereas for large time step it is subdiffusive. Note also that in the second case, the direct cascade intensifies, decreasing the large-scale tracer energy as encoded by the multiplicative constant of (9.3.20):

$$\left(\frac{\overline{\|\boldsymbol{\nabla} T_0\|^2}}{\overline{\|\boldsymbol{\nabla} T\|^2}}\right)^{1/2} \approx \left(1 + \left(\frac{\Delta t}{\tau_G}\right)^2\right)^{-1/2} = \begin{cases} 1 & \text{if } \Delta t \ll \tau_G \\ \frac{\tau_G}{\Delta t} & \text{if } \Delta t \gg \tau_G \end{cases} . \tag{9.4.2}$$

Conclusion

Through this chapter, we analyzed how sets of points are stretched and folded by a smooth, possibly stationary, flow, creating strong tracer gradients and raising the spectral tail of the initial tracer spectral distribution.

Two characteristics of the flow influence the norm of advected tracer gradients: a local growth rate, associated with the eigenvalues of the Cauchy-Green tensor, and the orientation of the stable direction, eigenvector of the Cauchy-Green tensor. Integrated over space, the influence of local orientation disappears if the initial tracer is not correlated to the flow. The overall gradients can then only strengthen. Such a weak correlation applies for passive tracers, having negligible back effects on the flow, but such a weak correlation can also appear when the tracer is strongly smoothed before being advected. The growth rate can then be readily related to FTLE and mesochronic velocity. If the Eulerian velocity is stationary, the gradient strengthening is mainly due to locally uniform shears and stationary convective cells. The first process stretches pairs of points when the streamlines are straights and parallel. The second one folds and wraps the tracer around vortices. Indeed, different concentric orbits are associated with different temporal periods. A local, in space, temporal frequency can then be understood as an angular velocity, and its spatial variations deform the tracer structures. For both the velocity shears and angular velocity shears, the norm of tracer gradients linearly increases in time. Integrated over space, the square of this norm controls and specifies the time evolution of the tracer correlation length, the spectral tail and its slope. Two simplified models are proposed to approximate the tracer spectrum tail. These models only depend on the advection time, the Eulerian velocity and the initial tracer statistical characteristics. Numerical simulations on a toy model and real satellite images confirm the validity of these different approximations. In particular, we successfully estimate the spatial and spectral distribution, as well as the time evolution of the mixing processes. We apply the method to specify the advection time and the filter width of the Lagrangian advection method. Moreover, the proposed development can help the definition of an eddy diffusion coefficient as a function of the large-scale velocity. Finally, this work highlights the preponderant effect of folding in finite-time mixing.

Appendix

9.A Omnidirectional spectra of tracers for smooth flow

Hereafter, we will focus on the omnidirectional spectrum of a scalar f. This will enable us to replace $\boldsymbol{\nabla} f(\boldsymbol{\nabla} f)^T$ by $\|\boldsymbol{\nabla} f(\boldsymbol{x})\|^2$ in the expression of the spectrum (9.3.13).

$$\begin{aligned}
\overline{\Gamma}_f(\kappa) &= \kappa \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \Gamma_f(\boldsymbol{k}), \\
&= \kappa \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x}\ \gamma_f(\boldsymbol{\delta x}) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}, \\
&= \kappa \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(\gamma_f(0) - \frac{1}{2}\boldsymbol{\delta x}^T \frac{1}{S}\ \mathbb{E} \int_\Omega \mathrm{d}\boldsymbol{x}\ \boldsymbol{\nabla} f(\boldsymbol{x})(\boldsymbol{\nabla} f(\boldsymbol{x}))^T \boldsymbol{\delta x} + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}, \\
&= \frac{\kappa}{S}\ \mathbb{E} \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \int_\Omega \mathrm{d}\boldsymbol{x} \left(f^2(\boldsymbol{x}) - \frac{1}{2}\boldsymbol{\delta x}^T \boldsymbol{\nabla} f(\boldsymbol{x})(\boldsymbol{\nabla} f(\boldsymbol{x}))^T \boldsymbol{\delta x} + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}, \\
&= \frac{\kappa}{S}\ \mathbb{E} \int_\Omega \mathrm{d}\boldsymbol{x} \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(f^2(\boldsymbol{x}) - \frac{1}{2}\left(\frac{\boldsymbol{\nabla} f(\boldsymbol{x})}{\|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2} \cdot \boldsymbol{\delta x} \right)^2 \|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2^2 + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}.
\end{aligned}$$

Locally in $\boldsymbol{x}$, we can define a variable change for $\boldsymbol{\delta x} = (\delta x_1\ \delta x_2)^T$. We apply the rotation matrix $\boldsymbol{U}(\boldsymbol{x}) = \frac{1}{\|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2}\left[\boldsymbol{\nabla} f(\boldsymbol{x})\ \boldsymbol{\nabla}^\perp f(\boldsymbol{x})\right]$ to $\boldsymbol{\delta x}$ to align $\boldsymbol{\delta x}$ with the tracer gradient and denote $\theta(\boldsymbol{x})$ the angle of the associated rotation:

$$\begin{aligned}
\overline{\Gamma}_f(\kappa) &= \frac{\kappa}{S}\ \mathbb{E} \int_\Omega \mathrm{d}\boldsymbol{x} \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(f^2(\boldsymbol{x}) - \frac{1}{2}(\delta x_1)^2 \|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2^2 + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i(\boldsymbol{U}^T\boldsymbol{k})\cdot\boldsymbol{\delta x}}, \\
&= \frac{\kappa}{S}\ \mathbb{E} \int_\Omega \mathrm{d}\boldsymbol{x} \oint_{[\theta(\boldsymbol{x}),\theta(\boldsymbol{x})+2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(f^2(\boldsymbol{x}) - \frac{1}{2}(\delta x_1)^2 \|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2^2 + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}, \\
&= \frac{\kappa}{S}\ \mathbb{E} \int_\Omega \mathrm{d}\boldsymbol{x} \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(f^2(\boldsymbol{x}) - \frac{1}{2}(\delta x_1)^2 \|\boldsymbol{\nabla} f(\boldsymbol{x})\|_2^2 + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}, \\
&= \kappa \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x} \left(\overline{\|f\|^2} - \frac{1}{2}(\delta x_1)^2 \overline{\|\boldsymbol{\nabla} f\|^2} + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) e^{-i\boldsymbol{k}\cdot\boldsymbol{\delta x}}.
\end{aligned}$$

The third equality above is due to the averaging over the spatial frequency angle $\theta_{\boldsymbol{k}}$. Indeed, $\boldsymbol{U}^T\boldsymbol{k}$ is just a rotation of $\boldsymbol{k}$. And, integrating over $[0, 2\pi]$ or over $[\theta(\boldsymbol{x}), 2\pi + \theta(\boldsymbol{x})]$ is the same thing, since it leads to the same closed line: a circle of radius κ.

$$\begin{aligned}
\overline{\Gamma}_f(\kappa) &= \kappa \oint_{[0,2\pi]} \mathrm{d}\theta_{\boldsymbol{k}} \int_\Omega \mathrm{d}\boldsymbol{\delta x}_1 \mathrm{d}\boldsymbol{\delta x}_2 \left(\overline{\|f\|^2} \exp\left(-\frac{1}{2}\frac{\overline{\|\boldsymbol{\nabla} f\|^2}}{\overline{\|f\|^2}}(\delta x_1)^2 \right) + \underset{\|\boldsymbol{\delta x}\|\to 0}{o}\left(\|\boldsymbol{\delta x}\|^3\right) \right) \\
&\qquad \times e^{-ik_1\delta x_1} e^{-ik_2\delta x_2}, \\
&\underset{\|\boldsymbol{k}\|\to\infty}{\sim} \oint_{[0,2\pi]} \kappa \mathrm{d}\theta_{\boldsymbol{k}}\ \frac{C_f}{2(2\pi)}\ \exp\left(-\frac{1}{2}L_f^2 k_1^2 \right) (2\pi)\delta(k_2),
\end{aligned}$$

where $\mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} \kappa\cos(\theta_{\boldsymbol{k}}) \\ \kappa\sin(\theta_{\boldsymbol{k}}) \end{pmatrix}$, $L_f^2 = \frac{\overline{\|f\|^2}}{\overline{\|\boldsymbol{\nabla} f\|^2}}$ and $\frac{C_f}{2(2\pi)} = \left(\frac{2\pi\left(\overline{\|f\|^2}\right)^3}{\overline{\|\boldsymbol{\nabla} f\|^2}} \right)^{1/2}$. Note that the asymptotic equivalence is an approximation. Then, switching from cylindrical to Cartesian coordinates in each half rings $\{\boldsymbol{k} \in \mathbb{R}^2 | k_1 \leqslant 0, \|\boldsymbol{k}\| = \kappa\}$ and $\{\boldsymbol{k} \in \mathbb{R}^2 | k_1 \geqslant 0, \|\boldsymbol{k}\| = \kappa\}$ yields:

$$\begin{aligned}
\overline{\Gamma}_f(\kappa) &\underset{\|\boldsymbol{k}\|\to\infty}{\sim} 2\int_{-\kappa}^{\kappa} \frac{dk_2}{\sqrt{1 - \left(\frac{k_2}{\kappa}\right)^2}} \frac{C_f}{2}\ \exp\left(-\frac{1}{2}L_f^2\left(\kappa^2 - k_2^2\right) \right) \delta(k_2), \\
&\underset{\|\boldsymbol{k}\|\to\infty}{\sim} C_f \exp\left(-\frac{1}{2}L_f^2\kappa^2 \right).
\end{aligned}$$

Chapter 10

Additional results

In this chapter, we use both Itō (terms with $\mathrm{d}B_t$) and Stratonovich notations (terms with $\circ\mathrm{d}B_t$). For descriptions and comparisons of these notations, the reader can refer to the section 1.6 of this thesis.

10.1 Models under location uncertainty with Stratonovich notations

10.1.1 Flow

Thanks to the formula (1.6.1) of chapter 1, the SDE of the stochastic flow, $\boldsymbol{X}$, can be written in Stratonovich form as follow:

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}(\boldsymbol{X}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t,t)\mathrm{d}\boldsymbol{B}_t \Longleftrightarrow \begin{cases} \mathrm{d}\boldsymbol{X}_t = \boldsymbol{w}_S(\boldsymbol{X}_t,t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{X}_t,t)\circ\mathrm{d}\boldsymbol{B}_t, \\ \text{with } \boldsymbol{w}_S = \boldsymbol{w} - \frac{1}{2}\sum_{i=1}^{d}\partial_i\boldsymbol{\sigma}\boldsymbol{\sigma}_{i\bullet}^T. \end{cases} \tag{10.1.1}$$

10.1.2 Transport

The Stratonovich form of the stochastic material derivative is:

$$D_t\Theta = \mathrm{d}_t\Theta + ((\boldsymbol{w}_S\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla})\Theta. \tag{10.1.2}$$

Moreover, let us consider a stochastic transport equation of a quantity Θ:

$$\mathbb{D}_t\Theta = f\mathrm{d}t + g\mathrm{d}B_t', \tag{10.1.3}$$

where $\boldsymbol{B}_t$ and B_t' are independent, and g is differentiable in time and does not depend on Θ. Then, according to equation (2.2.10) of chapter 2, the stochastic transport operator $\mathbb{D}_t\Theta$ and the material derivative $D_t\Theta$ coincide and:

$$\mathbb{D}_t\Theta = \mathrm{d}_t\Theta + ((\boldsymbol{w}_S\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla})\Theta. \tag{10.1.4}$$

We propose two different proofs:

Proof 1, using the Stratonovich form of the Ito-Wentzell formula

To derive equation (10.1.2), we write the definition of the material derivative in Lagrangian coordinates and apply the Stratonovich form of the Ito-Wentzell formula (Kunita, 1997; Chow, 2014) to Θ with the Stratonovich flow (10.1.1):

$$D_t\Theta(\boldsymbol{X}_t,t) = \mathrm{d}\left[\Theta(\boldsymbol{X}_t,t)\right] = \left[\mathrm{d}_t\Theta + ((\boldsymbol{w}_S\mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t)\cdot\boldsymbol{\nabla})\Theta\right](\boldsymbol{X}_t,t), \tag{10.1.5}$$

Equation (10.1.4) follows by (2.2.10).

Proof 2, using formula (1.6.3) **to go from Itō to Stratonovich notations**

Here, we need to assume a particular form for the forcing (equation (10.1.3)). By definition of the stochastic transport operator (2.2.9) in chapter 2 and by the formula (1.6.3) of chapter 1:

$$\begin{aligned}
\mathrm{d}_t\Theta &= -\left(\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta + \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)\mathrm{d}t + f\mathrm{d}t + g\mathrm{d}B'_t, && (10.1.6)\\
&= -\left(\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma}\circ\mathrm{d}\boldsymbol{B}_t\right)\cdot\boldsymbol{\nabla}\Theta + \boldsymbol{\nabla}\cdot\left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\Theta\right)\mathrm{d}t + f\mathrm{d}t + g\circ\mathrm{d}B'_t && (10.1.7)\\
&\quad -\frac{1}{2}\sum_k \mathrm{d}<-(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla})\Theta, B_k> -\frac{1}{2}\mathrm{d}<g, B'>. && (10.1.8)
\end{aligned}$$

The first quadratic covariation is itself evaluated using the Itō evolution law (10.1.6):

$$\begin{aligned}
\sum_k \mathrm{d}<-(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla})\Theta, B_k>_t &= \sum_k (\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla})\left[(\boldsymbol{\sigma}_{\bullet k}\cdot\boldsymbol{\nabla})\Theta\right]\mathrm{d}t && (10.1.9)\\
&= \boldsymbol{\nabla}\cdot\{\boldsymbol{a}\boldsymbol{\nabla}\Theta\}\,\mathrm{d}t - (\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T)\cdot\boldsymbol{\nabla}\Theta\mathrm{d}t. && (10.1.10)
\end{aligned}$$

The second quadratic covariation is trivial since we assumed that g is differentiable w.r.t. time (g has finite variations):

$$\mathrm{d}<g, B'_t>=0. \tag{10.1.11}$$

After cancellation of the diffusion terms in (10.1.8), we recognize the expression of the Stratonovich drift (10.1.1):

$$\boldsymbol{w}_S = \boldsymbol{w} - \frac{1}{2}\sum_{i=1}^{d}\partial_i\boldsymbol{\sigma}\boldsymbol{\sigma}_{i\bullet}^T = \boldsymbol{w}^\star - \frac{1}{2}\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T. \tag{10.1.12}$$

This leads to the result.

Note that the Stratonovich form (10.1.2) of the stochastic transport equation can also be derived from the Itō form (10.1.3) with the more classical formula (1.6.1) of chapter 1. Nevertheless, since the scalar Θ is a function living in an infinite-dimensional space, the noise term, $\boldsymbol{\Sigma}(\Theta)\mathrm{d}\boldsymbol{B}_t = \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\Theta$, is here a functional of the function Θ. Thus, in formula (1.6.1), the partial derivatives, $\partial_{\boldsymbol{x}_i}$, and the scalar product of $\mathbb{R}^d$, $\boldsymbol{X}\cdot\boldsymbol{Y}$, appearing in $((\boldsymbol{\Sigma}_{\bullet p}\cdot\boldsymbol{\nabla})\boldsymbol{\Sigma}_{\bullet q})$ have to be replaced by the variational derivatives, $\frac{\delta}{\delta\Theta(\boldsymbol{y})}$, and the scalar product of $\mathcal{L}^2(\Omega)$, $\int_\Omega fg$, respectively.

