



HAL
open science

Environnement logiciel et étalonnage de l'échelle en énergie des jets dans l'expérience ATLAS

S. Binet

► **To cite this version:**

S. Binet. Environnement logiciel et étalonnage de l'échelle en énergie des jets dans l'expérience ATLAS. Modélisation et simulation. Université Blaise Pascal - Clermont-Ferrand II, 2006. Français. NNT : . tel-00140524

HAL Id: tel-00140524

<https://theses.hal.science/tel-00140524>

Submitted on 6 Apr 2007

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.

UNIVERSITE BLAISE PASCAL
(U.F.R. de Recherche Scientifique et Technique)

ECOLE DOCTORALE DES SCIENCES FONDAMENTALES

THESE

présentée pour obtenir le grade de

DOCTEUR D'UNIVERSITE
(SPECIALITE : PHYSIQUE DES PARTICULES)

par

Sébastien BINET

Maître ès-Sciences, Diplômé d'Etudes Approfondies

**Environnement logiciel et étalonnage de l'échelle
en énergie des jets dans l'expérience ATLAS**

Thèse soutenue le 19 janvier 2006, devant la commission d'examen :

Président :	M.	A.	BALDIT
Examineurs :	M.	K.	ASSAMAGAN
	M.	Ph.	GRIS
	M.	D.	PALLIN
	M.	D.	ROUSSEAU
Rapporteurs :	Mme	U.	BASSLER
	M.	E.	LANÇON

Remerciements

Je remercie tout d'abord Bernard Michel, ancien directeur du Laboratoire de Physique Corpusculaire de Clermont-Ferrand, pour m'avoir accueilli au sein de son établissement, ainsi qu'Alain Baldit, actuel directeur et président de ce jury de thèse.

J'exprime toute ma gratitude à Dominique Pallin et Philippe Gris pour avoir dirigé et encadré ce travail de recherche ainsi que la liberté précieuse qu'ils ont su m'accorder. Je remercie également François Vazeille et l'ensemble de l'équipe ATLAS du LPC pour m'avoir intégré au sein du groupe.

Je suis extrêmement redevable aux membres du jury, Ursula Bassler et Eric Lançon qui ont accepté de rapporter sur ce travail, leurs critiques et suggestions ont grandement amélioré le manuscrit initial, et enfin Kétévi Assamagan et David Rousseau, tout d'abord pour leur présence le jour de la soutenance, mais également pour leurs judicieuses remarques.

Je voudrais également remercier mon compagnon de bureau, David Calvet, pour avoir assuré le bon déroulement informatique de ma thèse et pour les nombreuses discussions et conseils en programmation C++ qui m'ont été prodigués. Mon seul regret sera de n'avoir pas pu le convertir au Python. . . il m'aurait sans doute fallu juste une année de plus. J'exprime ma reconnaissance envers Fabrice Podlisky et Claudio Santoni pour les multiples discussions que nous avons eues autour d'un café ou *ristretto*, elles ont indubitablement contribué à la gest(at)ion de la thèse.

Je voudrais de plus profiter de cette opportunité pour remercier Hélène Fonvieille pour m'avoir permis de réaliser mes premiers pas dans la recherche, en DEUG avec Claudio, en Licence avec Jean Orloff et en maîtrise avec Alexandre Deur. Sans l'ombre d'un doute, cette thèse a commencé avec un stage Janus.

Je remercie mes amis ou amies, thésards ou non, pour m'avoir supporté pendant ces trois années et un peu plus. Un simple salut de la main serait bien pauvre pour vous exprimer toute ma gratitude.

Je ne pouvais finir qu'en remerciant mes parents et grands-parents ainsi que tous les autres membres de ma famille pour leur soutien inconditionnel, sans faille et continu, tout au long de cette aventure. Ce travail vous doit, à tous, beaucoup. En un mot : *merci*.

Table des matières

Résumé	ii
Abstract	ii
Remerciements	iii
1 Le quark <i>top</i> et le Modèle Standard	3
1.1 Éléments du Modèle Standard	3
1.1.1 La quête des constituants élémentaires de la matière	4
1.1.2 Interactions fondamentales	5
1.1.2.1 Interactions électrofaibles	5
1.1.2.2 Interaction forte	8
1.2 Mesures de précision	9
1.3 Le quark <i>top</i>	10
1.3.1 Historique	12
1.3.2 Caractéristiques du quark <i>top</i>	12
1.3.3 Mesure de la masse	13
1.3.3.1 Mesure indirecte	13
1.3.3.2 Mesure directe	13
2 Le détecteur ATLAS au LHC	17
2.1 Le grand collisionneur de hadrons	17
2.1.1 L'accélérateur	17
2.1.1.1 Les détecteurs	19
2.2 Le détecteur ATLAS	20
2.2.1 Objectifs de physique et sous-détecteurs	20
2.2.2 Détecteurs internes	23
2.2.3 Calorimètres	24
2.2.3.1 Calorimètres électromagnétiques	25
2.2.3.2 Calorimètres hadroniques	27
2.2.4 Spectromètre à muons	30
2.2.5 Système d'acquisition	31
2.2.5.1 Niveau 1	33
2.2.5.2 Niveau 2	34
2.2.5.3 Filtre d'événements	35

3	Environnement logiciel	37
3.1	Éléments de C++	37
3.2	Introduction à ATHENA	38
3.3	Organisation d'ATHENA	39
3.3.1	Objets de base d'ATHENA	40
3.3.2	Accès aux données : StoreGateSvc	40
3.4	Flux de données	43
3.4.1	Génération	44
3.4.2	Simulation	44
3.4.2.1	Simulation complète : Geant4	45
3.4.2.2	Simulation rapide : AtlFast	46
3.4.3	Reconstruction	48
3.4.3.1	Reconstruction des traces	48
3.4.3.2	Reconstruction des dépôts d'énergie	49
3.4.3.3	Reconstruction combinée	49
3.4.4	Analyse	51
3.5	Outils pour l'analyse de données	52
3.5.1	Classes pour l'analyse	53
3.5.2	Outils pour l'analyse	54
3.5.3	Algorithmes d'analyse	54
3.6	Outils de filtrage	54
3.6.1	Prédicats de sélection	55
3.6.2	Filtrage de vertex	55
3.7	Outils d'associations	57
3.8	Réorganisation de l'implémentation du Monte Carlo	59
3.9	Au-delà de l'analyse avec des n-tuples	59
3.10	Grilles de calcul	60
3.11	Conclusion	60
4	Étalonnage en énergie des jets	61
4.1	Différentes méthodes d'étalonnage	63
4.1.1	Étalonnage en faisceau test	63
4.1.2	Étalonnage <i>in situ</i>	63
4.1.2.1	E/p	64
4.1.2.2	Balance en impulsion transverse	64
4.1.2.3	Résonance	64
4.2	Étalonnage des jets légers <i>via</i> la résonance du boson W	65
4.2.1	Sélection des événements	65
4.2.2	Méthode d'étalonnage	73
4.2.3	Extraction des coefficients d'étalonnage	74
4.2.3.1	Méthode du χ^2	74
4.2.3.2	Multiplicateurs de Lagrange	77
4.2.4	Comparaison des résultats	83
4.2.5	Validation avec la reconstruction complète	83
4.3	Étalonnage des jets de b <i>via</i> la résonance du boson Z^0	85
4.3.1	Étude du canal $Z^0\gamma \rightarrow b\bar{b}\gamma$	88
4.3.1.1	Etude de stratégie de reconstruction	96

4.4 Conclusion	100
Bibliography	103

Table des figures

1.1	Ecart entre les valeurs mesurées des observables électrofaibles et les valeurs prédites par le Modèle Standard. Les valeurs préliminaires présentées dans ce tableau sont tirées de [1].	11
1.2	Masse du boson W en fonction de la masse du quark top . Les contours à 68% de niveau de confiance pour les mesures directes et indirectes sont indiquées. Différentes hypothèses pour la masse du boson de Higgs sont superposées. Cette figure est extraite de [1].	14
2.1	Le complexe des accélérateurs du CERN.	19
2.2	Le détecteur ATLAS.	20
2.3	Sensibilité d'après [10] pour la découverte du boson de Higgs du Modèle Standard en fonction de sa masse, avec le détecteur ATLAS. La luminosité intégrée de $10 fb^{-1}$ correspond à un an de prise de données à basse luminosité.	21
2.4	Système de trajectographie interne.	24
2.5	Système de calorimétrie d'ATLAS.	25
2.6	Schéma de la structure en accordéon du LArg d'ATLAS.	26
2.7	Coupe d'un module du TileCal. Nous pouvons noter l'orientation des tuiles selon la direction de propagation des hadrons. Ce design original facilite le routage des fibres optiques pour la lecture et permet également la redondance de cette information (les tuiles sont lues de chaque côté).	28
2.8	Les cellules du TileCal. Nous retrouvons le tonneau central sur la partie gauche du schéma, alors que la partie droite est consacrée au tonneau étendu. Les cellules A composent la première couche. Les cellules B et C (qui sont reliées) constituent la deuxième couche. La troisième et dernière couche est composée des cellules D.	29
2.9	Vue d'ensemble du spectromètre à muons. Nous pouvons distinguer les quatre types de chambres à muons : TGC, RPC, CSC et MDT.	30
2.10	Sections efficaces et taux de productions de différents processus pour des collisions proton-(anti)proton en fonction de l'énergie disponible dans le centre de masse. Ces résultats ont été obtenus pour un fonctionnement à haute luminosité.	32
2.11	Vue schématique des différents niveaux de déclenchement constituant le système d'acquisition.	33

- 3.1 Les différents objets de base qu'ATHENA met à disposition et en relation. On retrouve au centre de cette architecture la classe `Algorithm`. Les instances de cette classe sont chargées par le gestionnaire d'application (*Application Manager*) et appelées à chaque nouvel événement par le contrôleur de boucle (*Event Loop Manager*). Chaque instance d'une classe dérivant d'`Algorithm` a ensuite accès aux différents services mis à disposition par Gaudi/ATHENA : le *Message Service* pour contrôler le niveau de verbosité de l'algorithme, le *JobOptions Service* pour configurer les paramètres d'exécution de cet algorithme et le service de gestion des histogrammes. Les instances de chaque `Algorithm` peuvent également récupérer des pointeurs vers des `AlgTools` via le service des outils. Ces `AlgTools` peuvent être appelés plusieurs fois par événement, à la discrétion de l'implémentation de chaque algorithme (au contraire des instances de classes dérivant de `Algorithm` qui ne sont appelées qu'une et une seule fois par le contrôleur de boucle d'événements). 41
- 3.2 Modèle du tableau noir (`StoreGateSvc`) : flux apparent et flux réel des données entre algorithmes. L'algorithme A récupère les données T1 depuis `StoreGate` et enregistre les données T2 et T3 à la fin de son exécution. Puis l'algorithme B récupère les données T2 et publie les données T4 après traitement. Vient ensuite l'algorithme C dans l'ordre d'exécution qui a besoin des données T3 et T4 afin de produire les données T5. Vu de l'extérieur, tout se passe comme si l'algorithme C demandait les données T3 à l'algorithme A et les données T4 à l'algorithme B mais le seul vrai interlocuteur de chaque algorithme est `StoreGate`. On peut ainsi remplacer l'algorithme B par un algorithme B' qui publiera d'autres données T4 sans pour cela modifier toute la chaîne d'exécution en aval. 42
- 3.3 Schéma simplifié des différents flux de données lors de la simulation rapide (`AtlFast` [35]) et de la simulation complète (`Geant4` [33]). Les rectangles représentent les différentes étapes du processus de gestion du flux de données. Les ovoïdes représentent les types d'objets échangés entre les différentes étapes. 43
- 3.4 Flot des données lors de la simulation complète du détecteur. Les différentes étapes de la simulation sont représentées par les rectangles et les objets produits au cours de chaque étape, par des ovoïdes. Les rectangles en pointillés sont des étapes optionnelles de la simulation. La première ligne décrit l'étape génération qui fournit les événements de physique. 46
- 3.5 Facteurs de réjection d'`AtlFastB` (en haut) et fonctions d'étalonnage d'`AtlFastB` (en bas) pour différentes valeurs d'efficacité d'étiquetage des jets de b 47
- 3.6 Chaîne de reconstruction des traces. Les boîtes de la ligne du haut représentent les objets de données, celles du bas étant les algorithmes. Les flèches représentent la direction du flux de données. 49
- 3.7 Chaîne de reconstruction des dépôts d'énergie. Les boîtes de la ligne du haut représentent les objets de données, celles du bas étant les algorithmes. Les flèches représentent la direction du flux de données. 50