10.1.3 Modified drift

The results (10.1.1) and (10.1.4) enable to better understand the form of the modified drift, $\boldsymbol{w}^\star$, involved in the models under location uncertainty:

$$\begin{aligned}
\boldsymbol{w}^\star &\triangleq \boldsymbol{w} - \frac{1}{2}\boldsymbol{\nabla}\cdot\boldsymbol{a} + \boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T, && (10.1.13)\\
&= \underbrace{\boldsymbol{w} - \frac{1}{2}\sum_{i=1}^{d}\partial_i\boldsymbol{\sigma}\boldsymbol{\sigma}_{i\bullet}^T}_{\substack{=\boldsymbol{w}_S\\ \text{(Lagrangian)}\\ \text{Stratonovich drift}}} + \underbrace{\frac{1}{2}\boldsymbol{\sigma}(\boldsymbol{\nabla}\cdot\boldsymbol{\sigma})^T}_{\substack{\text{Due to Ito notation}\\ \text{in the evolution law}\\ \text{(as the diffusion)}}}. && (10.1.14)
\end{aligned}$$

In particular, for incompressible flow, the operator $\boldsymbol{\sigma}$ is divergence-free and the modified drift, $\boldsymbol{w}^\star$, is the Stratonovich large-scale velocity.

10.2 Comparison with Holm (2015)

From a Lagrangian mechanics approach, Holm (2015) derived several stochastic fluid dynamics models. Hereafter, we detail their similarities and the distinctions with the models under location uncertainty. To simplify the comparison, we work with the Stratonovich notations.

For a scalar tracer – like the buoyancy, b, under the Boussinesq approximation – Holm (2015) prescribes the same type of evolution law than the modeling under location uncertainty:

$$\mathbb{D}_t b = 0. \tag{10.2.1}$$

Nonetheless, the incompressible stochastic transports of velocity and vorticity differ from our approach. First, Holm (2015) considers the transport – up to some forcings – of the Stratonovich large-scale linear momentum, $\rho \boldsymbol{w}_S = \rho \boldsymbol{w}^\star$, whereas our Navier-Stokes under location uncertainty (8.1.1) assumes the transport of the Itō large-scale linear momentum, $\rho \boldsymbol{w}$. Furthermore, due to the choice of its Lagrangian, the stochastic representation of Holm (2015) involves an additional term. Specifically, our stochastic Navier-Stokes equation (8.1.1) with neither viscosity nor Coriolis force reads:

$$[\mathrm{d}_t + ((\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla})] \underbrace{\boldsymbol{w}}_{\substack{\text{Due to the}\\ \text{transport}\\ \text{of } \rho\boldsymbol{w}}} = -\frac{1}{\rho}\boldsymbol{\nabla} p_1 \mathrm{d}t, \tag{10.2.2}$$

whereas the model of Holm (2015) is:

$$[\mathrm{d}_t + ((\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla})] \underbrace{\boldsymbol{w}^\star}_{\substack{\text{Due to the}\\ \text{transport}\\ \text{of } \rho\boldsymbol{w}^\star}} + \underbrace{\sum_{k=1}^{d} \boldsymbol{\nabla}(w_k^\star \mathrm{d}t + \boldsymbol{\sigma}_{k\bullet} \circ \mathrm{d}\boldsymbol{B}_t) w_k^\star}_{\text{Additional term}} = -\frac{1}{\rho}\boldsymbol{\nabla} p_2 \mathrm{d}t, \tag{10.2.3}$$

where p_1 and p_2 both represent the pressure but do not necessarily coincide. The additional term has at least two major consequences.

First, as derived in Holm (2015), this term leads to an evolution law of the vorticity, $\boldsymbol{\omega}^\star = \boldsymbol{\nabla} \times \boldsymbol{w}^\star$, similar to the deterministic one. Accordingly, the models of Holm (2015) conserves the helicity and the Kelvin theorem remains. In particular, in 2D, the transport of 2D vorticity $\omega^\star = \boldsymbol{\nabla}^\perp \cdot \boldsymbol{w}^\star$ reads:

$$[\mathrm{d}_t + ((\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla})]\, \omega^\star = 0, \tag{10.2.4}$$

whereas the transport of vorticity $\omega = \boldsymbol{\nabla}^\perp \cdot \boldsymbol{w}$ under location uncertainty involves a source term:

$$[\mathrm{d}_t + ((\boldsymbol{w}^\star \mathrm{d}t + \boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t) \cdot \boldsymbol{\nabla})]\, \omega = -\mathrm{tr}\left[\boldsymbol{\nabla}^\perp (\boldsymbol{\sigma} \circ \mathrm{d}\boldsymbol{B}_t)^T \boldsymbol{\nabla} \boldsymbol{w}^T\right]. \tag{10.2.5}$$

As expected, under moderate uncertainty assumptions (see chapter 3 for more details), the QG PV evolution law (see equations (3.3.12) and (3.B.38)) involves a similar forcing term. The paragraph 3.3.2 describes the physical meaning and the influence of this term for the QG PV dynamics.

The second consequence of the additional term in the Navier-Stokes model (10.2.3) is the lost of the large-scale kinetic energy conservation. In the following, we will proof this claim. To be coherent with the modeling, the kinetic energy is defined as $KE = 1/2\|\boldsymbol{w}^\star\|^2_{(\mathcal{L}^2(\Omega))^d}$. Using the Stratonovich Navier-Stokes model (10.2.3) and several integrations by parts:

$$\mathrm{d}\ KE = \int_\Omega \boldsymbol{w}^\star \cdot \mathrm{d}_t \boldsymbol{w}^\star = \sum_k \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \boldsymbol{w}^\star\ (\circ \mathrm{d}\boldsymbol{B}_t)_k, \tag{10.2.6}$$

where $S_k^\sigma = \frac{1}{2}\left(\left[\boldsymbol{\nabla}\left(\boldsymbol{\sigma}_{\bullet k}\right)^T\right] + \left[\boldsymbol{\nabla}\left(\boldsymbol{\sigma}_{\bullet k}\right)^T\right]^T\right)$ defined the strain-rate tensor of the unresolved velocity. The kinetic energy variation has *a priori* non-zero mean since it is a Stratonovich noise. With the formula (1.6.3), we can turn to the Itō notation in order to explicit this noise:

$$\mathrm{d}\ KE = \sum_k \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \boldsymbol{w}^\star\ (\mathrm{d}\boldsymbol{B}_t)_k + \frac{1}{2}\sum_k \mathrm{d}\left\langle \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \boldsymbol{w}^\star, (\boldsymbol{B}_t)_k \right\rangle. \tag{10.2.7}$$

Then the Navier-Stokes model (10.2.3) implies:

$$\frac{1}{2}\mathrm{d}\left\langle \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \boldsymbol{w}^\star, (\boldsymbol{B}_t)_k \right\rangle = \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \mathrm{d}\left\langle \boldsymbol{w}^\star, (\boldsymbol{B}_t)_k \right\rangle, \tag{10.2.8}$$

$$= \int_\Omega (\boldsymbol{w}^\star)^T S_k^\sigma \left((\boldsymbol{\sigma}_{\bullet k} \cdot \boldsymbol{\nabla})\boldsymbol{w}^\star + \boldsymbol{\nabla}\boldsymbol{\sigma}_{\bullet k}^T \boldsymbol{w}^\star\right) \mathrm{d}t, \tag{10.2.9}$$

To simplify this expression, we may assume that the unresolved velocity is isotropic and homogeneous in space. In this case, all terms of the form $\partial_p \sigma_{qk} \sigma_{jk}$ and of the form $\sum_q \partial_p \sigma_{qk} \partial_q \sigma_{jk} = \sum_q \partial_q (\partial_p \sigma_{qk} \sigma_{jk})$ cancel and the expectation of the kinetic energy budget gives:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\{KE\} = \frac{1}{2}\sum_k \mathbb{E}\|\boldsymbol{\nabla}\boldsymbol{\sigma}_{\bullet k}^T \boldsymbol{w}^\star\|^2_{(\mathcal{L}^2(\Omega))^d} > 0. \tag{10.2.10}$$

To conclude, the model of Holm (2015) does not conserve mean kinetic energy in the general case. Thus, it does not conserve the kinetic energy of each realization. In contrast, the model under location uncertainty always conserves the mean kinetic energy and the kinetic energy of each realization (see equation (8.1.9)).

10.3 Review of the derivation of the models under location uncertainty

In this section, we recall the derivation of geophysical models under location uncertainty presented in chapters 2, 3 and 4. After the general fluid dynamics models, we detail the Boussinesq, QG and SQG models.

All along the following development, the small-scale random flow component will be assume incompressible, *i.e.* associated with a divergence-free diffusion tensor:

$$\boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = 0. \tag{10.3.1}$$

This assumption remains realistic for the geophysical and incompressible flows considered in this thesis, and does not prevent the resolved velocity component (and therefore the whole field) to be compressible.

10.3.1 Mass conservation

Using the transport theorem (2.2.28), mass conservation for arbitrary volumes rules the stochastic transport of the fluid density, denoted ρ:

Continuity equation

$$D_t \rho + \rho \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t = 0. \tag{10.3.2}$$

10.3.2 Active scalar conservation law

The transport theorem (2.2.28) applied to a quantity $\rho\Theta$ describes the rate of change of the scalar Θ and is generally balanced by a production/dissipation term, as:

$$D_t(\rho\Theta) + \rho\Theta \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t = \rho \mathcal{F}_\Theta(\Theta)\mathrm{d}t. \tag{10.3.3}$$

We can apply the product rule

$$D_t(\rho\Theta) = D_t\rho\Theta + \rho D_t\Theta, \tag{10.3.4}$$

since $\frac{D(\rho\Theta)}{Dt}$ and $\frac{D(\rho)}{Dt}$ are correlated in time. Then, the mass conservation (10.3.2) simplifies the transport evolution of the scalar:

$$D_t\Theta = \mathbb{D}_t\Theta = \mathcal{F}_\Theta(\Theta)\mathrm{d}t. \tag{10.3.5}$$

As in the deterministic case, the 1st law of thermodynamics implies temperature conservation ($\Theta = T$) and the conservation of the amount of substance implies the conservation of salinity

$(\Theta = S)$:

Thermodynamics equations

$$\mathbb{D}_t T = \mathcal{F}_T(T)\mathrm{d}t, \tag{10.3.6a}$$
$$\mathbb{D}_t S = \mathcal{F}_S(S)\mathrm{d}t. \tag{10.3.6b}$$

10.3.3 Navier-Stokes model under location uncertainty

2nd Newton's law in the distribution sense

Then, we involve an informal expression of the 2nd Newton's law of motion. In a Galilean frame, the variation of the linear momentum is due to some external actions:

$$\mathrm{d}\int_{\mathcal{V}(t)} \rho\, \frac{\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t} = \int_{\mathcal{V}(t)} \mathrm{d}_t \boldsymbol{F}. \tag{10.3.7}$$

The left-hand term must be interpreted in a distribution sense, the small-scale velocity, $\boldsymbol{\sigma}\dot{\boldsymbol{B}} = \frac{\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t}{\mathrm{d}t}$, being a white noise in time. For every test function h, we have:

$$\int_{\mathbb{R}_+} h(t)\mathrm{d}\int_{\mathcal{V}(t)} \rho\boldsymbol{w} - \int_{\mathbb{R}_+} \frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)} \rho\, \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)} \mathrm{d}_t\boldsymbol{F}. \tag{10.3.8}$$

Both sides of this equation must have the same structure and the forces can be written as:

$$\int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)} \mathrm{d}_t\boldsymbol{F} = -\int_{\mathbb{R}_+} \frac{\mathrm{d}h}{\mathrm{d}t}(t)\int_{\mathcal{V}(t)} \rho\, \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t + \int_{\mathbb{R}_+} h(t)\int_{\mathcal{V}(t)} (\boldsymbol{\eta}\mathrm{d}t + \boldsymbol{\gamma}\mathrm{d}\boldsymbol{B}_t)\,. \tag{10.3.9}$$

The right-hand first term must compensate for the white-noise distributional differentiation of (10.3.8), whereas the last term of (10.3.9) provides the structure of the forces under location uncertainty.

The Reynolds transport theorem (2.2.28) applied to the (Itō) linear momentum, $\rho\boldsymbol{w}$, the incompressibility constraint (10.3.1) and the continuity equation (10.3.2) give as in the deterministic case:

$$\mathrm{d}\int_{\mathcal{V}(t)} \rho\boldsymbol{w} = \int_{\mathcal{V}(t)} (D_t(\rho\boldsymbol{w}) + \rho\boldsymbol{w}\boldsymbol{\nabla}\cdot\boldsymbol{w}^\star\mathrm{d}t) = \int_{\mathcal{V}(t)} \rho D_t\boldsymbol{w}. \tag{10.3.10}$$

This result being true for any volume $\mathcal{V}(t)$, we can remove the integral and use the previous results (10.3.8) and (10.3.9) :

$$D_t\boldsymbol{w} \;=\; \frac{1}{\rho}\boldsymbol{\eta}\mathrm{d}t + \frac{1}{\rho}\boldsymbol{\gamma}\mathrm{d}\boldsymbol{B}_t. \tag{10.3.11}$$

Eulerian formulation

The pressure is decomposed into a continuous-in-time component continuous in time, p, and time-uncorrelated component, $\dot{p}_\sigma = \frac{\mathrm{d}_t p_\sigma}{\mathrm{d}t}$. Accordingly the forcing reads:

$$\left\{\begin{array}{rcl} \boldsymbol{\eta} &=& -\rho(\boldsymbol{f}\times\boldsymbol{w}) + \rho\boldsymbol{g} - \boldsymbol{\nabla}p + \rho\nu\Delta\boldsymbol{w}, \\ \boldsymbol{\gamma}\mathrm{d}\boldsymbol{B}_t &=& -\rho(\boldsymbol{f}\times\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) - \nabla\mathrm{d}_t p_\sigma + \rho\nu\Delta\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t. \end{array}\right. \tag{10.3.12}$$

Then, from (10.3.11) and the formula (2.2.10) the stochastic Navier-Stokes model under location uncertainty is:

$$\begin{aligned} \mathbb{D}_t\boldsymbol{w} + \boldsymbol{f}\times(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \boldsymbol{g}\mathrm{d}t - \frac{1}{\rho}\boldsymbol{\nabla}(p\mathrm{d}t + \mathrm{d}_t p_\sigma) + \nu\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) \\ - \mathrm{tr}\left((\boldsymbol{\sigma}^T\boldsymbol{\nabla})\left(-\boldsymbol{f}\times\boldsymbol{\sigma} + \nu\Delta\boldsymbol{\sigma}\right)\right)\mathrm{d}t + \mathrm{d}_t\left\langle \int_0^t (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_{t'}\cdot\boldsymbol{\nabla}), \int_0^t \frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}_{t'}p_\sigma \right\rangle. \end{aligned} \tag{10.3.13}$$

In all the thesis, we neglect the correlations of the second line. With this simplification, the Navier-Stokes representation reads:

Navier-Stokes equations under location uncertainty with noise

$$\mathbb{D}_t \boldsymbol{w} + \boldsymbol{f} \times (\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t) = \boldsymbol{g}\mathrm{d}t - \frac{1}{\rho}\boldsymbol{\nabla}(p\mathrm{d}t + \mathrm{d}_t p_\sigma) + \nu\Delta(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t). \tag{10.3.14}$$

Note however that the correlations of (10.3.13) may be useful to describe a Stokes drift.

Another (possibly simpler) way to derive all the previous equations of this section is to use Stratonovich notations. Indeed, Stratonovich calculus is similar to the deterministic calculus. However, after that formal derivation, one would need to go from Stratonovich notations to Itō notations with the formula (1.6.3). We recall that only Itō stochastic differential equations can be simulated numerically.

Navier-Stokes model under location uncertainty without noise

For computational reasons, a deterministic LES-like model is sometime more adequate than a random model. For this purpose, Mémin (2014) assumed the drift $\boldsymbol{w}$ to be a smooth function of time. As such, its evolution law cannot involve white noise terms. The stochastic Navier-Stokes model (10.3.14) can hence be split. On one side, the evolution of the drift is determined by the smooth terms (in $\mathrm{d}t$). On the other side, the "noise terms" (in $\mathrm{d}\boldsymbol{B}_t$) must cancel each others. This intuitive argument is rigorously supported by the uniqueness of the semimartingale decomposition (Kunita, 1997) in stochastic calculus.