3.8	Diagrammes UML de l'arbre d'héritage des classes d'analyse : à gauche les classes de base et à droite les classes dédiées à la reconstruction de chaînes de désintégration. Les classes dont le nom est en italique sont abstraites et celles dont le nom commence par un I sont virtuelles [21].	53
3.9	Diagramme UML des classes de filtrages. Toutes les classes implémentent la méthode <code>isAccepted()</code> qui réalise l'opération de filtrage.	56
3.10	Exemple de code illustrant la simplicité des filtres de particules : seuls les éléments du vecteur <code>parts</code> contenant un quark b avec une énergie d'au moins 40 GeV seront copiés dans le vecteur <code>myBs</code> . Ce filtre pourra être exécuté indifféremment sur des jets ou des particules Monte-Carlo.	57
3.11	Exemple de <code>jobOption</code> illustrant la configuration d'un filtre Monte-Carlo sélectionnant différentes particules satisfaisant des motifs de désintégration. Des <code>patterns</code> peuvent être rajoutés ou enlevés sans nécessiter de recompilation du code d'analyse.	58
3.12	Diagramme UML des classes de filtrage d'associations. Toutes les classes implémentent la méthode <code>isAccepted()</code> qui réalise l'opération de filtrage.	58
4.1	Diagrammes de Feynman de production de paires $t\bar{t}$. Au LHC, environ 90% des paires $t\bar{t}$ seront produites par fusion de gluons.	66
4.2	Masse invariante de tous les candidats W en traits pleins rouges (haut et bas). Le bruit de fond combinatoire est représenté en pointillés bleus coiffant la zone hachurée (figure du haut). Dans la figure du bas, les traits pointillés bleus représentent la masse invariante des jets associés aux quarks originaires du W	69
4.3	Masse invariante des paires de jets réalisant la meilleure combinaison par rapport au pic en masse. Les traits pleins rouges représentent la meilleure combinaison par rapport au pic en masse (haut et bas). Le résultat de son ajustement par une gaussienne est donné sur la figure du haut. Le bruit de fond combinatoire est représenté en pointillés bleus sur la figure du haut. Les pointillés bleus de la figure du bas représentent la masse invariante reconstruite à partir des jets qui ont été associés aux quarks provenant du W Monte-Carlo. Le résultat de son ajustement par une gaussienne est donné sur la figure du bas.	70
4.4	Masse invariante des paires de jets réalisant la meilleure combinaison par rapport au pic en masse, pour la coupure avec au moins un jet de b dans l'événement (haut) et avec deux jets de b (bas). Les différents types de bruits de fonds combinatoires sont représentés par les zones hachurées en bleu. Les hachures inclinées sur la droite correspondent aux événements pour lesquels une meilleure combinaison jet-jet est trouvée, supplantant celle qu'il aurait fallu choisir. Les hachures inclinées sur la gauche correspondent aux événements pour lesquels au moins un des jets de la bonne combinaison ne passe pas les coupures de présélection.	71
4.5	Masses invariantes (en traits pleins rouges) de toutes les combinaisons $j\bar{j}b$ (figure du haut) et masse invariante de la combinaison maximisant l'impulsion transverse du top ainsi reconstruit (figure du bas). Le bruit de fond combinatoire est représenté par la zone hachurée en pointillés bleus.	72

4.6	Résolutions en E , η et ϕ obtenues pour les événements $t\bar{t}$ simulés et reconstruits avec <code>AtlFast-01-00-91</code>	76
4.7	Distribution des coefficients de correction en énergie obtenus événement par événement <i>via</i> un ajustement par MINUIT. Il y a 2 entrées par événement puisque l'ajustement d'un W porte sur les 2 jets le constituant.	77
4.8	Distribution des coefficients de correction pour l'intervalle en énergie [75 GeV; 80 GeV]. Ces coefficients de corrections sont ajustés par une gaussienne sur l'intervalle asymétrique [$pic - 2 \times \sigma$; $pic + \sigma$] afin de limiter les effets dûs aux radiations de gluons. La moyenne de la gaussienne sera ensuite utilisée pour paramétriser l'évolution des constantes de correction.	78
4.9	Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec l'ajustement par MINUIT. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4.	79
4.10	Distribution de la masse invariante m_{jj} avant (en traits pleins rouges) et après application des facteurs de correction événement par événement (en traits pointillés bleus).	80
4.11	Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec les multiplicateurs de Lagrange. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4.	82
4.12	Différence entre la masse du W après application des poids obtenus événement par événement pour l'ajustement <i>via</i> MINUIT et pour l'ajustement en utilisant les multiplicateurs de Lagrange.	84
4.13	Différence entre les facteurs de correction en énergie obtenus événement par événement pour l'ajustement <i>via</i> MINUIT et pour l'ajustement en utilisant les multiplicateurs de Lagrange.	84

- 4.14 Histogrammes de contrôle du traitement de la superposition des collections. La première ligne comporte les informations relatives à la collection d'électrons et les deux dernières relèvent des jets étiquetés b et non étiquetés, respectivement. La colonne de gauche représente la distribution de la différence en énergie entre un objet reconstruit et son association Monte-Carlo, avant (en bleu) et après (en vert) nettoyage. Dans la colonne de droite est reportée la distance ΔR entre un objet reconstruit et son association Monte-Carlo avant (en rouge) et après (en jaune) nettoyage. 86
- 4.15 Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec l'ajustement par MINUIT. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4. Ces rapports ont été obtenus à partir de l'échantillon $t\bar{t}$ 4100 simulé *via Geant4* et reconstruit avec la reconstruction complète à l'occasion de l'Atelier de Physique de Rome [30]. 87
- 4.16 Masse invariante de la paire jj avant (en rouge à gauche) et après (en bleu à droite) application de la paramétrisation de la fonction d'étalonnage. Les valeurs obtenues lors de l'ajustement par les gaussiennes sont données dans les equations 4.28 et 4.29. 88
- 4.17 Distribution de l'impulsion transverse du photon de plus haute énergie transverse avant coupure HLT. Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus. La coupure appliquée par le menu $\gamma 60i$ est représentée par le trait plein vert vertical. Nous pouvons noter l'effet de la génération du bruit de fond avec $p_T > 50 GeV$ 91
- 4.18 Multiplicités des jets (toutes saveurs confondues) après coupure HLT. Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus. 91
- 4.19 Distribution de l'impulsion transverse du jet de plus haute énergie transverse après coupure HLT et pour différentes efficacités d'étiquetage de jets de b ($\epsilon_b = 50\%$ en haut à gauche, 60% en haut à droite et 70% en bas à gauche). Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus. 92
- 4.20 Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 50% 93

4.21	Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 60%.	94
4.22	Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 70%.	95
4.23	Distributions de l'impulsion transverse des quarks provenant du Z^0 en haut à droite et du rayon du cône reliant les quarks b issus du boson Z^0 en haut à droite. Différence entre les impulsions transverses du Z^0 et du photon et leur superposition, respectivement en bas à droite et en bas à gauche. . .	96
4.24	Distributions de l'impulsion transverse des quarks issus du boson Z^0 lorsqu'une association est trouvée (en traits pleins rouges) et lorsqu'aucune association n'a été trouvée. Nous pouvons noter l'effet de la coupure sur les jets ($p_T > 40 GeV$) qui dégrade la probabilité d'association.	98
4.25	Différence entre les angles azimutaux du quark b et son compagnon pour le Monte-Carlo (vert), les jets associés aux quarks (bleu) et les jets appariés par les différentes stratégies de reconstruction bl (rouge).	99
4.26	Distribution du cosinus de l'angle de la paire de jets avec l'axe faisceau dans le référentiel du Z^0 reconstruit (en haut à gauche). Le signal est en traits pleins rouges et le bruit de fond en pointillés bleus. La figure en haut à droite (en bas à gauche) représente la distributions du cosinus de l'angle entre la direction de la poussée du Z^0 reconstruit avec le jet de plus haute impulsion transverse (de plus petit p_T). Les aires hachurées correspondent à l'application de la coupure $ \cos \theta^* < 0.8$	100

Liste des tableaux

1.1	Les différentes familles de fermions. Seules les particules sont représentées (il faut donc aussi compter avec les anti-particules).	5
2.1	Les principales caractéristique du grand collisionneur de hadrons (LHC) en mode proton-proton.	18
2.2	Performances prévues pour les différents sous-détecteurs d'ATLAS et leur couverture en pseudo-rapidité η , définie comme $\eta = -\ln(\tan \frac{\theta}{2})$ où θ est l'angle polaire défini dans le repère du laboratoire.	23
2.3	Tableau des principales signatures pour la sélection de niveau 1. [18]. Les menus sont de la forme $NoXXI$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet physique ($\mu, e/\gamma, \dots$) sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et I indique s'il faut ou non ajouter un critère d'isolation. XE représente l'énergie manquante dans le plan transverse. Notons qu'environ 20% de la bande passante est allouée pour les déclenchements de type étalonnage et autres. Le total de la bande passante inclut un facteur de sécurité effectif de 4 afin de pallier les incertitudes pouvant affecter les taux de déclenchement au niveau 1.	34
2.4	Tableau des menus du système de déclenchement de dernier niveau (<i>High Level Trigger : HLT</i>) [18]. Les menus sont de la forme $NoXXi$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et i indique s'il faut ou non ajouter un critère d'isolation. xE représente l'énergie manquante dans le plan transverse.	35
4.1	Tableau des menus du système de déclenchement de dernier niveau (<i>High Level Trigger : HLT</i>). Les menus sont de la forme $NoXXi$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et i indique s'il faut ou non ajouter un critère d'isolation. xE représente l'énergie manquante dans le plan transverse.	67
4.2	Tableau des efficacités séquentielles des différentes coupures de présélection pour les événements $t\bar{t}$ générés avec <code>AtlFast</code> . La dernière colonne donne le nombre d'événements attendus pour un an de prise de données à basse luminosité. Chaque ligne donne donc le nombre d'événements passant la totalité des coupures pour une étape donnée de la présélection.	67

4.3	Tableau des sections efficaces de production et des efficacités totales de sélection pour le signal et les différents bruits de fond. Tableau extrait de [75].	67
4.4	Pollutions des bruits de fonds de types I et II (respectivement : jet hors sélection et meilleure combinatoire) pour les événements satisfaisant la coupure Sel4 ou la coupure Sel5. La pureté en W pour chacune de ces classes d'événements est également reportée.	69
4.5	Sections efficaces obtenues pour la production de bosons $Z^0 \rightarrow b\bar{b}$. Chacun des processus a été généré avec Pythia 6.203 en demandant un $p_{\hat{T}} > 20 \text{ GeV}$ et au moins un Z^0 se désintégrant en une paire de quarks b	88
4.6	Tableau des sections efficaces des bruits de fond considérés et du signal ($Z(b\bar{b}) + \gamma$). Le nombre d'événements attendus pour un an de prise de données à basse luminosité est également reporté.	89
4.7	Efficacités de sélection des événements pour le bruit de fond et le signal. Les coupures sont appliquées de manière séquentielle : chaque ligne correspond au pourcentage d'événements survivant après la coupure par rapport au nombre total d'événements analysés. La première ligne correspond à l'application du déclenchement HLT, les deuxième et troisième lignes représentent les coupures " <i>2 jets dont au moins un étiqueté b</i> " et " <i>2 jets dont l'impulsion transverse est d'au moins 40 GeV/c</i> " respectivement. La ligne <i>Sel 2'</i> correspond à la coupure " <i>au moins 1 jet d'impulsion transverse d'au moins 40 GeV/c</i> ". Les deux dernières lignes correspondent aux événements pour lesquels la masse du boson candidat reconstruit est inférieure à $200 \text{ GeV}/c^2$ et comprise dans l'intervalle $[60 \text{ GeV}/c^2; 110 \text{ GeV}/c^2]$, respectivement. Le nombre d'événements retenus après chaque étape de la sélection a été normalisé au nombre d'événements attendus pour une luminosité $\mathcal{L} = 10 \text{ fb}^{-1}$	90
4.8	Valeurs des significances S/\sqrt{B} pour les efficacités d'étiquetage des jets issus de b pour les coupures $m_{bb} < 200 \text{ GeV}/c^2$ et $60 \text{ GeV}/c^2 \leq m_{bb} < 110 \text{ GeV}/c^2$	93
4.9	Efficacités d'association pour les différentes efficacités d'étiquetage des b pour le signal. La première et la deuxième ligne représentent le nombre d'associations quark-jet dans les événements signaux. Les deux lignes suivantes donnent la répartition de la saveur des jets associés aux quarks parmi les événements pour lesquels deux associations ont pu être trouvées. Les deux dernières lignes donnent enfin le pourcentage d'événements parmi les bl pour lesquels le jet léger associé est le jet de plus haute impulsion transverse ou pas.	97
4.10	Tableau des efficacités de reconstruction du Z^0 pour le signal et le bruit de fond en fonction de l'efficacité d'étiquetage des jets de b . Ces efficacités correspondent au nombre d'événements compris dans la fenêtre en masse $[60; 110]$ et pour les classes d'événements $2b$ ou bl . La dernière colonne représente la signification correspondant aux nombres d'événements attendus dans la fenêtre en masse pour 10 fb^{-1}	98

Introduction

Le Modèle Standard est le cadre théorique qui permet de décrire la physique des particules. Confirmé par de nombreux tests expérimentaux, il n'a encore jamais été mis en défaut depuis sa formulation. Cependant plusieurs zones d'ombres subsistent, notamment le mécanisme de brisure de la symétrie électrofaible qui est indispensable à la génération de la masse des particules. De nombreuses théories au-delà du Modèle Standard ont vu le jour pour proposer des réponses aux questions en suspens. La plupart de ces théories prédisent l'apparition de phénomènes nouveaux à l'échelle d'énergie du TeV .

La nouvelle génération de collisionneurs de particules, tel que le Grand Collisionneur de Hadrons (LHC), a pour objectif d'explorer le domaine des énergies au-delà du TeV . Cette échelle d'énergie correspond en effet à la limite au-delà de laquelle le Modèle Standard n'est plus renormalisable. Grâce à la grande énergie disponible dans le centre de masse ($14 TeV$) ainsi que la très haute luminosité ($10^{34} cm^{-2} s^{-1}$), le LHC possède un important potentiel de découverte qui s'étend des mesures de précision des paramètres laissés libres par le Modèle Standard (masse du quark top) jusqu'à la recherche de nouvelles particules.

La très haute luminosité du LHC, nécessaire pour mesurer les processus à très faible section efficace de production, a cependant l'inconvénient de générer une très grande quantité de données qu'il faut ensuite être en mesure de traiter. Les collaborations de physiciens qui construisent les détecteurs qui seront installés auprès du LHC, ont donc dû développer, tester et mettre en production une plateforme logicielle compatible avec le volume de données attendu. Ce passage à l'échelle est un tour de force en soi, puisque plusieurs ordres de grandeur séparent le point de fonctionnement de cet environnement logiciel de celui des expériences passées ou présentes.

Le détecteur ATLAS est l'une des quatre expériences qui fonctionnera auprès du LHC. Les performances attendues de ce détecteur devraient permettre un enregistrement le plus fidèle possible des collisions de protons ayant lieu dans l'accélérateur. La fidélité de la reproduction des collisions dépend de la qualité et finesse des mesures, mais également de l'étalonnage précis des détecteurs. Le travail présenté dans ce mémoire s'inscrit dans ce contexte : il porte d'une part sur le développement d'outils logiciels permettant l'analyse des données recueillies et d'autre part sur l'étalonnage de l'échelle en énergie des jets.

Les deux premiers chapitres introduiront le cadre théorique et le dispositif expérimental du détecteur ATLAS. Le Modèle Standard sera présenté dans un premier temps. Ceci permettra d'extraire les principaux objectifs de physique du LHC, répondre aux questions encore ouvertes posées par le Modèle Standard, et d'en déduire les performances nécessaires afin de les atteindre. Une attention toute particulière sera donnée au quark top dont la connaissance précise est une porte d'entrée vers d'éventuels écarts aux prédictions

du Modèle Standard. Enfin le Grand Collisionneur de Hadrons sera présenté puis une description des différents éléments constituant le détecteur ATLAS sera donnée, mettant en lumière la figure de mérite évoquée au chapitre précédent.

Le troisième chapitre est consacré à la plateforme logicielle en charge de la simulation, de la reconstruction et de l'analyse des données. Les outils développés dans le cadre de ce mémoire et mis à disposition de l'ensemble de la collaboration seront présentés.

Enfin le quatrième chapitre est dédié à l'étalonnage de l'échelle en énergie des jets dans ATLAS. Il s'intéresse dans un premier temps à l'étalonnage *in situ* des jets légers en utilisant des processus de physique. Puis une étude de faisabilité de l'extraction de l'échelle en énergie des jets de b en appliquant la méthode présentée pour les jets légers sera exposée.

Chapitre 1

Le quark *top* et le Modèle Standard

Ce chapitre est dédié à la description du cadre théorique sur lequel repose le champ de la physique qui se consacre à l'étude des constituants fondamentaux de la matière ainsi que des lois la régissant : la physique des particules.

Nous donnerons un bref historique des différentes étapes (théoriques et expérimentales) ayant donné naissance au Modèle Standard. Nous aborderons ensuite la problématique des mesures de précision dans le cadre du Modèle Standard pour conclure sur la physique du quark *top* avec, en toile de fond, la Nouvelle Physique.

1.1 Éléments du Modèle Standard

À l'heure actuelle, notre connaissance de la matière et des interactions régissant celle-ci à l'échelle de l'infiniment petit, est décrite par un modèle appelé Modèle Standard. Il est la synthèse des théories développées pour décrire les particules fondamentales : les particules formant la matière et celles véhiculant les interactions. Ce modèle repose sur deux piliers : la théorie électrofaible et la chromodynamique quantique. L'interaction gravitationnelle n'est en effet pas prise en compte dans la description de la nature que donne le Modèle Standard. Dans ce modèle basé sur la théorie quantique des champs, chaque interaction dérive d'une invariance de jauge locale ¹ du Lagrangien la décrivant. Les médiateurs des interactions sont les générateurs des groupes de symétrie : ce sont les bosons de jauge, de spin entier. Les constituants de la matière sont appelés fermions, de spin demi-entier.

Le succès du Modèle Standard repose sur des prédictions confirmées par une large gamme de tests expérimentaux [1]. Les expériences de physique des particules ont ainsi pour buts de continuer à vérifier l'exactitude des prédictions du Modèle Standard, comme l'existence du boson de Higgs, de s'assurer de sa cohérence et de rechercher des signatures non-prédites par le Modèle Standard. Voyons tout d'abord comment les physiciens sont arrivés à l'élaboration d'un modèle aussi prédictif.

¹C'est à dire l'invariance par rapport aux transformations locales du groupe de symétrie associé à l'interaction.

1.1.1 La quête des constituants élémentaires de la matière

La matière ordinaire qui nous entoure peut être décrite à l'aide de seulement trois particules différentes : le proton p , l'électron e^- et le neutron n . Ils forment une structure appelée atome, introduite par Dalton au début du XIX^e siècle. Ce concept d'atome permit de classer les différents éléments (*via* la table de Mendeleïev). Cependant, bien que servant de guide pour la prédiction de l'existence de nouveaux éléments, l'atome ne permettait pas de véritablement comprendre la matière. Ce ne fut qu'avec la découverte de l'électron (1897), du proton (1911) et du neutron (1932), que cette compréhension put émerger. Les propriétés physico-chimiques des différents éléments de la table de Mendeleïev peuvent en effet être déduites à partir de l'arrangement de ces trois particules.

Cependant le modèle de l'atome n'est pas satisfaisant. En effet, en plus d'être instable ², il ne permet pas d'expliquer le comportement étrange de certains atomes. En 1896, Becquerel découvre que des sels d'Uranium émettent un rayonnement capable d'ioniser l'air et de sensibiliser une plaque photographique. Pierre et Marie Curie mettent ensuite en évidence que ce comportement peut être obtenu avec d'autres corps : c'est la radio-activité. En 1930, l'étude de la radio-activité β permet à Pauli de postuler l'existence d'une nouvelle particule, le neutrino ν , qui fut ensuite découverte en 1956. L'étude des rayons cosmiques permit en 1937 de mettre en évidence le muon μ , particule présentant les mêmes propriétés que l'électron, à l'exception de sa masse ($m_\mu \sim 200 \times m_e$). En 1947, la spectroscopie des particules s'élargit avec la découverte dans les rayons cosmiques de plusieurs nouvelles particules, des mésons de spin entier et des baryons de spin demi-entier. Ces nouvelles particules sont alors nommées particules étranges.

Gell-Mann, Nee'man et Zweig postulent en 1964 l'existence d'une structure composite aux hadrons : les quarks. Ces nouvelles particules élémentaires seraient la représentation fondamentale du groupe $SU(3)_S$ de saveur. Les trois quarks u (*Up*), d (*Down*) et s (*Strange*), et les anti-quarks \bar{u} , \bar{d} et \bar{s} , permettent la construction de tous les hadrons observés. Au début considérés comme un habile artifice mathématique, les quarks furent ensuite mis en évidence de manière indirecte dans la fin des années 60.

Sur des considérations théoriques, Glashow et Bjørken puis Glashow, Illiopoulos et Maiani postulent l'existence d'un quatrième quark, le quark c (*Charm*). Il est mis en évidence en 1974. Avec la découverte en 1962 à Brookhaven de l'existence d'un second neutrino, le ν_μ , compagnon du lepton μ , il apparaît que la Nature ait choisi de créer deux familles de quarks et de leptons ayant des caractéristiques et des propriétés similaires. Seules les masses diffèrent.

En 1975, la découverte d'un nouveau lepton appelé tau (noté τ) possédant une masse supérieure à celle du muon, permet de postuler l'existence d'une troisième famille de leptons. Pour compléter le modèle et expliquer la violation de la symétrie CP , l'existence d'une troisième famille de quarks est postulée. Les composants de cette famille additionnelle, le quark b (*Bottom* ou *Beauty*) et le quark t (*Top* ou *Truth*), sont découverts à Fermilab en 1977 et 1994, respectivement. Enfin, le neutrino de la troisième génération a été mis en évidence en 2000 par la collaboration DONuT à Fermilab [2].

L'ensemble des particules connues à ce jour et constituant la matière, est reporté dans le tableau 1.1. Les 12 particules de matière que compte le Modèle Standard, sont réparties

²Dans le cadre de l'électromagnétisme classique, les électrons "gravitent" autour du noyau devraient en effet entrer en collision avec le noyau, sous l'effet de l'attraction électrique.