Navier-Stokes equations under location uncertainty without noise

$$\partial_t \boldsymbol{w} + (\boldsymbol{w}^\star \cdot \boldsymbol{\nabla})\boldsymbol{w} - \boldsymbol{\nabla} \cdot \left(\frac{1}{2}\boldsymbol{a}\boldsymbol{\nabla}\boldsymbol{w}\right) + \boldsymbol{f} \times \boldsymbol{w} = \boldsymbol{g} - \frac{1}{\rho}\boldsymbol{\nabla}p + \nu\Delta\boldsymbol{w}, \tag{10.3.15a}$$

$$(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t \cdot \boldsymbol{\nabla})\boldsymbol{w} + \boldsymbol{f} \times \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t = -\frac{1}{\rho}\boldsymbol{\nabla}\mathrm{d}_t p_\sigma + \nu\Delta\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t. \tag{10.3.15b}$$

10.3.4 Boussinesq equations

For ocean and atmosphere flows, a partition of the density and pressure is generally considered:

$$\rho = \rho_b + \rho_0(z) + \rho'(x, y, z, t), \tag{10.3.16a}$$

$$p = \widetilde{p}(z) + p'(x, y, z, t). \tag{10.3.16b}$$

Fields $\widetilde{\rho}(z) = \rho_b + \rho_0(z)$ and $\widetilde{p}(z)$ correspond to the density and the pressure at equilibrium (without any motion), respectively; they are deterministic functions and depend on the height only. The pressure and density departures, p' and ρ', are random functions, depending on the uncertainty component. From the expression of the vertical velocity component (10.3.14), the equilibrium fields are related through an hydrostatic balance:

$$\frac{\partial \widetilde{p}}{\partial z} = -g\widetilde{\rho}(z). \tag{10.3.17}$$

Within small density fluctuations (*i.e.* the Boussinesq approximation) as observed in the ocean, the stochastic mass conservation reads

$$0 = \mathbb{D}_t \rho + \rho\boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t \approx \rho_b \boldsymbol{\nabla} \cdot \boldsymbol{w}^* \mathrm{d}t. \tag{10.3.18}$$

This implies that the flow is volume-preserving.

According to equations (10.3.6a) and (10.3.6b), temperature and salinity are transported by the random flow. If those tracers do not oscillate too much, the density anomaly, $\rho - \rho_b$, can be approximated by a linear combination of these two properties. And thus, in the Boussinesq approximation, this anomaly is transported:

$$0 = D_t(\rho - \rho_b) = \mathbb{D}_t(\rho - \rho_b). \tag{10.3.19}$$

The Navier-Stokes equations coupling the Boussinesq and traditional approximations then read:

Simple Boussinesq equations under location uncertainty

Momentum equations

$$\mathbb{D}_t \boldsymbol{w} + f\boldsymbol{k} \times (\boldsymbol{u} + (\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H) = b\,\boldsymbol{k}\mathrm{d}t - \frac{1}{\rho_b}\boldsymbol{\nabla}(p'\mathrm{d}t + \mathrm{d}_t p_\sigma) + \mathcal{F}(\boldsymbol{w}\mathrm{d}t + \boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t), \tag{10.3.20a}$$

Buoyancy equation

$$\mathbb{D}_t(b + N^2 z) = 0, \tag{10.3.20b}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w} = \boldsymbol{\nabla}\cdot\big(\boldsymbol{\sigma}\dot{\boldsymbol{B}}\big) = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\cdot\boldsymbol{a} = 0. \tag{10.3.20c}$$

For this system, the thermodynamics equations are expressed through the buoyancy variable $b = -g\rho'/\rho_b$, and the stratification (Brunt-Väisälä frequency) $N^2(z) = -g\frac{1}{\rho_b}\partial_z\rho_0(z)$ is introduced. To obtained the Boussinesq model (2.3.15) of Resseguier et al. (2017a), the drift $\boldsymbol{w}$ has been assumed smooth in time (see 10.3.3 for more details).

10.3.5 Non-dimentionalized Boussinesq equations

To simplify the stochastic Boussinesq model (10.3.20), Quasi-Geostrophic (QG) models are developed for large horizontal length scales, L, such as:

$$\frac{1}{Bu} = \left(\frac{F_r}{R_o}\right)^2 = \left(\frac{L}{L_d}\right)^2 \sim 1 \text{ and } \frac{1}{R_o} = \frac{Lf_0}{U} \gg 1, \tag{10.3.21}$$

where U is the horizontal velocity scale, $L_d \triangleq \frac{Nh}{f}$ is the Rossby deformation radius and h is the characteristic vertical length scale.

That simplification necessitates a non-dimensionalistion of the Boussinesq equations under location uncertainty. To derive this non-dimensional version (10.3.20), we scale the horizontal coordinates $\tilde{\boldsymbol{x}}_h = L\boldsymbol{x}_h$, the vertical coordinate $\tilde{z} = hz$, the aspect ratio $D = h/L$ between the vertical and horizontal length scales. A characteristic time $\tilde{t} = Tt$ corresponds to the horizontal advection time U/L with horizontal velocity $\tilde{\boldsymbol{u}} = U\boldsymbol{u}$. A vertical velocity $\tilde{w} = (h/L)Uw$ is deduced from the divergence-free condition. We further take a scaled buoyancy $\tilde{b} = Bb$, pressure $\tilde{\phi}' = \Phi\phi'$ (with the density scaled pressures $\phi' = p'/\rho_b$ and $\mathrm{d}_t\phi_\sigma = \mathrm{d}_t p_\sigma/\rho_b$), and the earth rotation $\boldsymbol{f}^* = f\boldsymbol{k}$.

Hereafter, we explicit scaling assumptions to derive the non-dimensional version of the stochastic Boussinesq model.

Besides traditional ones, another dimensionless number, Υ, is introduced to compares horizontal advective and diffusive terms in the momentum and buoyancy equations. . In the following, $\boldsymbol{\sigma}_{H\bullet}$ stands for the horizontal component of $\boldsymbol{\sigma}$, $\boldsymbol{a}_H$ for $\boldsymbol{\sigma}_{H\bullet}\boldsymbol{\sigma}_{H\bullet}^T$ and A_u for its scaling. The new dimensionless number is defined by:

$$\Upsilon \triangleq \frac{UL}{A_u} = \frac{U^2}{A_u/T}. \tag{10.3.22}$$

In order to keep plausible geophysical orders of magnitude, we restrict $\Upsilon \geqslant R_o$.

The scaling to compare vertical to horizontal unresolved velocities is also considered:

$$\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim \frac{R_o}{B_u} D, \tag{10.3.23}$$

where $D = \frac{h}{L}$ is the aspect ratio and the subscript H indicates horizontal coordinates. For the justification of this scaling, we refer the reader to 3.3.1.

At mid-latitudes, the related term, given by $\beta \triangleq \partial_y f$, is much smaller than the constant part of the Coriolis frequency. Nevertheless, it can govern a large part of the relative vorticity at large scales. The following scaling is thus chosen (Vallis, 2006):

$$\beta y \sim \boldsymbol{\nabla}^{\perp} \cdot \boldsymbol{u} \sim \frac{U}{L} = R_o f_0. \tag{10.3.24}$$

The resulting non-dimensional Boussinesq system under location uncertainty becomes:

Nondimensional Boussinesq equations under location uncertainty

Momentum equations

$$\begin{aligned}
&\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)\boldsymbol{u} + \left(\frac{R_o}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z\boldsymbol{u} \\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\Big(a_{ij}\boldsymbol{u}\Big)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right) + \frac{1}{R_o}\left(1+R_o\beta y\right)\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right) \\
&\quad = -E_u\,\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned} \tag{10.3.25a}$$

$$\begin{aligned}
&\mathrm{d}_t w + (\boldsymbol{w}\cdot\boldsymbol{\nabla})\,w\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}_H\right)w + \left(\frac{R_o}{B_u\Upsilon^{1/2}}\right)(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\partial_z w \\
&\quad - \frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\Big(a_{ij}w\Big)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right) = \frac{\Gamma}{D^2}b\mathrm{d}t - \frac{E_u}{D^2}\,\partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),
\end{aligned} \tag{10.3.25b}$$

Buoyancy equation

$$\begin{aligned}
&\mathrm{d}_t b + \left(\boldsymbol{w}^*_\Upsilon\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)\right)\cdot\boldsymbol{\nabla}b - \frac{1}{2}\frac{1}{\Upsilon}\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{a}_H\boldsymbol{\nabla}b\right)\mathrm{d}t + O\left(\frac{R_o}{\Upsilon B_u}\right) \\
&\qquad + \frac{1}{(Fr)^2}\frac{1}{\Gamma}\left(w^*_{\Upsilon/2}\mathrm{d}t + \left(\frac{R_o}{B_u}\right)\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z\right) = 0,
\end{aligned} \tag{10.3.25c}$$

Effective drift

$$\begin{aligned}
&\boldsymbol{w}^*_\Upsilon = \left(\boldsymbol{u}^*_\Upsilon, w^*_\Upsilon\right)^T, \\
&\quad = \left(\left(\boldsymbol{w} - \frac{1}{2\Upsilon}\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right), \left(w - \left(\frac{R_o}{2\Upsilon B_u}\right)\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_{Hz} + O\left(\frac{R_o}{\Upsilon B_u}\right)^2\right)\right)^T,
\end{aligned} \tag{10.3.25d}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{w} = 0, \tag{10.3.25e}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\right) = 0, \tag{10.3.25f}$$

$$\boldsymbol{\nabla}_H\cdot\left(\boldsymbol{\nabla}_H\cdot\boldsymbol{a}_H\right)^T + 2\frac{R_o}{B_u}\boldsymbol{\nabla}_H\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{R_o}{B_u}\right)^2\right) = 0. \tag{10.3.25g}$$

Here, we do not separate the time-correlated components and the time-uncorrelated components in the momentum equations. The terms in $O\left(\frac{R_o}{B_u}\right)$ and $O\left(\frac{R_o}{B_u}\right)^2$ are related to the time-uncorrelated vertical velocity. These terms are too small to appear in the final QG models ($B_u = O(1)$ in QG approximation) and not explicitly shown. We only make appear the big

O approximations. Traditional non-dimensional numbers are introduced : the Rossby number $R_o = U/(f_0 L)$ with f_0 the average Coriolis frequency; the Froude number ($F_r = U/(Nh)$), ratio between the advective time to the buoyancy time; E_u, the Euler number, ratio between the pressure force and the inertial forces, $\Gamma = Bh/U^2 = D^2BT/W$ the ratio between the mean potential energy to the mean kinetic energy. To scale the buoyancy equation, the ratio between the buoyancy advection and the stratification term has also been introduced:

$$\frac{B/T}{N^2W} = \frac{B}{N^2h} = \frac{U^2}{N^2h^2}\frac{Bh}{U^2} = Fr^2\Gamma. \tag{10.3.26}$$

10.3.6 Quasi-Geostrophic model under Moderate Uncertainty (QG_{MU})

Hereafter, we consider the QG approximation ($R_o \ll 1$ and $B_u \sim 1$), for $\Upsilon \sim 1$. We focus on solutions of the Boussinesq model with Rossby number going to zero. To derive the evolution equations corresponding to this limit, the solution of the non-dimentional Boussinesq model (10.3.25) is developed as a power series of the Rossby number:

$$\begin{pmatrix} \boldsymbol{w} \\ b \\ \phi \end{pmatrix} = \sum_{k=0}^{\infty} {R_o}^k \begin{pmatrix} \boldsymbol{w}_k \\ b_k \\ \phi_k \end{pmatrix}. \tag{10.3.27}$$

According to the horizontal momentum equation (10.3.25a), the scaling of the pressure still corresponds to the usual geostrophic balance. This sets the Euler number as:

$$E_u \sim \frac{1}{R_o}. \tag{10.3.28}$$

For the ocean, the aspect ratio, $D \triangleq H/L$, is small and $D^2 \ll 1$. As a consequence,

$$\frac{D^2}{E_u} \sim D^2 R_o \ll D^2 \ll 1 \text{ and } \frac{D^2}{E_u\Upsilon} \sim \frac{D^2 R_o}{\Upsilon} \leqslant D^2 \ll 1. \tag{10.3.29}$$

Therefore, the inertial and diffusion terms are negligible in the vertical momentum equation. The hydrostatic assumption is still valid. This leads to the classical QG scaling of the buoyancy equation:

$$\Gamma \sim E_u \sim \frac{1}{R_o} \text{ and } \frac{1}{{F_r}^2\Gamma} \sim \frac{R_o}{{F_r}^2} = \frac{B_u}{R_o}. \tag{10.3.30}$$

In the following, the subscript H is omitted for the differential operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ. They all represent 2D operators. Only keeping terms of order 0 and 1, we get the following system:

Momentum equations

$$R_o\left(\mathrm{d}_t\boldsymbol{u}+(\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{u}-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\big(a_{ij}\boldsymbol{u}\big)\mathrm{d}t+O\left(\frac{R_o}{\Upsilon B_u}\right)\right)$$
$$+(1+R_o\beta y)\,\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right)=-\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),\tag{10.3.31}$$

$$b\,\mathrm{d}t+O\left(R_oD^2\right)=\partial_z\left(\phi'\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right),\tag{10.3.32}$$

Buoyancy equation

$$\frac{R_o}{B_u}\left(\mathrm{d}_tb+\boldsymbol{\nabla}b\cdot\left(\boldsymbol{u}\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)_H\right)+\partial_zb\;w\mathrm{d}t-\frac{1}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}b\right)\mathrm{d}t\right)$$
$$+w\mathrm{d}t-\frac{1}{\Upsilon}\frac{R_o}{B_u}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t+\frac{R_o}{B_u}\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z+O\left(\frac{{R_o}^2}{\Upsilon{B_u}^2}\right)=0,\tag{10.3.33}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}+\partial_zw=0,\tag{10.3.34}$$

$$\boldsymbol{\nabla}\cdot\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_H+\frac{R_o}{B_u}\partial_z\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_z=0,\tag{10.3.35}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T+2\frac{R_o}{B_u}\boldsymbol{\nabla}\cdot\partial_z\boldsymbol{a}_{Hz}+O\left(\left(\frac{R_o}{B_u}\right)^2\right)=0.\tag{10.3.36}$$

The thermodynamic equation (10.3.33) at 0 order leads to :

$$w_0=0,\tag{10.3.37}$$

and then, by the large-scale incompressibility equation (10.3.34), the 0-order horizontal velocity is divergence-free. Following the scaling assumption, the horizontal small-scale velocity is also divergence-free (10.3.35). The horizontal momentum equation (10.3.31) at the 0-th order leads to:

$$\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp\phi_0'\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\boldsymbol{\nabla}^\perp\mathrm{d}_t\phi_\sigma,\tag{10.3.38}$$

where time-correlated and time-uncorrelated components have been separated by the mean of uniqueness of the semi-martingale decomposition (Kunita, 1997). Being divergent-free, both components can be expressed with two stream functions ψ_0 and $\mathrm{d}_t\psi_\sigma$:

$$\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp\psi_0\text{ and }(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=\boldsymbol{\nabla}^\perp\mathrm{d}_t\psi_\sigma,\tag{10.3.39}$$

exactly corresponding to the dimensionless pressure terms:

$$\psi_0=\phi_0'\text{ and }\mathrm{d}_t\psi_\sigma=\mathrm{d}_t\phi_\sigma.\tag{10.3.40}$$

Deriving these equations along z and introducing the hydrostatic equilibrium (10.3.32) – decomposed between correlated and uncorrelated components – yields the classical thermal wind balance at large-scale for the 0-th order terms. The buoyancy variable does not involve any white noise term, and the small-scale random velocity is thus almost constant along z, as

$$\partial_z\boldsymbol{u}_0=\boldsymbol{\nabla}^\perp b_0\text{ and }\partial_z(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H=O\left(R_oD^2\right).\tag{10.3.41}$$

Accordingly the variance tensor scales as:

$$\forall i,j\in H,\;\partial_za_{ij}=O\left({R_o}^2D^4\right),\tag{10.3.42}$$

which is negligible in all equations, and the uncertain random field solely depends on the horizontal coordinates. Since $R_o/B_u\sim R_o$, the 1-st order term of the buoyancy equation must be kept to describe the evolution of b_0:

$$\frac{1}{B_u}\mathbb{D}^H_{0t}b_0+w_1\mathrm{d}t-\frac{1}{\Upsilon}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t+\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z=0,\tag{10.3.43}$$

where, for all functions h,

$$\mathbb{D}_{0t}^{H} h \triangleq \mathrm{d}_t h + \boldsymbol{\nabla} h \cdot \left(\boldsymbol{u}_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial_{ij}^2 \left(a_{ij} h \right) \mathrm{d}t. \tag{10.3.44}$$