Famille	Leptons	Quarks
1	$e_L, e_R, (\nu_e)_L$	d_L, d_R, u_L, u_R
2	$\mu_L, \mu_R, (\nu_\mu)_L$	s_L, s_R, c_L, c_R
3	$\tau_L, \tau_R, (\nu_\tau)_L$	b_L, b_R, t_L, t_R

TAB. 1.1 – Les différentes familles de fermions. Seules les particules sont représentées (il faut donc aussi compter avec les anti-particules).

en 3 familles de 2 leptons et 2 quarks. Nous pouvons noter une première limitation du Modèle Standard. Même s'il permet de décrire correctement le comportement de ces particules de spin demi-entier (appelées fermions), le Modèle Standard ne fournit aucune explication quant au nombre de familles, ni à la disparité des masses entre les différentes familles de quarks et de leptons.

En théorie quantique des champs, les fermions sont représentés par des spineurs de Dirac. Les fermions peuvent alors être décomposés en particules de chiralité droite (*right*, R) ou gauche (*left*, L). Expérimentalement, seuls les neutrinos gauches ont pu être observés, au contraire des autres fermions pour qui les deux composantes existent.

Nous avons donné le contenu en matière du Modèle Standard. Voyons maintenant comment le Modèle Standard explique les phénomènes physiques observés à ce jour mettant en relation ces particules de matière.

1.1.2 Interactions fondamentales

A l'heure actuelle, tous les phénomènes de physique sont interprétés comme la manifestation de quatre interactions fondamentales : la gravitation, l'interaction électromagnétique, l'interaction nucléaire faible et l'interaction nucléaire forte. Le Modèle Standard permet d'interpréter les trois dernières interactions fondamentales dans un même cadre théorique. L'incorporation de la gravitation à la physique des particules reste en effet un problème de physique théorique qui n'est pas encore résolu. Ce n'est cependant pas un problème numérique pour la description de la physique des particules dans le cadre du Modèle Standard. En effet, l'intensité de la gravitation aux énergies accessibles en physique des particules expérimentale est négligeable.

Le Modèle Standard est une théorie de champ quantique. Dans ce cadre théorique, chaque force fondamentale est le résultat de l'interaction entre les particules de matière, les fermions, et les particules médiatrices de l'interaction, les bosons de jauge. Le boson d'échange de l'interaction électromagnétique est le photon. Les bosons W^\pm et Z^0 sont les particules échangées par interactions faibles, responsable de la désintégration β . Huit gluons sont nécessaires pour décrire l'interaction forte, responsable de la cohésion des nucléons. L'invariance de jauge est la base de ce modèle théorique. Cette invariance requiert que la théorie demeure invariante sous l'application en chaque point de l'espace-temps d'une transformation : la transformation de jauge.

1.1.2.1 Interactions électrofaibles

Les interactions électrofaibles sont le résultat de quarante années de recherches théoriques et expérimentales. Ces recherches ont abouti à l'unification des interactions élec-

tromagnétique et nucléaires faibles. Dans les années 40, Feynman, Schwinger et Tomonaga formulent la première théorie de champ quantique, l'Electrodynamique Quantique (*Quantum Electrodynamics*, QED). QED est la théorie quantique des interactions électromagnétiques. Dans ce cadre théorique, tout phénomène électromagnétique résulte de l'échange de photons entre des fermions portant une charge électrique. De part son succès, QED devint le modèle de toute théorie de champ quantique.

Se calquant sur QED, Fermi élabore un modèle des interactions faibles pour décrire la désintégration β . Dans le modèle de Fermi, la désintégration β est l'interaction ponctuelle couplant 4 fermions. La désintégration β est alors la manifestation de la désintégration d'un neutron en 3 particules : un proton, un électron et un anti-neutrino électronique.

$$n \rightarrow p + e^- + \bar{\nu}_e \quad (1.1)$$

Cependant, contrairement à QED, ce modèle n'est pas renormalisable. En 1957, la mise en évidence de la violation de parité nécessite la reformulation du modèle de Fermi pour prendre en compte l'aspect chiral des interactions faibles. Ce travail est réalisé en 1958 par Sakurai, Gell-Mann et Feynman. Le modèle de Fermi est alors modifié en une théorie $V - A$ (couplage vecteur, couplage axial-vecteur).

Il faut attendre 1961 pour que Glashow jette les bases de la première tentative d'unification entre les interactions électromagnétique et faibles. Le cadre théorique de cette théorie électrofaible repose sur le groupe de jauge :

$$SU(2)_L \otimes U(1)_Y \quad (1.2)$$

La théorie électrofaible de Glashow contient en fait 2 interactions :

- l'interaction d'isospin faible, basée sur le groupe non-abélien $SU(2)_L$,
- l'interaction d'hypercharge faible, basée sur le groupe $U(1)_Y$.

L'application de la relation de Gell-Mann Nishijima permet de relier l'isospin faible et l'hypercharge faible à la charge électrique :

$$Q = T_3 + \frac{Y}{3} \quad (1.3)$$

Ainsi, tous les fermions portent une hypercharge faible (et subissent donc l'interaction d'hypercharge faible) alors que seuls les fermions de chiralité gauche sont sensibles à l'interaction d'isospin faible. Un problème majeur résulte de cette tentative d'unification : la symétrie exacte du Lagrangien électrofaible par rapport à $SU(2)_L \otimes U(1)_Y$ implique des masses nulles pour les bosons médiateurs de l'interaction, ainsi que pour les fermions. Or, la courte portée de l'interaction faible implique la présence de bosons massifs. Il faut donc introduire un mécanisme capable de donner une masse aux bosons de l'interaction faible, tout en préservant une masse nulle pour le photon.

L'introduction du mécanisme de Higgs dans le modèle de Glashow, permet à Weinberg et Salam de donner une masse aux particules sans briser l'invariance de jauge. En 1967, le modèle GSW pose les premiers éléments de ce qui deviendra le Modèle Standard. Enfin, en 1970, le modèle des interactions électrofaibles est achevé par Glashow, Iliopoulos et Maiani qui intègrent les quarks au cadre théorique. Ce modèle permet en outre de prédire l'existence de l'interaction faible par courant neutre. Les courants neutres sont ensuite découverts en 1973 au CERN. La validation complète de ce modèle vient finalement avec la découverte des bosons W de l'interaction faible en 1983 au CERN.

Nous pouvons tout d'abord remarquer que le modèle de Fermi est une théorie effective du modèle GSW. Ceci implique qu'à basse énergie, les deux modèles doivent être identiques. Ceci permet de relier la constante de couplage de Fermi G_F à la constante de couplage d'isospin faible :

$$\frac{G_F}{\sqrt{2}} = \frac{g_2^2}{8m_W^2} \quad (1.4)$$

D'autre part, nous pouvons relever un point clef de la théorie électrofaible : l'origine de la masse des particules. Le mécanisme introduit par Higgs, une fois utilisé pour palier le problème de la portée de l'interaction faible, nécessite l'introduction d'un doublet de champs scalaires complexes de $SU(2)_L$:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} \quad (1.5)$$

L'isospin de ce doublet est $T = \frac{1}{2}$ et son hypercharge $Y = 1$, ce qui impose les charges des composantes. Le Lagrangien décrivant l'évolution de ce doublet peut s'écrire sous la forme :

$$\mathcal{L}_{scalaire} = (D_\mu \Phi)(D^\mu \Phi^*) - V(\Phi) \quad (1.6)$$

où $V(\Phi)$ est un terme de potentiel de la forme :

$$V(\Phi) = \frac{1}{2}\mu^2(\Phi^\dagger \Phi) + \frac{\lambda}{4}(\Phi^\dagger \Phi)^2 \quad (1.7)$$

où μ et λ sont deux constantes arbitraires ($\lambda \geq 0$). Au minimum du potentiel, nous avons :

$$\Phi^\dagger \Phi = -\frac{\mu^2}{\lambda} = v^2 \quad (1.8)$$

La valeur moyenne de ce champ dans le vide doit minimiser la densité d'énergie et correspondre à un minimum stable. Le signe du paramètre μ^2 gouverne la forme du potentiel satisfaisant ces contraintes. Pour $\mu^2 > 0$, le minimum de potentiel est unique et trivial. Il est à champ nul et laisse la symétrie invariante. Cette solution ne permet cependant pas d'attribuer une masse aux bosons de l'interaction faible.

Pour $\mu^2 < 0$, le potentiel n'a plus de minimum stable en zéro. Le minimum stable est à champ non-nul et brise spontanément la symétrie. Le choix de Higgs a été de prendre :

$$\langle 0|\Phi|0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.9)$$

Le paramètre v est la valeur moyenne dans le vide, non-nulle lorsque la symétrie est brisée. Ainsi, en utilisant l'invariance de jauge sous une transformation locale, nous pouvons reparamétriser le champ Φ sous la forme :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (1.10)$$

où H est le champ du boson de Higgs physique. D'autre part, les champs physiques associés aux bosons vecteurs états propres de masse Z^0 , W^\pm et γ s'expriment comme des combinaisons linéaires des champs de jauge W_μ^i et B_μ :

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} [W_\mu^1 \mp iW_\mu^2] \\ Z_\mu &= \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu \\ A_\mu &= \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \end{aligned} \quad (1.11)$$

ce qui permet d'établir la relation entre charge électrique e et constantes de couplage g_2 d'isospin faible et g_1 d'hypercharge faible :

$$e = g_2 \sin \theta_W = g_1 \cos \theta_W \quad (1.12)$$

caractérisant l'unification des interactions faibles et électromagnétique. Finalement, en injectant l'expression du champ de Higgs dans le Lagrangien électrofaible, nous pouvons exprimer les masses des bosons vecteurs états propres de masse par :

$$\begin{aligned} m_W &= \frac{1}{2} v g_2 \\ m_Z &= \frac{1}{2} v \sqrt{g_2^2 + g_1^2} \end{aligned} \quad (1.13)$$

Trois des quatre degrés de liberté introduits avec le champ défini dans l'équation 1.5 seront absorbés par les bosons Z^0 et W^\pm leur conférant une composante longitudinale et de fait, une masse. Le degré de liberté restant suggère l'existence d'une nouvelle particule scalaire fondamentale, le boson de Higgs. Ainsi, au prix de deux nouveaux paramètres libres (λ et μ) et d'une nouvelle particule H , le mécanisme de Higgs attribue une masse aux bosons de l'interaction faible et aux autres particules du Modèle Standard. De plus, le modèle permet de relier, sans prédire leur valeur, les masses de ces bosons :

$$m_Z = \frac{m_W}{\cos \theta_W} \quad (1.14)$$

Enfin, à partir de la masse du boson W et connaissant la valeur de la constante de Fermi, nous pouvons estimer la valeur de l'échelle de brisure électrofaible :

$$v^2 = 4 \frac{m_W^2}{g_2^2} = 4 m_W^2 \frac{\sqrt{2}}{8 m_W^2 G_F} = \frac{\sqrt{2}}{2 G_F} \sim (250 \text{ GeV}/c^2)^2 \quad (1.15)$$

Cependant, l'origine de ce mécanisme de brisure n'est pas comprise (le doublet de Higgs est introduit "à la main"). De plus, le boson de Higgs n'a toujours pas été découvert à ce jour, bien qu'il soit le garant de la cohérence et de la renormalisabilité du Modèle Standard.

1.1.2.2 Interaction forte

L'interaction forte est la force responsable de la cohésion du noyau. En prenant comme modèle QED, elle a tout d'abord été décrite par Yukawa en 1935 comme l'échange de mésons (les pions) entre les nucléons constituant le noyau. Cependant, avec la mise en évidence d'une sous-structure aux nucléons *via* des expériences de diffusion électron-nucléon profondément inélastique, confirmant le postulat de Gell-Mann, Nee'man et Zweig, il fallait trouver un autre cadre théorique. La découverte du baryon Δ^{++} , composé de trois quarks up , et l'application du principe de Pauli, nécessitaient l'introduction d'un nouveau nombre quantique : la charge de couleur.

Le début des années 70 voit la naissance de la Chromodynamique Quantique (*Quantum Chromodynamics*, QCD) qui est elle-aussi calquée sur QED. Elle est basée sur un groupe de jauge plus complexe, $SU(3)_C$, dont les 8 générateurs sont les 8 matrices λ de Gell-Mann, associées aux 8 bosons de jauge de masse nulle, les gluons. Le groupe de jauge $SU(3)_C$ permet de classer les particules en fonction de leur charge de couleur. Les leptons sont des singulets de $SU(3)_C$, ils ne sont ainsi pas sensibles à l'interaction forte. Au

contraire, les quarks sont des triplets : ils possèdent une charge de couleur et sont donc soumis à l'interaction forte. Le postulat de base de QCD est que chaque quark existe en trois couleurs différentes, rouge (\mathcal{R}), vert (\mathcal{V}) et bleu (\mathcal{B}). Les anti-quarks portent alors les anti-couleurs correspondantes. Enfin, tous les états physiques observables (les hadrons) doivent être des singulets de couleur ("blancs"). Les quarks interagissent entre eux par interaction forte en s'échangeant leur couleur *via* les gluons, eux-mêmes porteurs de couleurs reflétant ainsi le caractère non-abélien du groupe de jauge $SU(3)_C$.

Le postulat de base de QCD, stipulant que les états physiques observables sont de couleur blanche, conditionne le confinement des quarks à l'intérieur des hadrons. La raison de ce confinement provient des interactions gluons-gluons. En effet, l'écrantage de la charge électrique à grande distance qui avait lieu dans QED, devient un anti-écrantage. De fait, la charge de couleur augmente avec la distance : plus les quarks sont éloignés et plus ils ressentent l'interaction forte de manière intense, ce qui tend à les confiner à l'intérieur des hadrons. *A contrario*, à la limite d'une distance nulle, les quarks sont totalement libres : ce phénomène est appelé liberté asymptotique. La constante de couplage de l'interaction forte peut alors s'écrire en fonction de l'échelle en énergie Q considérée :

$$\alpha_S(Q^2) = \frac{12\pi}{(33 - 2n_f) \ln\left(\frac{Q^2}{\Lambda_{QCD}^2}\right)} \quad (1.16)$$

où n_f est le nombre de saveurs de quarks à l'échelle en énergie considérée, et Λ_{QCD}^2 est l'échelle en énergie de QCD ($\sim 200 \text{ MeV}$). Pour $n_f < 17$, α_S tend vers zéro lorsque Q tend vers l'infini, ce qui reproduit bien le caractère de liberté asymptotique à distance nulle.

Après avoir introduit les différentes interactions fondamentales prises en compte dans le cadre du Modèle Standard, nous allons nous intéresser aux mesures de précision effectuées afin de contraindre plus fortement le Modèle Standard.

1.2 Mesures de précision

Le Modèle Standard des interactions forte et électrofaibles permet de décrire le comportement de la matière et de trois des interactions connues à ce jour. Le groupe de jauge associé au Modèle Standard est $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, qui correspond au produit direct des groupes de jauge que nous avons introduits précédemment.

Malgré les nombreux succès expérimentaux de ce modèle, il semble très probable qu'il ne soit pas la théorie fondamentale mais simplement une excellente approximation à basse énergie d'une théorie plus fondamentale. En effet, la première limite de ce modèle apparaît tout d'abord dans le fait que la gravitation ne soit pas prise en compte. De plus, si le principe de brisure de symétrie électrofaible est maintenant bien établi, le mécanisme à son origine est encore mal compris : l'ingrédient fondamental utilisé dans le Modèle Standard pour cette brisure, le boson de Higgs, reste encore à découvrir. Enfin, le Modèle Standard n'a aucun pouvoir prédictif quant aux valeurs numériques de plusieurs paramètres (*ex* : les masses des particules fondamentales, les valeurs des constantes de couplage des différentes interactions, ...).

Ce faisceau d'indices tend à nous conforter dans l'idée que le Modèle Standard soit une théorie effective, c'est à dire la réalisation à basse énergie d'une théorie plus fondamentale,

une théorie au-delà du Modèle Standard. Plusieurs cadres théoriques ont été développés et il est donc important d'effectuer des mesures de précision des paramètres du Modèle Standard. En effet, ces mesures permettent d'appliquer davantage de contraintes sur ces paramètres afin d'éventuellement révéler le ou les secteurs dans lesquels des écarts à la théorie apparaissent. La cohérence du Modèle Standard est testée en mesurant les 18 paramètres laissés libres par ce dernier :

- les 3 constantes de couplages g_1 , g_2 et g_3 ,
- les 2 paramètres du potentiel du champ de Higgs μ et λ ,
- les 9 couplages de Yukawa qui contribuent à établir une masse pour les fermions, et
- les 4 paramètres indépendants de la matrice de Cabibbo-Kobayashi-Maskawa qui établissent le mélange entre les états propres de saveur et de masse pour les quarks.

Cependant, les cinq premiers paramètres ne correspondent pas à des quantités physiques directement mesurables. Les paramètres mesurés sont alors :

- la constante de couplage électromagnétique $\alpha = \frac{e^2}{4\pi}$, e étant reliée à g_1 et g_2 ,
- la masse du boson Z^0 ,
- la constante de Fermi G_F ,
- la masse du boson de Higgs m_H ,
- la constante de couplage forte $\alpha_S = \frac{g_3^2}{4\pi}$

Ces paramètres sont accessibles à l'expérience et font ainsi l'objet de mesures de précision (sauf m_H , la masse du boson de Higgs). Bien que tous les paramètres du Modèle Standard ne puissent être mesurés, il est néanmoins possible d'effectuer un test de cohérence de ce dernier. En effet, les observables du Modèle Standard sont reliées entre elles *via* les corrections radiatives. La figure 1.1 (tirée de [1]) représente l'écart entre les valeurs mesurées et les valeurs prédites dans le cadre du Modèle Standard. Nous pouvons noter que l'écart le plus important correspond à l'asymétrie avant-arrière pour les quarks b . La probabilité de χ^2 relative à l'ajustement des paramètres du Modèle Standard est de $\sim 15\%$ pour $\chi^2/ndf \sim 18/13$.

La comparaison entre les valeurs mesurées et celles obtenues dans le cadre du Modèle Standard permet de repérer le jeu de paramètres le moins bien décrit dans le cadre du Modèle Standard. La vérification de la cohérence entre mesures directes et indirectes constitue ainsi un test de la validité du Modèle Standard. Les écarts observés entre les deux mesures peuvent alors être attribués soit à des contributions des ordres supérieurs non inclus lors du calcul théorique, soit à de nouveaux phénomènes non prédits par le Modèle Standard. Menées de concert avec une meilleure connaissance des incertitudes théoriques, les mesures de précisions permettent de lever ces ambiguïtés.

1.3 Le quark *top*

Le quark *top* est le sixième quark à avoir été découvert, en 1994, à Fermilab, auprès du TeVatron. Le quark *top* complète ainsi la troisième génération de quarks. Selon le Modèle Standard, c'est une particule de spin 1/2, charge 2/3 et de partenaire d'isospin faible le quark b . Il se désintègre presque exclusivement en un boson W et un quark b . Le quark *top* est le seul fermion fondamental ayant une masse de l'ordre de l'échelle électrofaible. Il constitue une excellente sonde du secteur électrofaible et de la brisure de symétrie associée. Le quark *top* pourrait être un révélateur de nouvelle physique à travers ses modes de production ou de désintégration.

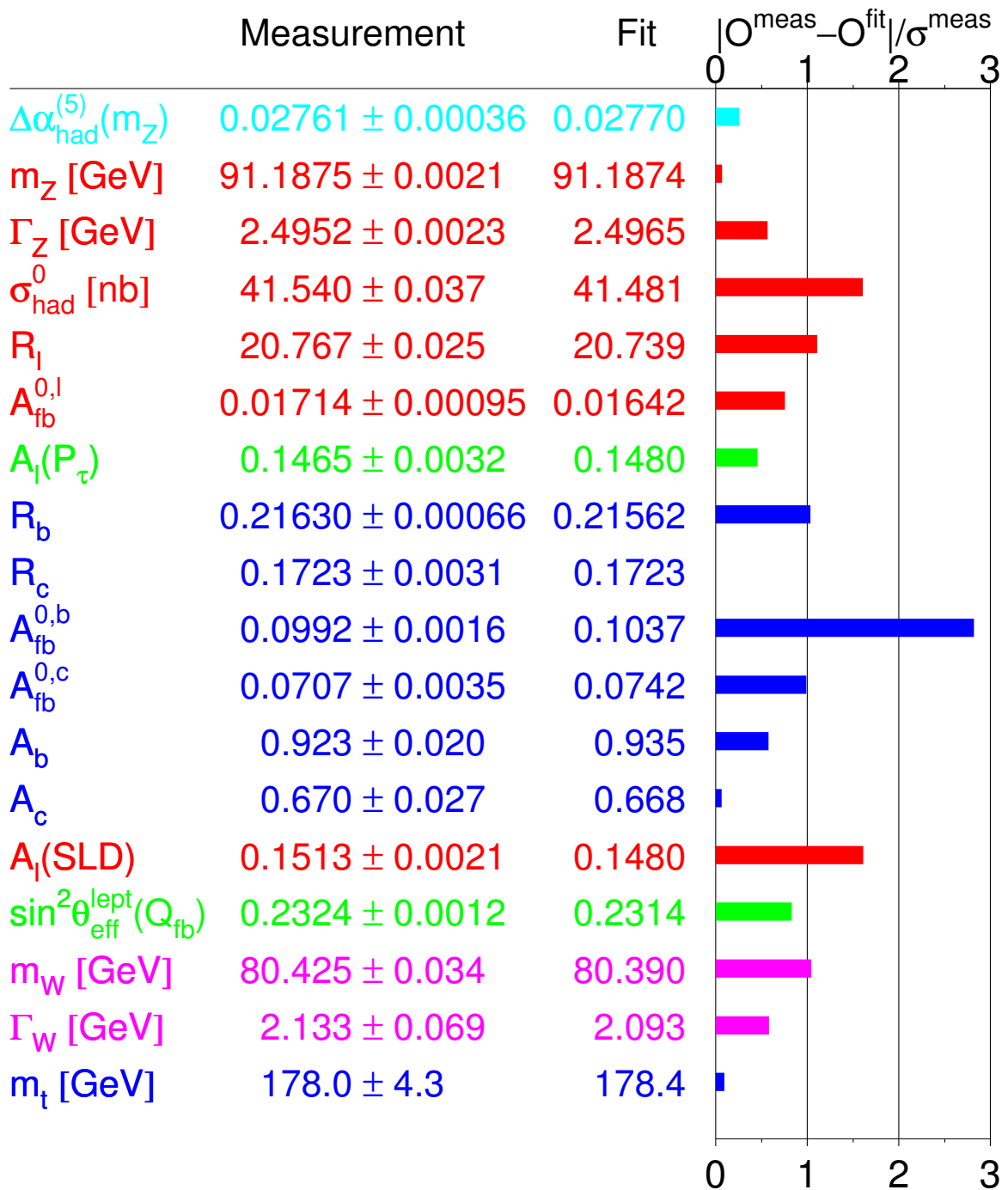


FIG. 1.1 – Ecart entre les valeurs mesurées des observables électrofaibles et les valeurs prédites par le Modèle Standard. Les valeurs préliminaires présentées dans ce tableau sont tirées de [1].

La première partie fera l'objet d'un bref historique de la découverte du quark *top* en mettant l'accent sur le rôle prépondérant que ce dernier joue dans le cadre du Modèle Standard. Nous donnerons ensuite les différentes caractéristiques de ce fermion, puis nous décrirons les différentes méthodes de mesure de la masse du *top*.

1.3.1 Historique

Avec la découverte du lepton τ en 1975 et celle du quark *b* en 1977, l'existence du quark *top* devenait un élément crucial de la cohérence du Modèle Standard. En effet, l'existence d'une troisième famille de quarks est nécessaire afin d'inclure la violation de *CP* observée dans le cadre du Modèle Standard. Ceci provient du fait que le nombre minimal de générations de quarks est 3 afin de garder la matrice de mélange des états propres de masse et des états propres de saveur, complexe.