Taking the derivative along z leads to:

$$\frac{1}{B_u} \left(\mathbb{D}_{0t}^{H} \partial_z b_0 + \boldsymbol{\nabla} b_0 \cdot \partial_z \left(\boldsymbol{u}_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} d\boldsymbol{B}_t)_H \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial_{ij}^2 \left(\partial_z a_{ij} b_0 \right) \mathrm{d}t \right)$$
$$+ \partial_z w_1 \mathrm{d}t - \frac{1}{\Upsilon} \left(\boldsymbol{\nabla} \cdot \partial_z a_{Hz} \right)^T \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} \partial_z (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_z = 0. \tag{10.3.45}$$

The introduction of the thermal wind equations (10.3.41) and incompressibility conditions (10.3.34-10.3.36) helps simplifying this equation as:

$$\frac{1}{B_u} \mathbb{D}_{0t}^{H} \partial_z b_0 - \boldsymbol{\nabla} \cdot \boldsymbol{u}_1 \mathrm{d}t + \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{\Upsilon^{1/2}} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H = 0. \tag{10.3.46}$$

Note the factor $\left(\frac{R_o}{B_u} \right)^{-1}$ appears. It comes from the incompressible conditions (10.3.35) and (10.3.36), leading $\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$ to both scale as $\frac{R_o}{B_u}$. The hydrostatic balance at 0-order links the buoyancy to the pressure, and then to the stream function

$$\partial_z b_0 = \partial_z^2 \phi_0 = \partial_z^2 \psi_0. \tag{10.3.47}$$

The 1-st order term of the vertical velocity is not known. Yet, the system can be closed using the vorticity equation at order 1:

$$\boldsymbol{\nabla}^{\perp} \cdot \left(\mathbb{D}_{0t}^{H} \boldsymbol{u}_0 \right) + \left(\boldsymbol{\nabla} \cdot \boldsymbol{u}_1 + \left(\frac{R_o}{B_u} \right)^{-1} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) + \boldsymbol{\nabla} (\beta y) \cdot \left(\boldsymbol{u}_0 \mathrm{d}t + (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H \right) = 0, \tag{10.3.48}$$

where the divergence terms come from the constant Coriolis term.

Again, factors $\left(\frac{R_o}{B_u} \right)^{-1}$ compensate the order of magnitude of $\boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H$ and $\boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$. Then,

$$\mathbb{D}_{0t}^{H} \left(\Delta \psi_0 \right) + \boldsymbol{\nabla} \cdot u_1 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} \left(\frac{R_o}{B_u} \right)^{-1} \boldsymbol{\nabla} \cdot (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H + \beta \left(v_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_y \right)$$
$$+ \frac{1}{\Upsilon^{1/2}} \mathrm{tr} \left(\boldsymbol{\nabla}^{\perp} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H^T \boldsymbol{\nabla} \boldsymbol{u}_0^T \right) - \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial_{ij}^2 \left(\boldsymbol{\nabla}^{\perp} a_{ij} \cdot \boldsymbol{u}_0 \right) \mathrm{d}t = 0. \tag{10.3.49}$$

To make appear the transport of PV, we note that:

$$\begin{aligned} \mathbb{D}_{0t}^{H} \left(1 + \beta y \right) &= -\beta \left(v_0 \mathrm{d}t + \frac{1}{\Upsilon^{1/2}} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_y \right) + \boldsymbol{\nabla} \cdot a_{Hy} \beta \mathrm{d}t \\ &\quad - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{2\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t. \end{aligned} \tag{10.3.50}$$

Then, using (10.3.46), (10.3.47) and (10.3.49), we get:

$$\mathbb{D}_{0t}^{H} \left(\Delta \psi_0 + 1 + \beta y + \frac{1}{B_u} \partial_z^2 \psi_0 \right) = -\boldsymbol{\nabla} \cdot a_{Hy} \beta \mathrm{d}t - \left(\frac{R_o}{B_u} \right)^{-1} \frac{1}{2\Upsilon} \boldsymbol{\nabla} \cdot \left(\boldsymbol{\nabla} \cdot \boldsymbol{a}_H \right)^T \mathrm{d}t$$
$$- \mathrm{tr} \left(\boldsymbol{\nabla}^{\perp} (\boldsymbol{\sigma} \mathrm{d}\boldsymbol{B}_t)_H^T \boldsymbol{\nabla} \boldsymbol{u}_0^T \right) + \frac{1}{2\Upsilon} \sum_{i,j \in H} \partial_{ij}^2 \left(\boldsymbol{\nabla}^{\perp} a_{ij} \cdot \boldsymbol{u}_0 \right) \mathrm{d}t. \tag{10.3.51}$$

We recall that coefficients $\left(\frac{Ro}{Bu}\right)^{-1}$ are still present since

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H \sim \boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H)^T\,\mathrm{d}t \sim \left(\frac{Ro}{Bu}\right). \tag{10.3.52}$$

If we rewrite the equation with dimensional quantities, the evolution equation for $u_0 = \lim_{Ro\to 0} u$ is obtained (dropping the index 0 for clarity):

$$\mathbb{D}_t^H Q = -\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right) + \frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_H f))^T\,\mathrm{d}t, \tag{10.3.53}$$

where Q is the QG potential vorticity:

$$Q \triangleq \Delta\psi + f + \left(\frac{1}{N}\right)^2\partial_z^2\psi. \tag{10.3.54}$$

For a given $\boldsymbol{\sigma}_H$, equations (10.3.53) and (10.3.54) constitute a stochastic Eulerian QG model that can be simulated numerically.

Note, (10.3.38) provides the geostrophic balance for the small-scale velocity component. To express the material derivative of Q, the noise term is expanded:

$$-\mathrm{tr}\left(\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H^T\boldsymbol{\nabla}\boldsymbol{u}^T\right) = -\sum_{k,j\in H}\partial_{kj}^2\psi\partial_k\boldsymbol{\sigma}_{j\bullet}\mathrm{d}\boldsymbol{B}_t. \tag{10.3.55}$$

According to (2.2.10), the difference between the material derivative, D_tQ, and the stochastic transport operator $\mathbb{D}_tQ$, is a function of the time-uncorrelated forcing:

$$\left\{\begin{array}{lcl}\mathbb{D}_tQ &=& f_1\mathrm{d}t + \boldsymbol{h}_1^T\mathrm{d}\boldsymbol{B}_t,\\ D_tQ &=& f_2\mathrm{d}t + \boldsymbol{h}_2^T\mathrm{d}\boldsymbol{B}_t,\end{array}\right. \Longleftrightarrow \left\{\begin{array}{lcl} f_2 &=& f_1 + \mathrm{tr}\big((\boldsymbol{\sigma}^T\boldsymbol{\nabla})\,\boldsymbol{h}_1^T\big),\\ \boldsymbol{h}_1 &=& \boldsymbol{h}_2.\end{array}\right. \tag{10.3.56}$$

The expression of h_1 is given by equation (10.3.53) and the above formulas give:

$$D_tQ - \mathbb{D}_tQ = \sum_{i\in H}\boldsymbol{\sigma}_{i\bullet}\partial_i\left(-\sum_{j,k\in H}\partial_k\boldsymbol{\sigma}_{j\bullet}\partial_{kj}^2\psi\right)^T, \tag{10.3.57}$$

$$= -\sum_{i,j,k\in H}\left(\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_{kj}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ijk}^3\psi\right). \tag{10.3.58}$$

With the use of the small-scale incompressibility, we obtain:

$$\frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ik}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_{jk}^2\psi + \boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ijk}^3\psi\right). \tag{10.3.59}$$

From (10.3.58) and (10.3.59), it yields:

$$D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H}\partial_{ij}^2\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial_{ik}^2\boldsymbol{\sigma}_{j\bullet}^T\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}_{j\bullet}^T\partial_{ik}^2\psi\right). \tag{10.3.60}$$

Denoting, α, the following matrix

$$\alpha_{ij} \triangleq \sum_{k\in H}\partial_k\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}_{k\bullet}^T = \sum_{k\in H}\partial_k(\boldsymbol{\sigma}_{i\bullet}\partial_j\boldsymbol{\sigma}_{k\bullet}^T), \tag{10.3.61}$$

we have

$$\boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi) = \sum_{i,j,k\in H}\left(\partial_j\boldsymbol{\sigma}_{i\bullet}\partial^2_{ik}\boldsymbol{\sigma}^T_{j\bullet}\partial_k\psi + \partial_j\boldsymbol{\sigma}_{i\bullet}\partial_k\boldsymbol{\sigma}^T_{j\bullet}\partial^2_{ik}\psi\right), \tag{10.3.62}$$

$$= D_tQ - \left(\mathbb{D}_tQ - \frac{1}{2}\sum_{i,j\in H}\partial^2_{ij}\left(\boldsymbol{\nabla}^\perp a_{ij}\cdot\boldsymbol{u}\right)\right), \tag{10.3.63}$$

and the material derivative of the PV finally reads:

$$D^H_tQ = \boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\psi)\mathrm{d}t - \frac{1}{2}\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot(\boldsymbol{a}_Hf))^T\,\mathrm{d}t - \mathrm{tr}\left[\boldsymbol{\nabla}^\perp(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)^T_H\boldsymbol{\nabla}\boldsymbol{u}^T\right]. \tag{10.3.64}$$

To note, the transpose of the matrix α has a compact expression:

$$\alpha^T = \sum_p\left(\boldsymbol{\nabla}\boldsymbol{\sigma}^T_{Hp}\right)^2. \tag{10.3.65}$$

10.3.7 Surface Quasi-Geostrophic model under Moderate Uncertainty (SQG_{MU})

A classical boundary conditions choice for the QG model considers a vanishing solution in the deep ocean and a buoyancy transport (10.3.20b) at the surface (Vallis, 2006; Lapeyre and Klein, 2006):

$$\psi \underset{z\to-\infty}{\longrightarrow} 0 \text{ and } D^H_tb_{|z=0} = \mathbb{D}^H_tb_{|z=0} = 0. \tag{10.3.66}$$

Assuming zero PV in the interior but keeping these boundary conditions leads to the Surface Quasi-Geostrophic model (SQG) (Blumen, 1978; Held et al., 1995; Lapeyre and Klein, 2006; Constantin et al., 1994, 1999, 2012). Under the stochastic framework, the derivation is similar. The PV is indeed identical to the classical one (see equation (10.3.54)), assuming zero PV in the interior and vanishing solution as $z\to-\infty$ unsurprisingly yields the same SQG relationship:

$$\hat{b} = N\|\boldsymbol{k}\|\ \hat{\psi}. \tag{10.3.67}$$

The top boundary condition, equation (10.3.66), provides an evolution equation, namely the horizontal transport of surface buoyancy, in the stochastic sense:

$$\mathbb{D}^H_tb = 0. \tag{10.3.68}$$

The time-uncorrelated component of the velocity, $\boldsymbol{\sigma}\dot{\boldsymbol{B}}$, is divergence-free. Its inhomogeneous and anisotropic spatial covariance has then to be specified. The time-correlated component of the velocity is also divergence-free, with a stream function specified by the SQG relation (10.3.67). The buoyancy is randomly advected, and the resulting smooth velocity component is random as well.

10.3.8 QG and SQG models under Strong Uncertainty (SQG_{SU})

For the case Υ close to the Rossby number, the diffusion term is not negligible anymore and the geostrophic balance is modified. As the terms of the geostrophic balance remain large ($R_o\leqslant\Upsilon\ll 1$), the scaling of the pressure can still be done with the Coriolis force. This leads to an Euler number scaling as

$$E_u\sim\frac{1}{R_o}. \tag{10.3.69}$$

Keeping a small aspect ratio $D^2\ll 1$, we get

$$\frac{E_u}{D^2}\sim\frac{1}{R_oD^2}\gg\frac{1}{R_o}\geqslant\frac{1}{\Upsilon}. \tag{10.3.70}$$

As the Rossby number and the ratio Υ are both small in the vertical momentum equation, the inertial terms are dominated by the diffusion term which is itself negligible in front of the pressure term. The hydrostatic balance is hence conserved. The buoyancy scaling still correspond to the thermal winds relation:

$$\Gamma \sim E_u \sim \frac{1}{R_o}. \tag{10.3.71}$$

Considering the scaling $\frac{(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z}{\|(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_H\|} \sim D\frac{R_o}{B_u}$ for the vertical small-scale velocity, the non-dimensional evolution equations are now given by:

Momentum equations

$$R_o\left(\mathrm{d}_t\boldsymbol{u} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\left(\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\cdot\boldsymbol{\nabla}\right)\boldsymbol{u} + O\left(\frac{R_o}{\Upsilon B_u}\right)\right) - \frac{R_o}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\big(a_{ij}\boldsymbol{u}\big)\mathrm{d}t$$
$$+ (1+R_o\beta y)\,\boldsymbol{k}\times\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\boldsymbol{\sigma}_H\mathrm{d}\boldsymbol{B}_t\right) = -\boldsymbol{\nabla}_H\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{10.3.72}$$

$$b\,\mathrm{d}t + O\left(\frac{R_oD^2}{\Upsilon^{1/2}}\right) = \partial_z\left(\phi'\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}\mathrm{d}_t\phi_\sigma\right), \tag{10.3.73}$$

Buoyancy equation

$$\frac{R_o}{B_u}\left(\mathrm{d}_tb + \boldsymbol{\nabla}b\cdot\left(\boldsymbol{u}\mathrm{d}t + \frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}d\boldsymbol{B}_t)_H\right) + \partial_zb\;w\mathrm{d}t\right) - \frac{R_o}{2\Upsilon}\sum_{i,j\in H}\partial^2_{ij}\left(a_{ij}b\right)\mathrm{d}t$$
$$+ w\mathrm{d}t - \frac{1}{\Upsilon}\frac{R_o}{B_u}\left(\boldsymbol{\nabla}\cdot a_{Hz}\right)^T\mathrm{d}t + \frac{R_o}{B_u}\frac{1}{\Upsilon^{1/2}}(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t)_z + O\left(\frac{{R_o}^2}{\Upsilon {B_u}^2}\right) = 0, \tag{10.3.74}$$

Incompressibility

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} + \partial_zw = 0, \tag{10.3.75}$$

$$\boldsymbol{\nabla}\cdot\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_H + \frac{R_o}{B_u}\partial_z\big(\boldsymbol{\sigma}\mathrm{d}\boldsymbol{B}_t\big)_z = 0, \tag{10.3.76}$$

$$\boldsymbol{\nabla}\cdot\left(\boldsymbol{\nabla}\cdot\boldsymbol{a}_H\right)^T + 2\frac{R_o}{B_u}\boldsymbol{\nabla}\cdot\partial_z\boldsymbol{a}_{Hz} + O\left(\left(\frac{R_o}{B_u}\right)^2\right) = 0. \tag{10.3.77}$$

Again, the operators Del, $\boldsymbol{\nabla}$, and Laplacian, Δ represent 2D operators. If $R_o \sim \Upsilon$, the system is not anymore approximately in geostrophic balance. The large-scale velocity becomes divergent and decoupling the system is more involved. For sake of simplicity, we thus focus on the case of homogeneous and horizontally isotropic turbulence. As a consequence, the variance tensor a is constant in space and diagonal:

$$a = \begin{pmatrix} a_h & 0 & 0 \\ 0 & a_h & 0 \\ 0 & 0 & a_z \end{pmatrix}. \tag{10.3.78}$$

The time-correlated components of the horizontal momentum (10.3.72) at the 0-th order can be written as:

$$-\frac{a_H}{2}\Delta\boldsymbol{u}_0 + \boldsymbol{k}\times\boldsymbol{u}_0 = -\boldsymbol{\nabla}\phi'_0. \tag{10.3.79}$$