L'étude des propriétés du quark *b* laissait supposer qu'il appartenait à un doublet d'isospin. L'absence d'observation directe et l'accès à des énergies de plus en plus importantes, permettaient d'imposer des limites inférieures sur la masse du quark *top*. En 1994, la recherche directe du quark *top* permettait ainsi de mettre $131 \text{ GeV}/c^2$ comme limite inférieure. La recherche indirecte permet de prédire la masse de ce quark en mesurant les observables électrofaibles au pôle du Z^0 auprès du collisionneur e^+e^- LEP :

$$m_{top}^{indirecte} = 177 \pm 11_{-19}^{+18} \text{ GeV}/c^2 \quad (1.17)$$

Les expériences CDF et DØ détectèrent les premiers indices de l'existence du quark *top* en 1994 et annoncèrent sa découverte en 1995, dans une gamme de masse en accord avec les prédictions de LEP :

$$\begin{aligned} CDF : m_{top} &= 176 \pm 8 \pm 10 \text{ GeV}/c^2 \\ D0 : m_{top} &= 199_{-21}^{+19} \pm 22 \text{ GeV}/c^2 \end{aligned} \quad (1.18)$$

Ces expériences ont par la suite mesuré plus précisément la masse du *top*, dans différents canaux. En utilisant les données actuellement recueillies pour le Run-II, l'analyse combinée des résultats de ces deux expériences permet de donner comme nouvelle valeur mondiale [3] (préliminaire) pour la masse du *top* :

$$m_{top} = 172.7_{-1.7}^{+1.7} \pm 2.9 \text{ GeV}/c^2 \quad (1.19)$$

avec une erreur systématique de $2.9 \text{ GeV}/c^2$.

1.3.2 Caractéristiques du quark *top*

Le mode de production principal du quark *top* auprès des machines hadroniques est celui faisant intervenir l'interaction forte. Le quark *top* est ainsi produit par paires $t\bar{t}$ puisque l'interaction forte conserve la saveur. Au LHC, l'absence d'anti-quarks de valence et l'abondance de gluons dans le proton, expliquent que les paires soient majoritairement produites par fusion de gluons ($\sim 90\%$). Cette situation est inversée auprès du TeVatron, cette machine mettant en jeu des collisions $p\bar{p}$.

Une fois produit, le quark *top* du Modèle Standard se désintègre presque exclusivement en Wb ($> 99\%$). Du fait de la brièveté de son temps de vie $\tau \sim 10^{-24} \text{ s}$, le quark *top* se désintègre avant de s'hadroniser. Les propriétés (impulsion, spin) de ce quark peuvent alors être directement observées au travers de ses produits de désintégration.

1.3.3 Mesure de la masse

La connaissance précise de la masse du quark *top* est primordiale pour mieux comprendre (et contraindre) la physique électrofaible, la physique de la beauté et la physique du Higgs. En effet, les observables électrofaibles (m_W , $\sin \theta_W^{\text{lept}}$, ...) sont très sensibles à la masse du *top* qui intervient dans le calcul des corrections radiatives des autres observables du Modèle Standard. De fait, nous allons nous intéresser à la précision nécessaire sur la mesure de la masse du *top* pour contraindre efficacement le Modèle Standard, lorsqu'elle sera couplée aux informations d'autres observables.

1.3.3.1 Mesure indirecte

A l'ordre le plus bas, la masse du boson W est donnée par :

$$m_W = \sqrt{\frac{\pi \alpha_{QED}}{\sqrt{2}} \frac{1}{G_F \sin^2 \theta_W}} \quad (1.20)$$

Lorsque les ordres supérieurs sont pris en compte, cette expression devient :

$$m_W = \sqrt{\frac{\pi \alpha_{QED}}{\sqrt{2}} \frac{1}{G_F \sin^2 \theta_W} \left(\frac{1}{1 - \Delta r} \right)} \quad (1.21)$$

où le terme Δr contient les corrections issues des diagrammes aux ordres supérieurs. Ce terme contient les corrections issues du quark *top* ($\propto m_{top}^2$) et les corrections provenant du boson de Higgs ($\propto \log(m_H)$). Ainsi, lorsque l'ajustement des données collectées au pôle du Z^0 et celles des mesures directes de la masse du W et de sa largeur, est effectué [1], nous obtenons la meilleure estimation indirecte de la masse du quark *top* : $m_t^{\text{Fit}} = 179_{-9}^{+12} \text{ GeV}/c^2$ qui est en accord avec la valeur moyenne mondiale.

La figure 1.2 présente la masse du boson W en fonction de la masse du quark *top*. Plusieurs hypothèses de masse pour le boson Higgs sont également reportées. Nous pouvons remarquer que les mesures actuelles privilégient un boson de Higgs léger. Ainsi, la mesure précise des masses du *top* et du W participe à la contrainte sur la masse du boson de Higgs. Cette contrainte teste donc la cohérence interne du Modèle Standard. Si le boson de Higgs venait à être découvert, la connaissance même approximative de sa masse permettrait de grandement réduire la bande de la figure 1.2.

1.3.3.2 Mesure directe

La mesure directe de la masse du quark *top* réalisée auprès du TeVatron par les expériences CDF et DØ est le résultat de la combinaison de plusieurs canaux de désintégration.

Une fois la paire $t\bar{t}$ produite, trois modes de désintégration sont accessibles pour le système $WbWb$:

- le canal dilepton dans lequel les deux bosons W se désintègrent en lepton et neutrino associé,
- le canal semi-leptonique dans lequel un des deux bosons se désintègre en une paire qq' , et
- le canal hadronique dans lequel les deux bosons se désintègrent en une paire qq' .

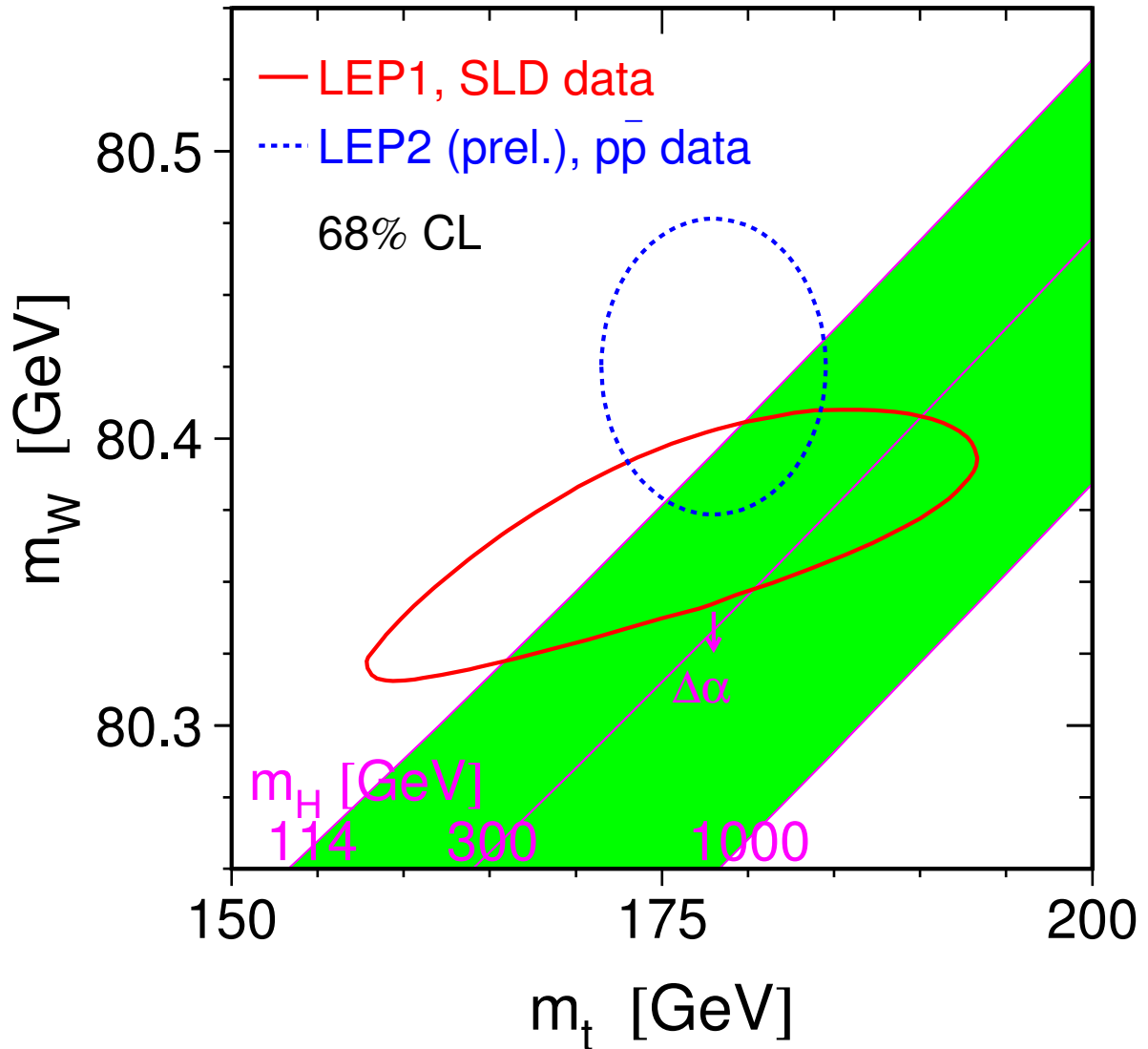


FIG. 1.2 – Masse du boson W en fonction de la masse du quark *top*. Les contours à 68% de niveau de confiance pour les mesures directes et indirectes sont indiquées. Différentes hypothèses pour la masse du boson de Higgs sont superposées. Cette figure est extraite de [1].

La signature du canal dilepton est la présence de deux leptons de charges opposées ainsi qu'une impulsion transverse manquante importante, sans oublier les deux jets issus de la fragmentation des quarks b . La signature de ce type d'événements est très propre mais du fait de la topologie de l'état final, ce mode permet la reconstruction complète d'un seul des deux top (cinématique sous-contrainte). Le mode hadronique ne fait pas intervenir de neutrino dans l'état final et rend donc possible la reconstruction complète des deux top (cinématique sur-contrainte). Cependant l'extraction du signal est compliquée par la présence de bruits de fond QCD. Le mode semi-leptonique permet la reconstruction complète d'un des deux top . Le boson W se désintégrant en une paire lepton-neutrino rend la reconstruction du top correspondant difficile ³.

La stratégie mise en place pour extraire la masse du top est la même, indépendamment du canal étudié :

- sélection des événements à partir des critères les moins sensibles à m_{top} ,
- reconstruction d'une masse par événement en utilisant un ajustement cinématique correspondant à l'hypothèse d'un événement $t\bar{t}$,
- comparaison de distribution de la masse du top reconstruit avec les distributions Monte-Carlo obtenues avec différentes valeurs de masses de top .

La valeur de la masse du top est alors celle obtenue au maximum de vraisemblance.

³Cette reconstruction est cependant possible.

Chapitre 2

Le détecteur ATLAS au LHC

2.1 Le grand collisionneur de hadrons

Le CERN (Organisation Européenne pour la Recherche Nucléaire) est en train de construire le LHC (*Large Hadron Collider*) qui devrait permettre d'étudier la physique des particules dans une gamme d'énergie encore jamais atteinte. Associé à quatre expériences de physique des particules, cet outil devrait permettre par exemple de mettre en évidence le boson de Higgs ou de nouvelles particules jusqu'à des masses de l'ordre du TeV ainsi que contraindre plus fortement le Modèle Standard *via* des mesures de précision des paramètres laissés libres.

Nous donnerons dans la suite une brève description du collisionneur de hadrons du CERN puis nous traiterons des généralités d'un détecteur installé auprès du LHC. Enfin nous décrirons les différents sous-détecteurs constituant le détecteur de la collaboration ATLAS.

2.1.1 L'accélérateur

Le LHC est un collisionneur proton-proton qui est logé dans l'ancien tunnel du LEP (*Large Electron Positron collider*) au CERN à Genève. Il devrait entrer en fonctionnement pendant l'été 2007 et il produira des collisions proton-proton ayant une énergie dans le centre de masse de $\sqrt{s} = 14 TeV$ avec une luminosité nominale de $\mathcal{L} = 10^{34} cm^{-2}.s^{-1}$. Cette luminosité instantanée devrait permettre d'atteindre $100 fb^{-1}$ pour une année de prise de données. Les principales caractéristiques du LHC sont reportées dans le tableau 2.1.

Contrairement au LEP, le LHC est une machine hadronique. Ce mode de fonctionnement est particulièrement bien adapté à la recherche de nouveaux phénomènes¹. En effet, l'énergie emportée par les partons est une fraction variable de l'énergie du proton initial. Ceci permet de couvrir une gamme d'énergie allant de la dizaine de GeV à quelques TeV , sans rien changer au point de fonctionnement de la machine. De plus, les pertes d'énergie par rayonnement synchrotron qui ont lieu dans les accélérateurs circulaires, sont négligeables dans le cas des protons. Ces pertes sont inversement proportionnelles à la

¹Il faut cependant pondérer cette affirmation. Plusieurs facteurs peuvent en effet gêner cette recherche. Nous pouvons noter par exemple, le fond QCD mal connu ou alors des biais dans le système de déclenchement.

Paramètre	Valeur
Circonférence	26659 <i>m</i>
Energie de faisceau	7 <i>TeV</i>
Champ des dipôles à 7 <i>TeV</i>	8.33 <i>T</i>
Luminosité	$10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$
Nombre de paquets/faisceau	2835
Nombre de protons/paquet	10^{11}
Espace entre les paquets	7.48 <i>m</i>
Temps entre deux croisements	24.95 <i>ns</i>
Energie perdue par révolution	7 <i>keV</i>
Extension latérale du faisceau	$\sigma_x = \sigma_y = 16 \text{ } \mu\text{m}$
Extension longitudinale du faisceau	$\sigma_z = 5.4 \text{ cm}$

TAB. 2.1 – Les principales caractéristique du grand collisionneur de hadrons (LHC) en mode proton-proton.

puissance quatrième de la masse de la particule accélérée. Cet effet ne sera alors pas un facteur limitant l'énergie disponible dans le centre de masse comme ce fut le cas pour le LEP.

Cependant, l'utilisation de partons lors des collisions ², couplée à la rareté des phénomènes qui seront étudiés, nécessitent de délivrer une très grande luminosité lors de la période de prise de données. L'utilisation d'un collisionneur $p\bar{p}$ est défavorisée car la technologie actuelle ne permet pas d'obtenir des sources d'antiprotons suffisamment lumineuses. Par contre, le choix du mode proton-proton pour le LHC nécessite de produire un champ magnétique propre à chacun des deux sens de propagation. Le CERN a donc développé des dipôles "2-en-1" dans lesquels deux champs magnétiques antiparallèles de 8.33 *T* délivrés par des bobines supraconductrices, partagent les mêmes cryostats et la même structure métallique pour le retour de champ.

Par ailleurs, la grande luminosité nécessaire aux mesures de phénomènes rares impose des contraintes fortes sur les détecteurs en terme de tenue aux radiations. Les détecteurs sont conçus dans l'optique de pouvoir résister au moins à dix ans de fonctionnement à haute luminosité ($\mathcal{L} = 10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$). Cette luminosité ne sera d'ailleurs atteinte qu'après trois ans de prise de données pendant lesquels la luminosité sera de l'ordre de $10^{33} \text{ cm}^{-2} \cdot \text{s}^{-1}$. De plus, pour atteindre la luminosité requise, il est nécessaire de délivrer des paquets de protons d'une densité importante (10^{11} protons par paquet) ainsi que de réduire l'espacement entre les paquets (25 *ns* entre chaque croisement de paquets). Ce faible espacement nécessite alors l'utilisation d'une électronique rapide dans les différents détecteurs installés auprès du LHC, afin de limiter les empilements d'événements provenant de croisements différents. Cet empilement d'événements est en effet fonction de la multiplicité des collisions par croisement de paquets (23 interactions inélastiques en moyenne à haute luminosité).

Nous avons donné les principales caractéristiques du LHC, les facteurs qui ont motivé ces choix et les avantages et inconvénients qui en découlent. Nous nous intéresserons dans

²La section efficace des collisions parton-parton évolue en $1/s$.

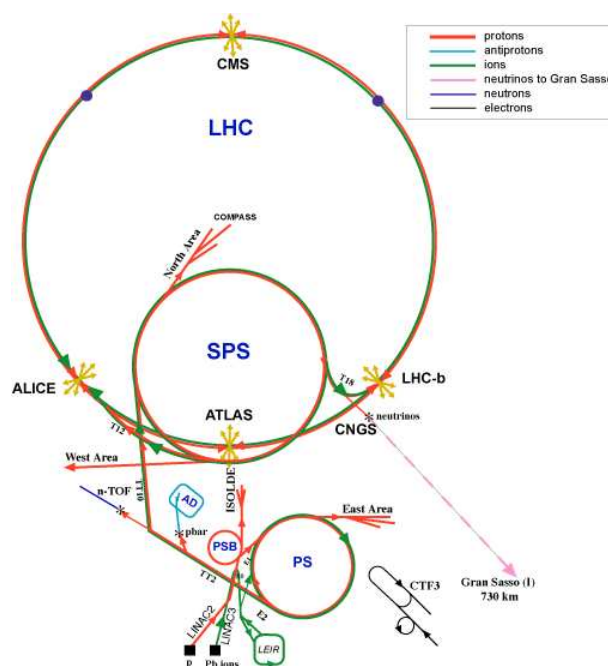


FIG. 2.1 – Le complexe des accélérateurs du CERN.

la suite aux détecteurs qui sont en cours d'installation auprès du grand collisionneur de hadrons.

2.1.1.1 Les détecteurs

Quatre expériences auront lieu auprès du grand collisionneur de hadrons. L'emplacement de ces expériences autour de l'anneau de collision du LHC, ainsi que le complexe des accélérateurs du CERN sont reportés sur la figure 2.1. ATLAS [4] (*A Toroidal LHC Apparatus*) et CMS [5] (*Compact Muon Solenoid*) sont des expériences généralistes couvrant un vaste programme de physique sur et au-delà du Modèle Standard. Le détecteur ALICE [7] (*A Large Ion Collider Experiment*) est dédié à l'étude de la formation d'un nouvel état de la matière, le plasma de quarks et de gluons (*Quark and Gluon Plasma*, QGP). Ce détecteur mettra à profit les périodes de prise de données pendant lesquelles le LHC produira des collisions d'ions lourds³ ($Pb - Pb$). Le détecteur LHCb [6] (*The Large Hadron Collider beauty experiment*) étudiera la physique de la beauté et la violation de la symétrie CP . Une cinquième expérience, Totem [8], sera intégrée au détecteur CMS et sera en charge de la mesure précise de la section efficace totale d'interaction proton-proton ainsi que de l'étude des diffusions élastiques et diffractives.

Dans la suite, nous nous intéresserons exclusivement au détecteur ATLAS.

³ATLAS possède également un programme de physique des ions lourds et sera en mesure de collecter des données pendant cette même période.

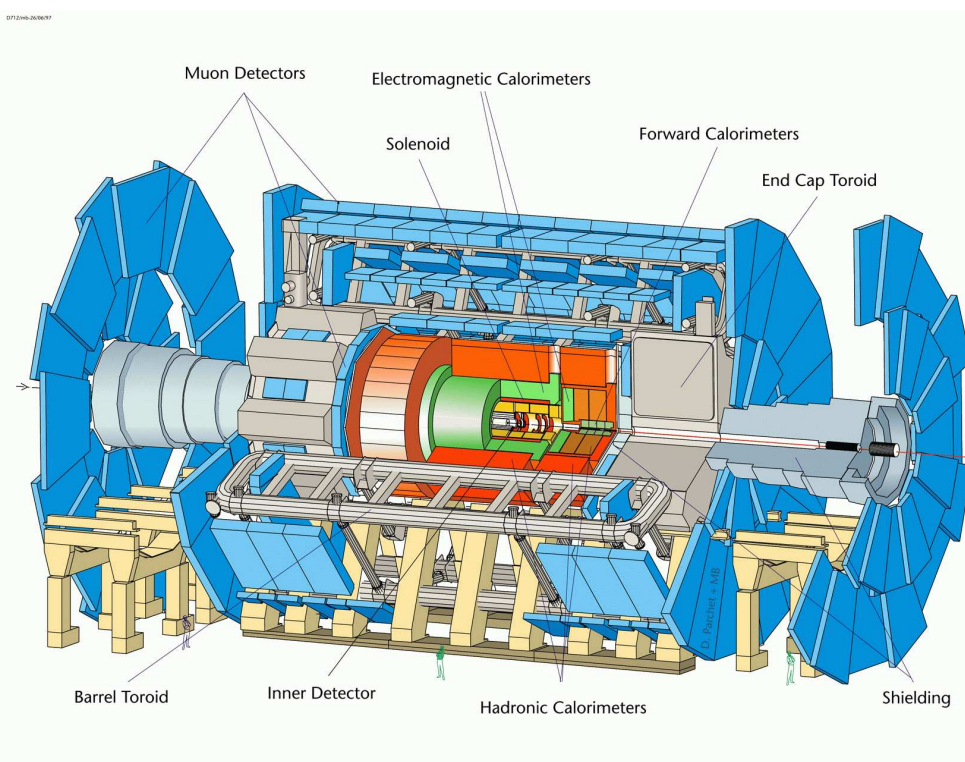


FIG. 2.2 – Le détecteur ATLAS.

2.2 Le détecteur ATLAS

Le détecteur ATLAS (figure 2.2) a été approuvé en 1996. C'est un ensemble de sous-détecteurs de formes cylindriques placés autour du faisceau de protons délivrés par le LHC. Les sous-détecteurs sont agencés en strates enrobant le faisceau de protons.

La collaboration ATLAS regroupe 1770 scientifiques (physiciens et ingénieurs), 34 pays et 152 instituts. Avant de donner une brève description des différents sous-détecteurs composant ATLAS, nous allons donner les grandes lignes de l'optimisation des sous-détecteurs, guidée par les objectifs du programme de physique.

2.2.1 Objectifs de physique et sous-détecteurs

De manière schématique, un détecteur en physique des particules est un outil dont l'objectif est de reconstituer une image la plus fidèle possible des interactions entre particules. Ainsi, de la qualité de rendu de l'image que nous voulons obtenir, résultent des contraintes plus ou moins fortes sur différents sous-détecteurs d'ATLAS. La conception d'ATLAS a été réalisée avec comme optique l'optimisation du potentiel de découverte fourni par le LHC. Il en résulte qu'ATLAS est un détecteur de grande précision mais également polyvalent [29]. Un des buts les plus importants d'ATLAS est de comprendre le mécanisme de brisure spontanée de la symétrie dans le domaine électrofaible [9]. En effet, l'obtention d'informations dans ce domaine de la physique des particules permettrait de répondre à la question de l'origine des masses des différentes particules.