Let us note that $\boldsymbol{f}\times\boldsymbol{u}_0 = f\boldsymbol{J}\boldsymbol{u}_0$ with $\boldsymbol{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and that $\boldsymbol{J}^T = \boldsymbol{J}^{-1} = -\boldsymbol{J}$. For a constant Coriolis frequency, the previous equation can be solved in the horizontal Fourier space :

$$\hat{\boldsymbol{u}}_0 = \left(\boldsymbol{J} + \frac{a_H}{2}\|\boldsymbol{k}\|_2^2\mathbb{I}_d\right)^{-1}\left(-i\boldsymbol{k}\hat{\phi}'_0\right) = \left(\mathbb{I}_d - \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2\boldsymbol{J}\right)^{-1}\left(i\boldsymbol{k}^\perp\hat{\phi}'_0\right), \tag{10.3.80}$$

with $k_c = \sqrt{\frac{2}{a_H}}$. Expanding the right-hand side operator in Taylor series and using the properties $\boldsymbol{J}^{2p} = (-1)^p \mathbb{I}_d$ and $\boldsymbol{J}^{2p+1} = (-1)^p \boldsymbol{J}$,

$$\left(\mathbb{I}_d - \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right)^{-1} = \sum_{p=0}^{+\infty} \left(\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right)^p, \tag{10.3.81}$$

$$= \sum_{p=0}^{+\infty} (-1)^p \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^{4p} \mathbb{I}_d + \sum_{p=0}^{+\infty} (-1)^p \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^{4p+2} \boldsymbol{J}, \tag{10.3.82}$$

$$= \sum_{p=0}^{+\infty} \left(-\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4\right)^p \left(\mathbb{I}_d + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right), \tag{10.3.83}$$

$$= \frac{1}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(\mathbb{I}_d + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2 \boldsymbol{J}\right). \tag{10.3.84}$$

This leads to the following solution for the modified geostrophic balance:

$$\hat{\boldsymbol{u}}_0 = \frac{1}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(i\boldsymbol{k}^\perp \hat{\phi}_0'\right) + \frac{\left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^2}{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4} \left(-i\boldsymbol{k}\hat{\phi}_0'\right). \tag{10.3.85}$$

In the physical space, the solution reads:

$$\boldsymbol{u}_0 = \boldsymbol{\nabla}^\perp \underbrace{\left(1 + \frac{\Delta^2}{k_c^4}\right)^{-1} \phi_0'}_{=\psi_0} + \boldsymbol{\nabla} \underbrace{\left(1 + \frac{\Delta^2}{k_c^4}\right)^{-1} \frac{\Delta}{k_c^2} \phi_0'}_{=\tilde{\psi}_0} \text{ with } k_c = \sqrt{\frac{2}{a_H}} \tag{10.3.86}$$

which is the Helmholtz decomposition of the horizontal velocity u_0 into its rotational and divergent component with a stream function ψ_0 and a velocity potential $\tilde{\psi}_0$. Differentiating the buoyancy equation at the order 0 along z, we obtain

$$\frac{a_H}{2} \Delta \partial_z \left(\frac{b_0}{B_u}\right) = \partial_z w_0 = -\boldsymbol{\nabla} \cdot \boldsymbol{u}_0 = -\Delta \tilde{\psi}_0 = -\frac{\Delta^2}{k_c^4} \psi_0. \tag{10.3.87}$$

The time-correlated part of the 0-th order hydrostatic equation relates the buoyancy to the pressure ϕ_0':

$$\frac{a_H}{2} \Delta \partial_z \left(\frac{b_0}{B_u}\right) = \frac{a_H}{2} \Delta \partial_z^2 \phi_0' = \frac{a_H}{2} \Delta \partial_z^2 \left(1 + \frac{\Delta^2}{k_c^4}\right) \psi_0. \tag{10.3.88}$$

Gathering these two equations and dropping index 0 lead to:

$$\left(\Delta + \left(1 + \frac{\Delta^2}{k_c^4}\right) \partial_z \left(\left(\frac{f_0}{N}\right)^2 \partial_z\right)\right) \psi = 0, \tag{10.3.89}$$

with $k_c = \sqrt{\frac{2 f_0}{a_H}}$. Using the horizontal Fourier transform, it writes:

$$\left(-\|\boldsymbol{k}\|_2^2 + \left(1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4\right) \partial_z \left(\left(\frac{f_0}{N}\right)^2 \partial_z\right)\right) \hat{\psi} = 0. \tag{10.3.90}$$

Under an uniform stratification, with a fixed value at a specific depth ($z = \eta$), and a vanishing condition in the deep ocean ($z \to -\infty$), a solution is:

$$\hat{\psi}(\boldsymbol{k}, z) = \hat{\psi}(\boldsymbol{k}, \eta) \; exp\left(\frac{N\|\boldsymbol{k}\|_2}{f_0 \sqrt{1 + \left\|\frac{\boldsymbol{k}}{k_c}\right\|_2^4}} (z - \eta)\right). \tag{10.3.91}$$

Accordingly, the buoyancy is:

$$\hat{b} = \partial_z \hat{\phi}' = f_0 \left(1 + \left\| \frac{\boldsymbol{k}}{k_c} \right\|_2^4 \right) \partial_z \hat{\psi} = N \|\boldsymbol{k}\|_2 \sqrt{1 + \left\| \frac{\boldsymbol{k}}{k_c} \right\|_2^4}\, \hat{\psi}. \tag{10.3.92}$$

Conclusion

During this thesis, we have studied the dynamics of tracers in fluid flows changing either very fast or very slowly.

From a large scale point of view, fast velocity changes prevent a full tracking of the dynamics. A fully deterministic description cannot completely apprehend all the complexity generated by the nonlinear interactions between scales and the necessary physical approximations performed in terms of forcing or boundary conditions. In contrast, a stochastic physical framework can encode this partial *uncertainty* through the introduction of random fields. The difficulty consists in keeping important properties related for instance to energy transfer or transport relations. This challenge is addressed by the modeling under location uncertainty. To model very fast changes of an incompressible flow – or similarly a very large time scale of observation – the unresolved and highly-oscillating component of the velocity is assumed to be random and time-uncorrelated. As demonstrated in chapters 2, 3, 4 and 8, within such framework, general and simplified random fluid dynamics models can be inferred from fundamental conservations of mechanics: mass, momentum, energy, amount of substance, etc. Stochastic versions of Navier-Stokes, Boussinesq, quasi-geostrophic and surface quasi-geostrophic models have been rigorously derived. In these evolution laws, a tracer is transported in a stochastic sense. Accordingly, the tracer energy is redistributed by an effective drift and through the joint action of two other terms.

The effective drift may be understood as the difference between the Lagrangian and the Eulerian mean velocity. It is directed towards the minimum of turbulence and have many similarities with the Stokes drift and the so-called bolus velocity introduced in oceanic models. The two other terms express a diffusion contribution and a random forcing. The diffusion is possibly inhomogeneous and anisotropic; it stabilizes the unstable tracer modes through an energy dissipation. Chapters 6 and 7 have shown that these first terms enable us to derive stable deterministic POD reduced order models from a direct Galerkin projection of fluid dynamical equations even at relatively high Reynolds number. Adapted estimators have been proposed for this purpose. The physical effect of the residual velocity forgotten by the ROM have also been interpreted through the eyes of our models under location uncertainty. Accordingly, these models offer new data analysis tools complementary to the usual Empirical Orthogonal Functions (EOF). Furthermore, the Surface Quasi-Geostrophic model under Strong Uncertainty (SQG_{SU}) model of chapter 4 exhibited the importance of subgrid dissipation terms in physical model derivation. This model also revealed how the interactions between this diffusion and the earth rotation affect the mesoscale and submesocale dynamics of the ocean.

The random forcing is a multiplicative noise which destabilizes some tracer modes by injecting as much energy as the diffusion removes. It acts as a random energy bridge between different modes of the tracer. These energy transfers are recovered but are *uncertain*. The ensuing instabilities hence increase the tracer variance. These stochastic instabilities are at the heart of chapters 2, 3, 5 and 8. First, chapters 2 and 8 theoretically unveil these energy fluxes. Then, through numerical simulations of a simplified oceanic model, chapter 3 demonstrates that a single run of the model resolves physically relevant filament instabilities unlike the deterministic model simulated at the same resolution. These instabilities are triggered by the multiplicative noise. So a single realization of the model under location uncertainty can improve numerical simulations of fluid flows. Furthermore, in the context of forecasting, the random destabilization prevents an ensemble of simulations to align with unstable modes of the dynamic. In geophysical fluid dynamics, this well-known problem of degeneracy is due to a strong stabilization of some modes by the subgrid tensor. This problem has been illustrated in chapter 3. An ensemble generated by the deterministic

dynamics and random initial conditions shrunk and hence underestimated its error by an order of magnitude. In the transport under location uncertainty, the energy is drained from these unstable modes to stable ones. The stable modes remain energetic due to the random energy intake and prevent the ensemble degeneracy. The ensemble simulated according to our model accurately estimates the amplitude and the position of its own error in both spatial and spectral space. Chapter 8 also shows that a small ensemble evolving in a strongly reduced subspace succeeds in estimating its error. In a similar way, chapter 5 shows how the same ensemble better describes a bifurcation in a long term forecast. The bifurcation leads to two likely scenarios. Before the bifurcation, the dynamics is in a meta-stable state. Again the subgrid tensor over stabilizes some stable modes. As in chapter 3, it strongly weakens small-scale modes which encode either an accurate information on the initial condition or a spread due to random initial conditions. Moreover, this over stabilization biased the exits of the meta-stable state. In this case, these exits are mainly determined by infinitesimal numerical errors. This makes long-term forecasts inaccurate as illustrated by both deterministic simulations and ensemble generated by random initial conditions. On the other hand, the noise involved in the model under location uncertainty sufficiently destabilizes the metastable state to recover and accurately describe both scenarios with a small ensemble size. The small number of needed realizations is also due to the multiplicative structure of the noise. Indeed, in addition to the introduction of non-Gaussianity this structure enforces an inhomogeneity of the noise even though the random small-scale velocity component is homogeneous. Let us outline that for an homogeneous random velocity field, the associated turbulent diffusion is uniform but the total energy budget of the subgrid terms – diffusion plus noise – is inhomogeneous. In this model, only the physically relevant instabilities are triggered. Accordingly, each realization are physically plausible. From an UQ point of view, it means that only the over-stabilized modes are spread. As a consequence, much less directions have to be sampled and in this way only a few realizations should be needed. This constitutes a capital advantage in the perspective of geophysical flow based numerical simulation. Moreover, the model structure is independent of the noise amplitude. This should allow strong noise in order to fit constraints of realistic ensemble forecasting without strongly perturbing the physics. This is often hardly possible with empirical models which exhibit an inappropriate balance between the effective diffusion and the noise variance. Chapter 8 demonstrates the possibility of estimating at low computational cost the influence of stationary inhomogeneous and time-uncorrelated small-scale velocity in POD-ROMs derived from a model under location uncertainty. Furthermore, these stochastic POD-ROMs demonstrate several strong UQ skills. The ensemble variability of stable modes is maintained by the random energy fluxes. They are in this way sustained and do not correspond to evanescent modes as in the case of only dissipative subgrid stress models. In addition, by a skillful spreading, the ensemble remains very close to the reference. On the one hand, some realizations become closer to the reference; on the other hand, the ensemble variance increases only if necessary. The algebraical structure of the global random transport under location uncertainty has been explained in chapters 1 and 8. The semigroup of passive tracer dynamics shows some similarities with a unitary operator called free multiplicative Brownian matrix. In finite dimension, this corresponds to a unitary Brownian matrix. By construction, the solution is forced to remain in the manifold of constant energy. Nonetheless, chapter 8 has shown that a truncation in the number of modes describing a numerical solution yields a small energy leak toward unresolved modes. Yet this process seems physically relevant to us since it enables the energy to cascade toward molecular viscous scales.

Currently, we continue to analyze stochastic instabilities introduced by the modeling under location uncertainty through the work begun in chapter 8 as well as the study of Lagrangian dynamics of tracer gradients. This last study is not presented in this thesis. The study aims at better understanding and taking into account subgrid velocity influence in mixing diagnosis. For this purpose, we generalize to random flows the Eulerian mixing criterions of Okubo (1970); Weiss (1991) and Lapeyre et al. (1999). Several theoretical and numerical results have already been obtained. Depending on the value of few parameters, the time-uncorrelated velocity either promotes hyperbolicity, ellipticity or elliptic bursts. We are currently working on the application of this method to diagnose mixing in the ocean surface from altimetry-derived velocity.

A problem of these velocity fields measured from space is indeed their low spatial and temporal resolutions. This bring us to the second subject: fluid flows changing very slowly – or observed at a very small time scale. Mixing diagnoses relying on these smooth data often neglect the influence of

the unobserved or sparsely observed subgrid velocity. This sometimes leads to strong errors (Keating et al., 2011). By our stochastic Lagrangian study we aim at statistically describing the effect of this subgrid component on mixing. Conversely, studying the mixing induced by the smooth large-scale velocity measured from space can unveil the missing elements of the dynamics. These identified missing elements can then yield better subgrid parametrization as well as subgrid velocity statistics. This problem has been addressed in chapter 9. We have characterized the squared norm of tracer gradients integrated over the space after a finite-time advection. When a spatial smoothing decorrelates the tracer and the incompressible flow, advection can only strengthen the gradients and the local growth rate is independent of the tracer. For a very low temporal resolution, the main processes of gradient strengthening corresponds to velocity shears where streamlines are straight and to angular velocity shears where the streamlines are curved. This last process folds and wraps the tracer around vortices. From these two mechanisms, we derived a simplified model – linear in time – to quantify the norm of tracer gradients. After having demonstrated that this norm characterizes the correlation length of the tracer, we analyzed its influence on the tracer spectrum. All things considered, we have obtained a simple model to quantify the position of the tracer spectrum tail after a finite-time advection. Furthermore, the parameters of this model only depends on the Eulerian velocity. We presented simulations with a model of stationary flow and real satellite images of the ocean. We have successfully applied our model to the Lagrangian advection method (Sutton et al., 1994; Desprès et al., 2011a; Berti and Lapeyre, 2014; Dencausse et al., 2014) to set up the spatial filter width and the advection time. This simple model could also be valuable in numerical simulations of large-scale fluid flows to specify the turbulent diffusion as a function of the large-scale velocity.

When the models under location uncertainty rely on an inhomogeneous time-uncorrelated velocity, we have referred to the Itō drift $\boldsymbol{w}$ to as the "large-scale velocity". Yet it was a choice. Similarly to Stokes drift studies, the Itō drift, $\boldsymbol{w}$, may be considered as the mean Lagrangian velocity whereas the corrected drift (Stratonovich drift in the incompressible case), $\boldsymbol{w}^\star = \boldsymbol{w} + \frac{1}{2}\sum_{i=1}^{d}(\boldsymbol{\sigma}\partial_i\boldsymbol{\sigma}_{i\bullet}^T - \partial_i\boldsymbol{\sigma}\boldsymbol{\sigma}_{i\bullet}^T)$, may be understood as the mean Eulerian velocity. Note that for a constant density the mass conservation only imposes zero divergence for $\boldsymbol{w}^\star$ and $\boldsymbol{\sigma}$ but not for $\boldsymbol{w}$. Depending on the averaging method, further numerical studies on a simple flow generated by a known inhomogeneous stochastic differential equation may help us to know in which case we estimate $\boldsymbol{w}$ and in which case we estimate $\boldsymbol{w}^\star$. Moreover, in our different stochastic Navier-Stokes representations, we have assumed a transport of the Itō drift rather than a transport of the Stratonovich drift. Numerical procedures could be performed on simple inhomogeneous flows to rigorously test both assumptions.

The transport under location uncertainty is characterized by the stabilization and the destabilization of several phenomena. Thus, this formalism may lead to interesting dispersion equations with noises. Such equations could link the property of the operator $\boldsymbol{\sigma}$ to the stability of some waves as well as non-oscillating phenomenons such as the filament instabilities of chapter 3. Many tools used to study deterministic dynamical systems (Lyapunov exponents, stability, ellipticity, hyperbolicity, bifurcation, etc) have a stochastic version (Arnold and Kloeden, 1989) and may help handle these stochastic dispersion equations. Moreover, if the influence of a simulation resolution explicitly appears in such a dispersion relation – for instance through a hyperviscosity – we may be able to setup the parameters of $\boldsymbol{\sigma}$ as a function of the tracer and the resolution in order to trigger the right instabilities. Another way of studying the instabilities and the energy fluxes of our random models is through reduced order models. In particular, Dynamical Orthogonality (DO) with a constant or varying dimensionality is an attractive method (Sapsis, 2011; Sapsis and Lermusiaux, 2012; Sapsis and Majda, 2013b; Sapsis, 2013). First, it is adapted to random systems and naturally separates mean and random components. This could be interesting to study the random energy exchanges between the mean and random components and between the stable and unstable modes. Furthermore, it may also overcome the energy leak unveiled in chapter 8 by enabling the spatial modes to change on the fly. As suggested in chapter 1, the theories of random matrices and free probability could bring a lot of information on the eigenvalues of the dynamics of randomly transported passive tracers. Again, these eigenvalues – probably lying on the unit circle of $\mathbb{C}$ – may reveal interesting features about random energy exchanges. Potential analytical results on the statistics of the tracer after a finite time may also improve some data assimilation procedures.

Even without additional analytical results, the models under location uncertainty offer new procedures to learn from geophysical data. A natural application is filtering and this application has to be addressed. In addition, the Girsanov theorem (see chapter 1) together with a parametric model of the operator $\boldsymbol{\sigma}$ theoretically enables maximum likelihood estimation and Bayesian inference from a set of satellite images of the tracer. Another way of closing our models could rely on a physical relation. For instance, the product of the isolines curvature and the norm of the tracer gradient is conserved in many cases (Constantin et al., 1994; Lapeyre, 2000). With an adequate parametric model (see for instance chapters 3 and 5), some parameters of $\boldsymbol{\sigma}$ such as its correlation length can be locally setup in order to meet this physical constraint.

In the long term, several modifications of the models under location uncertainty could be envisaged. The assumption of conditional Gaussianity of the unresolved velocity may be a physical limitation. A work around could consist of averaging or homogenization procedures of our stochastic models. Fast modes would follow linearized models under location uncertainty. Since the evolution laws of fast modes already contain a noise, there would be no need to artificially add it before the homogenization as in the MTV framework (Majda et al., 1999, 2001; Franzke et al., 2005; Majda et al., 2008). Furthermore, Levy processes defined by noises with Laplace distributions could be used instead of continuous semimartingales defined by Gaussian noises. However, even though the stochastic calculus of Levy processes is well known we would need an analog of the Itō-Wentzell formula. Finally, without relying on jump processes, statistical physics may lead to an Itō-like formula with an infinite number of terms (Klyatskin, 2005). It could overcome the difficulty introduced by the time differentiation of white noise in the derivation of random stochastic models (see chapters 2 and 8). Notwithstanding, we would again need an analog of the Itō-Wentzell formula.

Bibliography

M. Allen and D. Stainforth. Towards objective probabilistic climate forecasting. *Nature*, 419(6903): 228–228, 2002.

J. Anderson and S. Anderson. A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, 127(12):2741–2758, 1999.

D. Andrews and M. McIntyre. Planetary waves in horizontal and vertical shear: the generalized Eliassen-Palm relation and the zonal mean acceleration. *J. Atmos. Sci.*, 33:2031–2048, 1976.

H. Aref. Stirring by chaotic advection. *Journal of fluid mechanics*, 143:1–21, 1984.

L. Arnold. *Stochastic differential equations: theory and applications*. Wiley, 1974.

L. Arnold and P. Kloeden. Lyapunov exponents and rotation number of two-dimensional systems with telegraphic noise. *SIAM Journal on Applied Mathematics*, 49(4):1242–1274, 1989.

G. Artana, A. Cammilleri, J. Carlier, and E. Mémin. Strong and weak constraint variational assimilations for reduced order fluid flow modeling. *J. Comp. Phys*, 231(8):3264–3288, 2012.

A. Aspden, N. Nikiforakis, S. Dalziel, and J. Bell. Analysis of implicit LES methods. *App. Math. Comp. Sci.*, 3(1):103–126, 2008.

N. Aubry, P. Holmes, J. Lumley, and E. Stone. The dynamics of coherent structures in the wall region of a turbulent boundary layer. *J. Fluid Mech.*, 192:115–173, 1988.

G. Badin. Surface semi-geostrophic dynamics in the ocean. *Geophysical & Astrophysical Fluid Dynamics*, 107(5):526–540, 2013.

J. Bardina, J. Ferziger, and W. Reynolds. Improved subgrid scale models for large eddy simulation. In AIAA, editor, *13th Fluid Mechanics & Plasma Dynamics Conference*, volume 80, page 1380, 1980.

S. Belan, A. Chernykh, and G. Falkovich. Phase transitions in the distribution of inelastically colliding inertial particles. *Journal of Physics A: Mathematical and Theoretical*, 49(3):035102, 2015.

P. Berloff. Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529:71–95, 2005.

P. Berloff and J. McWilliams. Material transport in oceanic gyres. part II: Hierarchy of stochastic models. *Journal of Physical Oceanography*, 32(3):797–830, 2002.

J. Berner, G. Shutts, M. Leutbecher, and T. Palmer. A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences*, 66(3):603–626, 2009.

J. Berner, S.-Y. Ha, J. Hacker, A. Fournier, and C. Snyder. Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Monthly Weather Review*, 139(6):1972–1995, 2011.

J. Berner, U. Achatz, L. Batte, A. De La Camara, D. Crommelin, H. Christensen, M. Colangeli, S. Dolaptchiev, C. Franzke, P. Friederichs, P. Imkeller, H. Jarvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. Palmer, C. Penland, J.-S. Von Storch, M. Sakradzija, M. Weniger, A. Weisheimer, P. Williams, and J.-I. Yano. Stochastic parameterization: towards a new view of weather and climate models. Technical report, arXiv:1510.08682 [physics.ao-ph], 2015.

S. Berti and G. Lapeyre. Lagrangian reconstructions of temperature and velocity in a model of surface ocean turbulence. *Ocean Modelling*, 76:59–71, 2014.

P. Biane. Free brownian motion, free stochastic calculus and random matrices. *Free probability theory (Waterloo, ON, 1995)*, 12:1–19, 1997.

W. Blumen. Uniform potential vorticity flow: part I. theory of wave interactions and two-dimensional turbulence. *Journal of the Atmospheric Sciences*, 35(5):774–783, 1978.

G. Boccaletti, R. Ferrari, and B. Fox-Kemper. Mixed layer instabilities and restratification. *Journal of Physical Oceanography*, 37(9):2228–2250, 2007.

M. Bocquet, P. Raanes, and A. Hannart. Expanding the validity of the ensemble Kalman filter without the intrinsic need for inflation. *Nonlin. Processes Geophys*, 22:645–662, 2015.

M. Bocquet, K. Gurumoorthy, A. Apte, A. Carrassi, C. Grudzien, and C. Jones. Degenerate Kalman filter error covariances and their convergence onto the unstable subspace. *arXiv preprint arXiv:1604.02578*, 2016.

J. Boris, F. Grinstein, E. Oran, and R. Kolbe. New insights into large-eddy simulation. *Fluid Dynamics Research*, 10:199–228, 1992.

J. Boussinesq. Essai sur la théorie des eaux courantes. Mémoires présentés par divers savants à l'Académie des Sciences, 23 (1): 1–680, 1877.

J. Brankart. Impact of uncertainties in the horizontal density gradient upon low resolution global ocean modeling. *Ocean Modelling*, 66:64–76, 2013.

M. Brooke, K. Kontomaris, T. Hanratty, and J. McLaughlin. Turbulent deposition and trapping of aerosols at a wall. *Phys. FLuids*, 4:825–834, 1992.

Z. Brzeźniak, M. Capiński, and F. Flandoli. Stochastic partial differential equations and turbulence. *Mathematical Models and Methods in Applied Sciences*, 1(01):41–59, 1991.

M. Budišić and J.-L. Thiffeault. Finite-time braiding exponents. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(8):087407, 2015.

R. Buizza, M. Miller, and T. Palmer. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal Royal Meteorological Society*, 125: 2887–2908, 1999.

A. Cammilleri, F. Gueniat, J. Carlier, L. Pastur, E. Mémin, F. Lusseyran, and G. Artana. POD-spectral decomposition for fluid flow analysis and model reduction. *Theor. and Comp. Fluid Dyn.*, 2013.

J. Candy. *Bayesian signal processing: Classical, modern and particle filtering methods*, volume 54. John Wiley & Sons, 2011.

X. Capet, P. Klein, B. Hua, G. Lapeyre, and J. McWilliams. Surface kinetic energy transfer in surface quasi-geostrophic flows. *J. Fluid Mech.*, 604:165–174, 2008.

M. Caporaloni, F. Tampieri, F. Trombetti, and O. Vittori. Transfer of particles in non-isotropic air turbulence. *J. Atmos. Sci.*, 32:565–568, 1975.

K. Carlberg, C. Bou-Mosleh, and C. Farhat. Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations. *International Journal for Numerical Methods in Engineering*, 86(2):155–181, 2011.

W. Cazemier, R. Verstappen, and A. Veldman. Proper orthogonal decomposition and low-dimensional models for driven cavity flows. *Phys. Fluids*, 10(7):1685–1699, 1998.

G. Cébron. *Processes on the unitary group and free probability.* PhD thesis, Université Pierre et Marie Curie-Paris VI, 2014.

P. Chandramouli, D. Heitz, S. Laizet, and E. Mémin. Coarse-grid large eddy simulations in a wake flow with new stochastic small scale models. ArXiv report, 2016.

B. Chapron, F. Collard, and F. Ardhuin. Direct measurements of ocean surface velocity from space: Interpretation and validation. *Journal of Geophysical Research: Oceans*, 110(C7), 2005.

J. Chasnov. Simulation of the Kolmogorov inertial subrange using an improved subgrid model. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 3(1):188–200, 1991.

M. Chertkov, G. Falkovich, and V. Lebedev. Nonuniversality of the scaling exponents of a passive scalar convected by a random flow. *Physical review letters*, 76(20):3707, 1996.

F. Chinesta, P. Ladeveze, and E. Cueto. A short review on model order reduction based on proper generalized decomposition. *Archives of Computational Methods in Engineering*, 18(4):395–404, 2011.

M. Choi, T. Sapsis, and G. Karniadakis. On the equivalence of dynamically orthogonal and bi-orthogonal methods: Theory and numerical simulations. *Journal of Computational Physics*, 270: 1–20, 2014.

P.-L. Chow. *Stochastic partial differential equations.* CRC Press, 2014.

F. Comte, V. Genon-Catalot, and Y. Rozenholc. Penalized nonparametric mean square estimation of the coefficients of diffusion processes. *Bernoulli*, pages 514–543, 2007.

P. Constantin and G. Iyer. A stochastic Lagrangian representation of the three-dimensional incompressible Navier–Stokes equations. *Communications on Pure and Applied Mathematics*, 61 (3):330–345, 2008.

P. Constantin and G. Iyer. A stochastic-Lagrangian approach to the Navier–Stokes equations in domains with boundary. *The Annals of Applied Probability*, 21(4):1466–1492, 2011.

P. Constantin, A. Majda, and E. Tabak. Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar. *Nonlinearity*, 7(6):1495, 1994.

P. Constantin, Q. Nie, and N. Schörghofer. Front formation in an active scalar equation. *Physical Review E*, 60(3):2858, 1999.

P. Constantin, M. Lai, R. Sharma, Y. Tseng, and J. Wu. New numerical results for the surface quasi-geostrophic equation. *Journal of Scientific Computing*, 50(1):1–28, 2012.

D. Crommelin and E. Vanden-Eijnden. Reconstruction of diffusions using spectral data from timeseries. *Communications in Mathematical Sciences*, 4(3):651–668, 2006.

G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1992. ISBN 9780521385299.

J. D'Adamo, N. Papadakis, E. Mémin, and G. Artana. Variational assimilation of POD low-order dynamical systems. *J. of Turb.*, 8(9):1–22, 2007.

A. Deane, I. Kevrekidis, G. Karniadakis, and S. Orszag. Low-dimensional models for complex geometry fows: Application to grooved channels and circular cylinders. *Phys. Fluids*, 3(10): 2337, 1991.

B. Delyon. Concentration inequalities for the spectral measure of random matrices. *Electronic Communications in Probability*, 15:549–562, 2010.

B. Delyon and J. Yao. On the spectral distribution of gaussian random matrices. *Acta Mathematicae Applicatae Sinica*, 22(2):297–312, 2006.

N. Demni. Free jacobi process. *Journal of Theoretical Probability*, 21(1):118–143, 2008.

N. Demni and T. Hmidi. Spectral distribution of the free unitary brownian motion: another approach. In *Séminaire de Probabilités XLIV*, pages 191–206. Springer, 2012.

G. Dencausse, R. Morrow, M. Rogé, and S. Fleury. Lateral stirring of large-scale tracer fields by altimetry. *Ocean Dynamics*, 64(1):61–78, 2014.

A. Desprès, G. Reverdin, and F. d'Ovidio. Mechanisms and spatial variability of meso scale frontogenesis in the northwestern subpolar gyre. *Ocean Modelling*, 39(1):97–113, 2011a.

A. Desprès, G. Reverdin, and F. d'Ovidio. Summertime modification of surface fronts in the north atlantic subpolar gyre. *Journal of Geophysical Research: Oceans*, 116(C10), 2011b.

A. Doucet and A. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12:656–704, 2009.

A. Doucet, N. De Freitas, and N. Gordon. *Sequential Monte Carlo methods in practice*. Springer, 2001.

F. d'Ovidio, J. Isern-Fontanet, C. López, E. Hernández-García, and E. García-Ladona. Comparison between eulerian diagnostics and finite-size Lyapunov exponents computed from altimetry in the algerian basin. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(1):15–31, 2009.

C. Dufau, M. Orsztynowicz, G. Dibarboure, R. Morrow, and P.-Y. Le Traon. Mesoscale resolution capability of altimetry: Present and future. *Journal of Geophysical Research: Oceans*, 2016.

A. Eliassen. *The quasi-static equations of motion with pressure as independent variable*, volume 17. Grøndahl & Sons boktr., I kommisjon hos Cammermeyers boghandel, 1949.

G. Falkovich, K. Gawędzki, and M. Vergassola. Particles and fields in fluid turbulence. *Reviews of modern Physics*, 73(4):913, 2001.

B. Farrell and P. Ioannou. Statistical state dynamics: a new perspective on turbulence in shear flow. *arXiv preprint arXiv:1412.8290*, 2014.

F. Flandoli. The interaction between noise and transport mechanisms in PDEs. *Milan Journal of Mathematics*, 79(2):543–560, 2011.

D. Florens-Zmirou. On estimating the diffusion coefficient from discrete observations. *Journal of applied probability*, pages 790–804, 1993.

Jason E Frank and Georg A Gottwald. Stochastic homogenization for an energy conserving multi-scale toy model of the atmosphere. *Physica D: Nonlinear Phenomena*, 254:46–56, 2013.

C. Franzke and A. Majda. Low-order stochastic mode reduction for a prototype atmospheric GCM. *Journal of the atmospheric sciences*, 63:457–479, 2006.

C. Franzke, A. Majda, and E. Vanden-Eijnden. Low-order stochastic mode reduction for a realistic barotropic model climate. *Journal of the atmospheric sciences*, 62(6):1722–1745, 2005.

C. Franzke, T. O'Kane, J. Berner, P. Williams, and V. Lucarini. Stochastic climate theory and modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 6(1):63–78, 2015.

B. Galetti, A. Botaro, C.-H. Bruneau, and A. Iollo. Accurate model reduction of transient and forced wakes. *Eur. j. Mech. B, Fluids*, 26(3):354–366, 2007.

C. Garrett and J. Loder. Dynamical aspects of shallow sea fronts. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 302(1472): 563–581, 1981.

K. Gawędzki and A. Kupiainen. Anomalous scaling of the passive scalar. *Physical review letters*, 75(21):3834, 1995.

V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 29(1): 119–151, 1993.

V. Genon-Catalot, C. Laredo, and D. Picard. Non-parametric estimation of the diffusion coefficient by wavelets methods. *Scandinavian Journal of Statistics*, pages 317–335, 1992.

P. Gent and J. Mcwilliams. Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, 20(1):150–155, 1990.

P. Gent, J. Willebrand, T. McDougall, and J. McWilliams. Parameterising eddy-induced tracer transports in ocean circulation models. *J. Phys. Oceanogr.*, 25:463–474, 1995.

M. Germano, U. Piomelli, P. Moin, and W. Cabot. A dynamic subgrid-scale eddy viscosity model. *Phys. of Fluids*, 3:1760–1765, 1991.

H. Giordani, L. Prieur, and G. Caniaux. Advanced insights into sources of vertical velocity in the ocean. *Ocean Dynamics*, 56(5-6):513–524, 2006.

D. Givon, R. Kupferman, and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55, 2004.

G. Gottwald and J. Harlim. The role of additive and multiplicative noise in filtering complex dynamical systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 469(2155):20130096, 2013.

G. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2156), 2013.

G. Gottwald, D. Crommelin, and C. Franzke. Stochastic climate theory. In *Nonlinear and Stochastic Climate Dynamics*. Cambridge University Press, 2015.

J. Gower, K. Denman, and R. Holyer. Phytoplankton patchiness indicates the fluctuation spectrum of mesoscale oceanic structure. *Nature*, 288:157–159, 1980.

I. Grooms and A. Majda. Efficient stochastic superparameterization for geophysical turbulence. *PNAS*, 110(12), 2013.

I. Grooms and A. Majda. Stochastic superparameterization in quasigeostrophic turbulence. *Journal of Computational Physics*, 271:78–98, 2014.

J. Gula, J. Molemaker, and J. McWilliams. Gulf stream dynamics along the southeastern us seaboard. *Journal of Physical Oceanography*, 45(3):690–715, 2015.

G. J Hakim, C. Snyder, and D. Muraki. A new surface model for cyclone-anticyclone asymmetry. *Journal of the atmospheric sciences*, 59(16):2405–2420, 2002.

G. Haller. An objective definition of a vortex. *Journal of fluid mechanics*, 525:1–26, 2005.

G. Haller and T. Sapsis. Lagrangian coherent structures and the smallest finite-time Lyapunov exponent. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(2):023115, 2011.

G. Haller and G. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 147(3):352–370, 2000.

S. Harouna and E. Mémin. A wavelet based numerical simulation of Navier-Stokes equations under uncertainty. Preprint 2014.

K. Hasselmann. Stochastic climate models. part I: theory. *Tellus*, 28:473–485, 1976.

D. Haworth and S. Pope. A generalized langevin model for turbulent flows. *Phys. Fluids*, 29: 387–405, 1986.

I. Held, R. Pierrehumbert, S. Garner, and K. Swanson. Surface quasi-geostrophic dynamics. *Journal of Fluid Mechanics*, 282:1–20, 1995.

M. Hofmann et al. L^p estimation of the diffusion coefficient. *Bernoulli*, 5(3):447–481, 1999.

D. Holm. Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2176), 2015.

P. Holmes, J. Lumley, and G. Berkooz. *Turbulence, coherence structures, dynamical systems and symetry.* Cambridge university press, 1996.

P. Holmes, J. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry.* Cambridge university press, 1998.

B. Hoskins. The geostrophic momentum approximation and the semi-geostrophic equations. *Journal of the Atmospheric Sciences*, 32(2):233–242, 1975.

B. Hoskins. Baroclinic waves and frontogenesis part I: Introduction and eady waves. *Quarterly Journal of the Royal Meteorological Society*, 102(431):103–122, 1976.

B. Hoskins and N. West. Baroclinic waves and frontogenesis. part II: Uniform potential vorticity jet flows-cold and warm fronts. *Journal of the Atmospheric Sciences*, 36(9):1663–1680, 1979.

Ankita Jain, Ilya Timofeyev, and Eric Vanden-Eijnden. Stochastic mode-reduction in models with conservative fast sub-systems. *arXiv preprint arXiv:1410.3004*, 2014.

S. Kadri-Harouna and E. Mémin. Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. Manuscript submitted for publication, 2016.

V. Kalb and A. Deane. An intrinsic stabilization scheme for proper orthogonal decomposition based low-dimensional models. *Phys. Fluids*, 19(5), 2007.

G. Karamanos and G. Karniadakis. A spectral vanishing viscosity method for large-eddy simulations. *J. Comp. Phys.*, 163(1):22–50, 2000.

S. Keating, S. Smith, and P. Kramer. Diagnosing lateral mixing in the upper ocean with virtual tracers: Spatial and temporal resolution dependence. *Journal of Physical Oceanography*, 41(8): 1512–1534, 2011.

S. Keating, A. Majda, and S. Smith. New methods for estimating ocean eddy heat transport using satellite altimetry. *Monthly Weather Review*, 140(5):1703–1722, 2012.

P. Klein, B. Hua, G. Lapeyre, X. Capet, S. Le Gentil, and H. Sasaki. Upper ocean turbulence from high-resolution 3D simulations. *Journal of Physical Oceanography*, 38(8):1748–1763, 2008.

P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer, Berlin, 1999.

V. Klyatskin. *Stochastic equations through the eye of the physicist: Basic concepts, exact results and asymptotic approximations.* Elsevier, 2005.

V. Klyatskin, W. Woyczynski, and D. Gurarie. Short-time correlation approximations for diffusing tracers in random velocity fields: A functional approach. In *Stochastic modelling in physical oceanography*, pages 221–269. Springer, 1996.

A. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. 30(4):299–303, 1941.

B. Koopman. Hamiltonian systems and transformation in Hilbert space. *Proceedings of the National Academy of Sciences*, 17(5):315–318, 1931.

R. Kraichnan. Small-scale structure of a scalar field convected by turbulence. *Physics of Fluids (1958-1988)*, 11(5):945–953, 1968.

R. Kraichnan. Eddy viscosity and diffusivity: exact formulas and approximations. *Complex Systems*, 1(4-6):805–820, 1987.

R. Kraichnan. Anomalous scaling of a randomly advected passive scalar. *Physical review letters*, 72(7):1016, 1994.

H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

A. Kupiainen. Statistical theories of turbulence. *workshop Random media 2000*, June 2000.

T. Kurtz. A limit theorem for perturbed operator semigroups with applications to random evolutions. *Journal of Functional Analysis*, 12(1):55–67, 1973.

S. Laizet and E. Lamballais. High-order compact schemes for incompressible flows: a simple and efficient method with the quasi-spectral accuracy. *J. Comp. Phys.*, 228(15):5989–6015, 2009.

S. Laizet and N. Li. Spectral large-eddy simulation of isotropic and stably stratified turbulence. *Incompact3d, a powerful tool to tackle turbulence problems with up to* $0(10^5)$ *computational cores*, 67(11):1735–1757, 2011.

E. Lamballais, V. Fortunè, and S. Laizet. Straightforward high-order numerical dissipation via the viscous term for direct and large eddy simulation. *J. Comp. Phys.*, 230:3270–3275, 2011.

G. Lapeyre. *Topologie du mélange dans un fluide turbulent géophysique*. PhD thesis, 2000.

G. Lapeyre and P. Klein. Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory. *Journal of physical oceanography*, 36(2):165–176, 2006.

G. Lapeyre, P. Klein, and B. Hua. Does the tracer gradient vector align with the strain eigenvectors in 2D turbulence? *Physics of Fluids*, 11(12):3729–3737, 1999.

G. Lapeyre, P. Klein, and B. Hua. Oceanic restratification forced by surface frontogenesis. *Journal of Physical Oceanography*, 36(8):1577–1590, 2006.

F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 1986.

O. Le Maitre, M. Reagan, H. Najm, R. Ghanem, and O. Knio. A stochastic projection method for fluid flow. II. random process. *Journal of Computational Physics*, 181:9–44, 2002.

B. Legras, I. Pisso, G. Berthet, and F. Lefèvre. Variability of the Lagrangian turbulent diffusion in the lower stratosphere. *Atmospheric Chemistry and Physics*, 5(6):1605–1622, 2005.

Y. Lehahn, F. d'Ovidio, M. Lévy, and E. Heifetz. Stirring of the northeast atlantic spring bloom: A Lagrangian analysis based on multisatellite data. *Journal of Geophysical Research: Oceans*, 112(C8), 2007.

S Leibovich. On wave-current interaction theories of Langmuir circulations. *Journal of Fluid Mechanics*, 99(04):715–724, 1980.

C. Leith. Atmospheric predictability and two-dimensional turbulence. *Journal of the Atmospheric Sciences*, 28(2):145–161, 1971.

C. Leith. Stochastic backscatter in a subgrid-scale model: Plane shear mixing layer. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 2(3):297–299, 1990.

C. Leith and R. Kraichnan. Predictability of turbulent flows. *J. Atmos. Sci*, 29:1041–1058, 1972.

M. Lesieur and O. Metais. New trends in large-eddy simulations of turbulence. *Annual Review of Fluid Mechanics*, 28(1):45–82, 1996.

M. Lesieur and R. Sardouny. Satellite-sensed turbulent ocean structure. *Nature*, 294:673, 1981.

T. Lévy. Schur–weyl duality and the heat kernel measure on the unitary group. *Advances in Mathematics*, 218(2):537–575, 2008.

D. Lilly. On the application of the eddy viscosity concept in the inertial subrange of turbulence. Technical Report 123, NCAR, 1966.

D. Lilly. A proposed modification of the Germano subgrid-scale closure. *Phys. Fluids*, 3:2746–2757, 1992.

E. Lorenz. The predictability of a flow which possesses many scales of motion. *Tellus*, 21:289–307, 1969.

V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, and J. Wouters. Mathematical and physical ideas for climate science. *Reviews of Geophysics*, 52(4):809–859, 2014.

T. Ma, N. Ouellette, and E. Bollt. Stretching and folding in finite time. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 26(2):023112, 2016.

X. Ma, G.-S. Karamanos, and G. Karniadakis. Dynamics and low-dimensionality of a turbulent near wake. *J. Fluid Mech.*, 410:29–65, 2000.

X. Ma, G. Karniadakis, H. Park, and M. Gharib. DPIV-driven simulation: a new computational paradigm. *Proc. R. Soc. Lond. A*, 459:547–565, 2002.

J. MacInnes and F. Bracco. Stochastic particle dispersion modeling and the tracer-particle limit. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 4(12):2809–2824, 1992.

A. Majda. Statistical energy conservation principle for inhomogeneous turbulent dynamical systems. *Proceedings of the National Academy of Sciences*, 112(29):8937–8941, 2015.

A. Majda and P. Kramer. Simplified models for turbulent diffusion: Theory, numerical modelling, and physical phenomena. *Physics report*, 314:237–574, 1999.

A. Majda, I. Timofeyev, and E. Eijnden. Models for stochastic climate prediction. *Proceedings of the National Academy of Sciences*, 96(26):14687–14691, 1999.

A. Majda, Ilya Timofeyev, and Eric Vanden Eijnden. A mathematical framework for stochastic climate models. *Communications on Pure and Applied Mathematics*, 54(8):891–974, 2001.

A. Majda, I. Timofeyev, and E. Vanden-Eijnden. A systematic strategies for stochastic mode reduction in climate. *Journ. Atmos. Sci.*, 60:1705–1722, 2003.

A. Majda, C. Franzke, and B. Khouider. An applied mathematics perspective on stochastic modelling for climate. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366(1875):2427–2453, 2008.

P. Mana and L. Zanna. Toward a stochastic parameterization of ocean mesoscale eddies. *Ocean Modelling*, 79:1–20, 2014.

A Mariotti, M. Moustaoui, B. Legras, and H. Teitelbaum. Comparison between vertical ozone soundings and reconstructed potential vorticity maps by contour advection with surgery. *Journal of Geophysical Research*, 102:6131–6142, 1997.

P. Mason and D. Thomson. Stochastic backscatter in large-eddy simulations of boundary layers. *J. of Fluid Mech.*, 242:51–78, 1992.

I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew-product flows. *Nonlinearity*, 24:1361–1367, 2011.

E. Mémin. Fluid flow dynamics under location uncertainty. *Geophysical & Astrophysical Fluid Dynamics*, 108(2):119–146, 2014. doi: 10.1080/03091929.2013.836190.

C. Mendoza and A. Mancho. Hidden geometry of ocean flows. *Physical review letters*, 105(3): 038501, 2010.

F. Menter. Improved two-equation k-omega turbulence models for aerodynamic flows. *NASA STI/Recon Technical Report N*, 93:22809, 1992.

O. Métais and M. Lesieur. Statistical predictability of decaying turbulence. *Journal of the atmospheric sciences*, 43(9):857–870, 1986.

O. Métais and M. Lesieur. Spectral large-eddy simulation of isotropic and stably stratified turbulence. *J. Fluid Mech.*, 239:157, 1992.

I. Mezic. Spectral properties of dynamical systems, model reduction and decompositions. *Nonlinear Dyn*, 41(1-3):309–325, 2005.

I. Mezić, S. Loire, V. Fonoberov, and P. Hogan. A new mixing diagnostic and gulf oil spill movement. *Science*, 330(6003):486–489, 2010.

Y. Michel. Estimating deformations of random processes for correlation modelling in a limited area model. *Quarterly Journal of the Royal Meteorological Society*, 139(671):534–547, 2013a.

Y Michel. Estimating deformations of random processes for correlation modelling: methodology and the one-dimensional case. *Quarterly Journal of the Royal Meteorological Society*, 139(672): 771–783, 2013b.

R. Mikulevicius and B. Rozovskii. Stochastic Navier–Stokes equations for turbulent flows. *SIAM Journal on Mathematical Analysis*, 35(5):1250–1310, 2004.

I. Mirouze and A. Weaver. Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Quarterly Journal of the Royal Meteorological Society*, 136(651): 1421–1443, 2010.

L. Mitchell and G. Gottwald. Data assimilation in slow-fast systems using homogenized climate models. *Journal of the atmospheric sciences*, 69(4):1359–1377, 2012.

D. Muraki, C. Snyder, and R. Rotunno. The next-order corrections to quasigeostrophic theory. *Journal of the atmospheric sciences*, 56(11):1547–1560, 1999.

N. Nakamura. A new look at eddy diffusivity as a mixing diagnostic. *Journal of the atmospheric sciences*, 58(24):3685–3701, 2001.

W. Neves and C. Olivera. Wellposedness for stochastic continuity equations with ladyzhenskaya–prodi–serrin condition. *Nonlinear Differential Equations and Applications NoDEA*, 22(5):1247–1258, 2015.

G.-H. Ng, D. McLaughlin, D. Entekhabi, and A. Ahanin. The role of model dynamics in ensemble Kalman filter performance for chaotic systems. *Tellus A*, 63(5):958–977, 2011.

B. Noack, P. Papas, and P. Monkevitz. The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.*, 523(523):339–365, 2005.

B. Noack, M. Morzynski, and G. Tadmor. *Reduced-Order Modelling for Flow Control*, volume 528 of *CISM Courses and Lectures*. Springer-Verlag, 2010.

I. Nourdin. *Calcul stochastique généralisé et applications au mouvement brownien fractionnaire; Estimation non-paramétrique de la volatilité et test d'adéquation.* PhD thesis, Université Henri Poincaré-Nancy I, 2004.

B. Oksendal. *Stochastic differential equations.* Spinger-Verlag, 1998.

A. Okubo. Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences. 17(3):445–454, 1970.

Y. Orsolini, G. Hansen, G. Manney, N. Livesey, and U. Hoppe. Lagrangian reconstruction of ozone column and profile at the arctic lidar observatory for middle atmosphere research (alomar) throughout the winter and spring of 1997–1998. *Journal of Geophysical Research: Atmospheres*, 106(D9):10011–10021, 2001.

S. Orszag. Analytical theories of turbulence. *Journal of Fluid Mechanics*, 41(02):363–386, 1970.

J. Östh, B. Noack, S. Krajnović, D. Barros, and J. Borée. On the need for a nonlinear subscale turbulence term in POD models as exemplified for a high-Reynolds-number flow over an Ahmed body. *Journal of Fluid Mechanics*, 747:518–544, 2014.

T. Palmer and P. Williams. Introduction. stochastic physics and climate modelling. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 366(1875):2419–2425, 2008.

O. Pannekoucke and S. Massart. Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 134(635):1425–1438, 2008.

O. Pannekoucke, E. Emili, and O. Thual. Modelling of local length-scale dynamics and isotropizing deformations. *Quarterly Journal of the Royal Meteorological Society*, 140(681):1387–1398, 2014.

N. Papadakis, E. Mémin, A. Cuzol, and N. Gengembre. Data assimilation with the weighted ensemble Kalman filter. *Tellus A*, 62(5):673–697, 2010.

G. Papanicolaou and W. Kohler. Asymptotic theory of mixing stochastic ordinary differential equations. *Communications on Pure and Applied Mathematics*, 27(5):641–668, 1974.

P. Parnaudeau, E. Lamballais, D. Heitz, and J.H. Silvestrini. *Combination of the immersed boundary method with compact schemes for DNS of flows in complex geometry.* Kluwer academic publishers, 2003.

P. Parnaudeau, J. Carlier, D. Heitz, and E. Lamballais. Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900. *Physics of Fluids*, 20(8):085101, 2008.

E. Parzen. On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.

R. Pasquetti. Spectral vanishing viscosity method for large-eddy simulation of turbulent flows. *J. Sci. Comp.*, 27(1-3):365–375, 2006.

G. Pavliotis and A. Stuart. *Multiscale methods: Averaging and homogeneization.* Springer, 2008.

D. Peavoy, C. Franzke, and G. Roberts. Systematic physics constrained parameter estimation of stochastic differential equations. *Computational Statistics & Data Analysis*, 83:182–199, 2015.

R. Peikert, A. Pobitzer, F. Sadlo, and B. Schindler. A comparison of finite-time and finite-size Lyapunov exponents. In *Topological Methods in Data Analysis and Visualization III*, pages 187–200. Springer, 2014.

C. Penland. Noise out of chaos and why it won't go away. *Bulletin of the American Meteorological Society*, 84(7):921, 2003a.

C. Penland. A stochastic approach to nonlinear dynamics: A review (extended version of the article-" noise out of chaos and why it won't go away"). *Bulletin of the American Meteorological Society*, 84(7):925–925, 2003b.

C. Penland and L. Matrosova. A balance condition for stochastic numerical models with application to the El Nino-southern oscillation. *Journal of climate*, 7(9):1352–1372, 1994.

C. Penland and P. Sardeshmukh. The optimal growth of tropical sea surface temperature anomalies. *Journal of climate*, 8(8):1999–2024, 1995.

R. Pierrehumbert and H. Yang. Global chaotic mixing on isentropic surfaces. *Journal of the atmospheric sciences*, 50(15):2462–2480, 1993.

U. Piomelli, W. Cabot, P. Moin, and S. Lee. Subgrid scale backscatter in turbulent and transitional flows. *Phys. Fluids*, 3(7):1766–1771, 1991.

A. Poje, A. Haza, T. Özgökmen, M. Magaldi, and Z. Garraffo. Resolution dependent relative dispersion statistics in a hierarchy of ocean models. *Ocean Modelling*, 31(1):36–50, 2010.

S. Pope. Lagrangian PDF methods for turbulent flows. *Annu. Rev. Fluid Mech.*, 26:23–63, 1994.

S. Pope. *Turbulent flows*. Cambridge University Press, 2000.

S. Pope. Self-conditioned fields for large-eddy simulations of turbulent flows. *Journal of Fluid Mechanics*, 652:139–169, 2010.

C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*, volume 1905. Springer, 2007.

J. Price, M. Reed, M. Howard, W. Johnson, Z.-G. Ji, C. Marshall, N. Guinasso, and G. Rainey. Preliminary assessment of an oil-spill trajectory model using satellite-tracked, oil-spill-simulating drifters. *Environmental Modelling & Software*, 21(2):258–270, 2006.

B. Protas, B. Noack, and J. Östh. Optimal nonlinear eddy viscosity in Galerkin models of turbulent flows. *Journal of Fluid Mechanics*, 766:337–367, 2015.

F. Ragone and G. Badin. A study of surface semi-geostrophic turbulence: freely decaying dynamics. *Journal of Fluid Mechanics*, 792:740–774, 2016.

P. Rao. *Statistical inference for diffusion type processes*. Arnold, 1999.

M. Reeks. The transport of discrete particles in inhomogeneous turbulence. *Journal of aerosol science*, 14(6):729–739, 1983.

D. Rempfer and H. Fasel. Evolution of three-dimensional coherent structures in a flat-plate boundary layer. *Journal of Fluid Mechanics*, 260:351–375, 1994.

V. Resseguier, E. Mémin, and B. Chapron. Reduced flow models from a stochastic Navier-Stokes representation. *Annales de l'ISUP*, 2015.

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part I: Random transport and general models. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017a.

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part II: Quasi-geostrophic models and efficient ensemble spreading. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017b.

V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, part III: SQG and frontal dynamics under strong turbulence. Manuscript submitted for publication in Geophysical & Astrophysical Fluid Dynamics, 2017c.

V. Resseguier, E. Mémin, and B. Chapron. Stochastic modeling and diffusion modes for POD models and small-scale flow analysis. Under consideration for Journal of Fluid Mechanics, 2017d.

P. Richardson. Eddy kinetic energy in the north atlantic from surface drifters. *Journal of Geophysical Research: Oceans (1978–2012)*, 88(C7):4355–4367, 1983.

M. Romito. Time regularity of the densities for the Navier–Stokes equations with noise. *Journal of Evolution Equations*, pages 1–16, 2016.

M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, 27(3):832–837, 1956.

C. Rowley, I. Mezic, S. Bagheri, P. Schlatter, and D. Henningson. Spectral analysis of nonlinear flows. *J. Fluid Mech.*, 641:115–127, 2009.

T. Sapsis. *Dynamically orthogonal field equations for stochastic fluid flows and particle dynamics.* PhD thesis, Massachusetts Institute of Technology, 2011.

T. Sapsis. Attractor local dimensionality, nonlinear energy transfers and finite-time instabilities in unstable dynamical systems with applications to two-dimensional fluid flows. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 469(2153), 2013. ISSN 1364-5021. doi: 10.1098/rspa.2012.0550.

T. Sapsis and P. Lermusiaux. Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty. *Physica D: Nonlinear Phenomena*, 241(1):60–76, 2012.

T. Sapsis and A. Majda. A statistically accurate modified quasilinear Gaussian closure for uncertainty quantification in turbulent dynamical systems. *Physica D: Nonlinear Phenomena*, 252: 34–45, 2013a.

T. Sapsis and A. Majda. Blending modified Gaussian closure and non-Gaussian reduced subspace methods for turbulent dynamical systems. *Journal of Nonlinear Science*, 23(6):1039–1071, 2013b.

T. Sapsis and A. Majda. Statistically accurate low-order models for uncertainty quantification in turbulent dynamical systems. *Proceedings of the National Academy of Sciences*, 110(34): 13705–13710, 2013c.

B. Sawford. Generalized random forcing in random-walk models of turbulent dispersion model. *Phys. Fluids*, 29:3582–3585, 1986.

P. Schmid. Dynamic mode decomposition of numerical and experimental data. *J. Fluid Mech.*, 656:5–28, 2010.

U. Schumann. Stochastic backscatter of turbulence energy and scalar variance by random subgrid-scale fluxes. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 451, pages 293–318. The Royal Society, 1995.

G. Sehmel. Particle deposition from turbulent air flow. *J. Geophys. Res.*, 75:1766–1781, 1970.

R. Semaan, P. Kumar, M. Burnazzi, G. Tissot, L. Cordier, and B. Noack. Reduced-order modeling of the flow around a high-lift configuration with unsteady coanda blowing. *J. Fluid Mech.*, 800: 72–110, 2016.

B. Shivamoggi and G. van Heijst. The Okubo-Weiss criteria in two-dimensional hydrodynamic and magnetohydrodynamic flows. *arXiv preprint arXiv:1110.6190*, 2011.

G. Shutts. A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 612:3079–3012, 2005.

L. Sirovich. Turbulence and the dynamics of coherent structures. *Quart. J. Applied Math.*, 45: 561–590, 1987.

J. Slingo and T. Palmer. Uncertainty in weather and climate prediction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1956):4751–4767, 2011.

J. Smagorinsky. General circulation experiments with the primitive equation: I. the basic experiment. *Monthly Weather Review*, 91:99–165, 1963.

C. Snyder, T. Bengtsson, and M. Morzfeld. Performance bounds for particle filters using the optimal proposal. *Monthly Weather Review*, 143:4750–4761, 2015.

H. Sørensen. Parametric inference for diffusion processes observed at discrete points in time: a survey. *International Statistical Review*, 72(3):337–354, 2004.

D. Stammer. Global characteristics of ocean variability estimated from regional TOPEX/Poseidon altimeter measurements. *Journal of Physical Oceanography*, 27(8):1743–1769, 1997.

R. Sutton, H. Maclean, R. Swinbank, A. O'Neill, and F. Taylor. High-resolution stratospheric tracer fields estimated from satellite observations using Lagrangian trajectory calculations. *Journal of the atmospheric sciences*, 51(20):2995–3005, 1994.

E. Tadmor. Convergence of spectral methods for nonlinear conservation laws. *SIAM J. Numer. Anal*, 26(1):30–44, 1989.

F. Takens. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381. Springer-Verlag, 1981.

J.-L. Thiffeault. Stretching and curvature of material lines in chaotic flows. *Physica D: Nonlinear Phenomena*, 198(3):169–181, 2004.

J.-L. Thiffeault and A. Boozer. Geometrical constraints on finite-time Lyapunov exponents in two and three dimensions. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(1):16–28, 2001.

G. Tissot, L. Cordier, N. Benard, and B. Noack. 4D variational data assimilation for POD reduced-order models. In *Turbulence and Shear Flow Phenomena (TSFP-8)*, 2013.

A.-M. Tréguier, I. Held, and V. Larichev. Parameterization of quasigeostrophic eddies in primitive equation ocean models. *Journal of Physical Oceanography*, 27(4):567–580, 1997.

A. Trevisan and L. Palatella. On the Kalman filter error covariance collapse into the unstable subspace. *Nonlinear Processes in Geophysics*, 18(2):243–250, 2011.

A. Trevisan and F. Uboldi. Assimilation of standard and targeted observations within the unstable subspace of the observation-analysis-forecast cycle system. *Journal of the atmospheric sciences*, 61(1):103–113, 2004.

C. Ubelmann, P. Klein, and L.-L. Fu. Dynamic interpolation of sea surface height and potential applications for future high-resolution altimetry mapping. *Journal of Atmospheric and Oceanic Technology*, 32(1):177–184, 2015.

M. Ueckermann, P. Lermusiaux, and T. Sapsis. Numerical schemes for dynamically orthogonal equations of stochastic fluid and ocean flows. *Journal of Computational Physics*, 233:272–294, 2013.

G. Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2006.

J. van Waaij and H. van Zanten. Gaussian process methods for one-dimensional diffusions: optimal rates and adaptation. *Electronic Journal of Statistics*, 10(1):628–645, 2016.

M. Veneziani, A. Griffa, A. Reynolds, and A. Mariano. Oceanic turbulence and stochastic models from subsurface Lagrangian data for the Northwest Atlantic Ocean. *Journal of physical oceanography*, 34(8):1884–1906, 2004.

A. Weaver and P. Courtier. Correlation modelling on the sphere using a generalized diffusion equation. *Quart. J. Roy. Meteor. Soc.*, 127:1815–846, 2001.

J. Weiss. The dynamics of enstrophy transfer in two-dimensional hydrodynamics. *Physica D: Nonlinear Phenomena*, 48(2):273–294, 1991.

P. Welander. Studies on the general development of motion in a two-dimensional, ideal fluid. *Tellus*, 7(2):141–156, 1955.

D. Wilcox. Reassessment of the scale-determining equation for advanced turbulence models. *AIAA journal*, 26(11):1299–1310, 1988.

K. Wyrtki, L. Magaard, and J. Hager. Eddy energy in the oceans. *Journal of Geophysical Research*, 81(15):2641–2646, 1976.

Résumé

Cette thèse concerne le développement, l'extension et l'application d'une formulation stochastique des équations de la mécanique des fluides introduite par Mémin (2014). La vitesse petite échelle, non-résolue, est modélisée au moyen d'un champ aléatoire décorrélé en temps. Cela modifie l'expression de la dérivée particulaire et donc les équations de la mécanique des fluides. Les modèles qui en découlent sont dénommés modèles sous incertitude de position. La thèse s'articule autour de l'étude successive de modèles réduits, de versions stochastiques du transport et de l'advection à temps long d'un champ de traceur par une vitesse mal résolue.

La POD est une méthode de réduction de dimension, pour EDP, rendue possible par l'utilisation d'observations. L'EDP régissant l'évolution de la vitesse du fluide est remplacée par un nombre fini d'EDOs couplées. Grace à la modélisation sous incertitude de position et à de nouveaux estimateurs statistiques, nous avons dérivé et simulé des versions réduites, déterministes et aléatoires, de l'équation de Navier-Stokes.

Après avoir obtenu des versions aléatoires de plusieurs modèles océaniques, nous avons montré numériquement que ces modèles permettaient de mieux prendre en compte les petites échelles des écoulements, tout en donnant accès à des estimés de bonne qualité des erreurs du modèle. Ils permettent par ailleurs de mieux rendre compte des évènements extrêmes, des bifurcations ainsi que des phénomènes physiques réalistes absents de certains modèles déterministes équivalents.

Nous avons expliqué, démontré et quantifié mathématiquement l'apparition de petites échelles de traceur, lors de l'advection par une vitesse mal résolu. Cette quantification permet de fixer proprement des paramètres de la méthode d'advection Lagrangienne, de mieux le comprendre le phénomène de mélange et d'aider au paramétrage des simulations grande échelle en mécanique des fluides.

Mots-clés : Qantification d'incertitude, prévision d'ensemble, bifurcation, calcul stochastique, mécanique des fluides, géophysique, dynamique de surface dans l'océan, modèle d'ordre réduit, proper orthogonal decomposition, statistique des processus, advection Lagrangienne, mélange, repliement.

Abstract

This thesis develops, analyzes and demonstrates several valuable applications of randomized fluid dynamics models referred to as under location uncertainty. The velocity is decomposed between large-scale components and random time-uncorrelated small-scale components. This assumption leads to a modification of the material derivative and hence of every fluid dynamics models. Through the thesis, the mixing induced by deterministic low-resolution flows is also investigated.

We first applied that decomposition to reduced order models (ROM). The fluid velocity is expressed on a finite-dimensional basis and its evolution law is projected onto each of these modes. We derive two types of ROMs of Navier-Stokes equations. A deterministic LES-like model is able to stabilize ROMs and to better analyze the influence of the residual velocity on the resolved component. The random one additionally maintains the variability of stable modes and quantifies the model errors.

We derive random versions of several geophysical models. We numerically study the transport under location uncertainty through a simplified one. A single realization of our model better retrieves the small-scale tracer structures than a deterministic simulation. Furthermore, a small ensemble of simulations accurately predicts and describes the extreme events, the bifurcations as well as the amplitude and the position of the ensemble errors. Another of our derived simplified model quantifies the frontolysis and the frontogenesis in the upper ocean.

This thesis also studied the mixing of tracers generated by smooth fluid flows, after a finite time. We propose a simple model to describe the stretching as well as the spatial and spectral structures of advected tracers. With a toy flow but also with satellite images, we apply our model to locally and globally describe the mixing, specify the advection time and the filter width of the Lagrangian advection method, as well as the turbulent diffusivity in numerical simulations.

Keywords : Uncertainty quantification, ensemble forecast, bifurcation, stochastic calculus, fluid dynamics, geophysics, upper ocean dynamics, reduced order model, proper orthogonal decomposition, processes statistics, Lagrangian advection, mixing, folding.