Le potentiel de découverte du boson de Higgs du Modèle Standard pour des masses

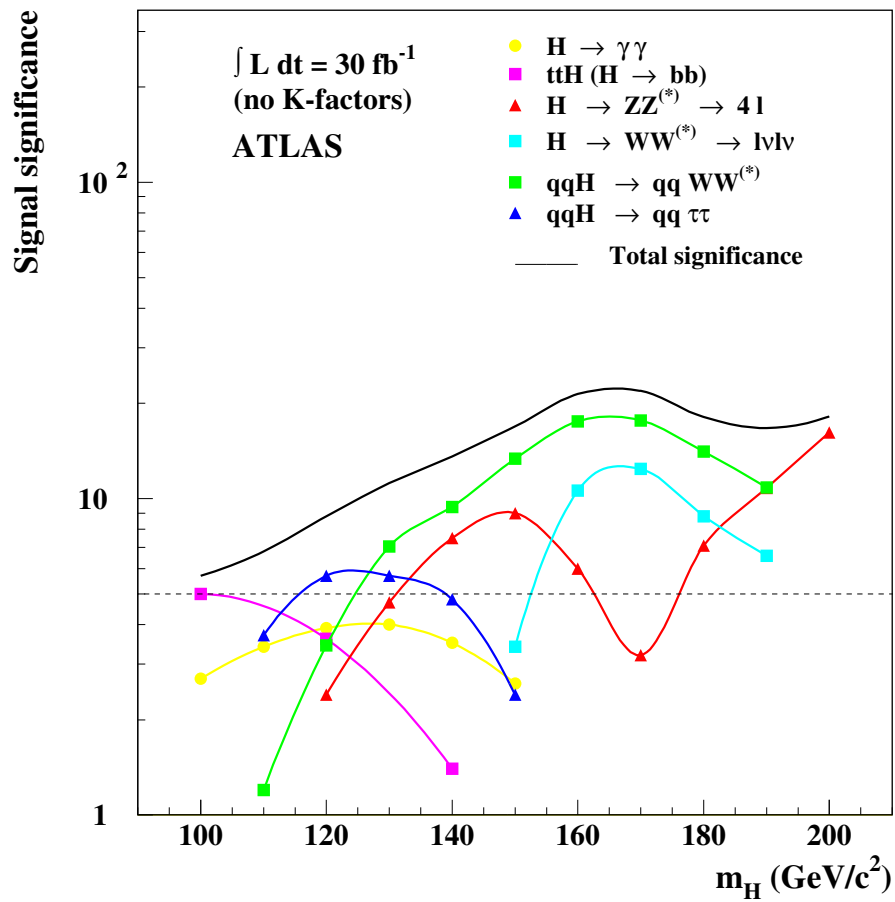


FIG. 2.3 – Sensibilité d'après [10] pour la découverte du boson de Higgs du Modèle Standard en fonction de sa masse, avec le détecteur ATLAS. La luminosité intégrée de 10 fb^{-1} correspond à un an de prise de données à basse luminosité.

allant de 80 GeV à 1 TeV est reporté sur la figure 2.3. Cette figure illustre la variété des canaux à considérer ainsi que les signatures associées. Ainsi, pour un Higgs léger (de masse inférieure à $130 GeV/c^2$), les canaux prépondérants sont $H \rightarrow b\bar{b}$ et $H \rightarrow \gamma\gamma$. Pour le premier canal, il faudra donc une bonne efficacité d'étiquetage des jets issus de b ainsi qu'une bonne réjection des jets légers. Quant à la désintégration en paire de photons, il faudra une bonne identification de ces derniers couplée à une mesure précise de leur angle d'ouverture afin d'obtenir une bonne reconstruction de la masse invariante de cette paire. Pour des masses plus importantes, les états finals comportent des leptons, des neutrinos et des jets, suivant les cas. La bonne reconstruction d'un Higgs dans cette gamme de masse requiert donc une bonne identification des leptons ainsi qu'une bonne mesure de leur énergie, une bonne herméticité afin de mesurer l'énergie transverse manquante et une bonne détection des jets vers l'avant.

Dans le cadre de la recherche de particules supersymétriques, les signatures de tels événements consistent en la détection (et la mesure la plus précise possible) des jets légers et issus de b , des leptons et de l'énergie manquante. Dans le cadre des mesures de précision des paramètres laissés libres par le Modèle Standard, un bon étalonnage de l'échelle en énergie des leptons (par exemple, pour la mesure de la masse du W) ainsi qu'une bonne reconstruction des jets (pour la mesure de la masse du top) sont primordiales.

Ces objectifs se traduisent alors par les contraintes suivantes sur les sous-détecteurs.

– **Détecteur interne :**

- mesure précise de l'impulsion des traces chargées,
- identification efficace des électrons,
- reconstruction précise des vertex secondaires pour l'étiquetage des b et des τ ,
- identification efficace des jets de b et de τ associée à une bonne réjection des jets légers.

– **Calorimètres :**

- bonne couverture angulaire et minimisation des zones mortes afin de garantir une bonne herméticité en vue d'une mesure précise de l'énergie manquante,
- bonne résolution en énergie, tant pour les électrons et photons (calorimètre EM) que pour les jets et hadrons isolés (calorimétrie hadronique),
- identification des électrons provenant de désintégrations semi-leptoniques de b pour améliorer l'étiquetage des jets de b ,
- bonne linéarité de la réponse couvrant tout le spectre en énergie,
- bonne séparation γ/π^0 et γ/jet ,
- bonne séparation angulaire (pour, par exemple, la bonne reconstruction de la masse du Higgs en $\gamma\gamma$).

– **Chambres à muons :**

- grande efficacité de détection des muons,
- bonne mesure des impulsions sur tout le spectre en énergie.

Les objectifs de chaque sous-détecteur ainsi que leur couverture angulaire sont résumés dans le tableau 2.2.

Le détecteur ATLAS est représenté sur la figure 2.2. En partant du point de collision entre les deux paquets de protons, nous trouvons le détecteur interne, qu'entoure le solénoïde fournissant le champ magnétique nécessaire à la mesure de l'impulsion des traces chargées. Les calorimètres électromagnétiques et hadroniques enveloppent le système de trajectographie interne. Nous retrouvons enfin les chambres à muons avec les aimants toroïdaux qui vont courber la trajectoire des muons pour la mesure de leur impulsion.

Sous-détecteur	Résolution (intrinsèque) prévue	Couverture en η
Détecteur interne	$\sigma/p_T \simeq 5 \times 10^{-4} \text{ GeV} \oplus 0.01$	± 2.5
Calorimètre EM	$\sigma/E \simeq 10\%/\sqrt{E} \oplus 0.01$	± 3
Calorimètre hadronique	$\sigma/E \simeq 50\%/\sqrt{E} \oplus 0.03$	± 3
Calorimètre avant	$\sigma/E \simeq 100\%/\sqrt{E} \oplus 0.1$	$3 < \eta < 5$
Spectromètre à muons	$\sigma/p_T \simeq 3\% \text{ } p_T < 100 \text{ GeV}$ et 10% à 1 TeV	± 3

TAB. 2.2 – Performances prévues pour les différents sous-détecteurs d’ATLAS et leur couverture en pseudo-rapacité η , définie comme $\eta = -\ln(\tan \frac{\theta}{2})$ où θ est l’angle polaire défini dans le repère du laboratoire.

Nous allons maintenant rentrer plus avant dans les détails de chaque sous-détecteur, en suivant le flux naturel des particules provenant d’une collision proton-proton.

2.2.2 Détecteurs internes

Le système de trajectographie interne [11] a pour rôle de détecter les traces des particules chargées, de mesurer leurs impulsions et de reconstruire les vertex primaire et secondaires (aussi appelés vertex déplacés). Ces mesures se font grâce à trois détecteurs de technologies différentes, au plus près du point d’interaction. Ces détecteurs sont plongés dans un champ magnétique de 2 Tesla fourni par un solénoïde.

Les mesures précises de l’impulsion des particules et du vertex primaire sont réalisées à l’aide du détecteur à pixel de silicium [12]. Cette mesure de précision est complétée grâce à un détecteur à micro-pistes de silicium (*SemiConductor Tracker*, SCT [13]). Nous pouvons noter que le nombre de couches utilisées pour ces détecteurs est limité en raison d’une part de la quantité de matière ainsi introduite et d’autre part de leur coût élevé. Le dernier élément du détecteur interne permet de suivre de manière quasiment continue les traces qui se propagent à l’intérieur de son volume, grâce à la détection du rayonnement de transition. Ce détecteur (*Transition Radiation Tracker*, TRT [13]) permet ainsi de minimiser la quantité de matière traversée par les particules. L’ensemble du système de trajectographie interne est représenté sur la figure 2.4 et s’inscrit dans un cylindre de 6.9 m de long et de 1.15 m de rayon. La quantité de matière utilisée par l’ensemble du détecteur interne est suffisamment faible pour que les performances de la calorimétrie ne soient pas dégradées. Pour une pseudo-rapacité nulle, elle équivaut à 0.28 X_0 .

Le détecteur à pixels, placé au plus près du point d’interaction, est conçu pour fournir trois points de mesure sur l’ensemble de la couverture angulaire ($|\eta| < 2.5$). Grâce à une segmentation bidimensionnelle en pixels de $50 \times 300 \mu\text{m}$, la localisation spatiale des dépôts d’énergie des particules chargées est donnée sans ambiguïté. Par ailleurs, cette granularité importante permet de réduire le taux d’occupation du détecteur à environ 1 pixel sur 10000 dans un environnement à haute luminosité. Le détecteur à pixels est constitué de trois tonneaux de rayons 4, 10 et 12 cm et de 5 disques de part et d’autre du point d’interaction. Ce détecteur étant le plus près du point d’interaction, il a fallu mettre au point un système permettant de remplacer le premier cylindre, à terme affecté par la grande quantité de radiations émises dans cette région d’ATLAS. Il est prévu de le remplacer après 3 ans de prise de données à basse luminosité et au bout d’un an à luminosité nominale.

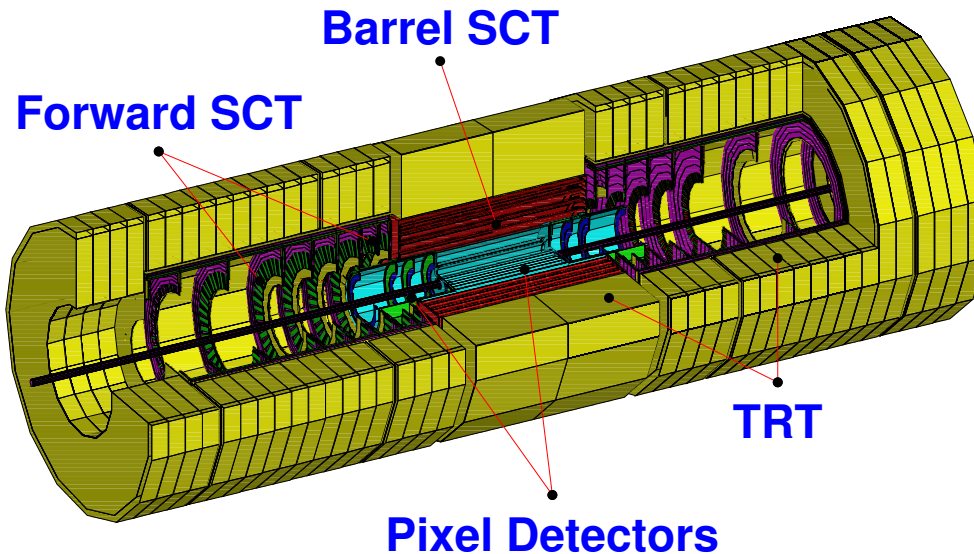


FIG. 2.4 – Système de trajectographie interne.

Le détecteur de traces à semi-conducteur est constitué de quatre cylindres à doubles couches de micro-pistes de silicium collées dos-à-dos. La largeur de ces pistes, l'une orientée suivant la direction azimutale et l'autre tournée de 40 mrad par rapport à la première, est de $80\ \mu\text{m}$. Ces bandes s'étalent sur 126 mm . Le SCT permet ainsi d'obtenir quatre mesures spatiales de grande qualité pour chaque trace, dans la région intermédiaire délimitée par les quatre tonneaux de rayons respectifs 30.0 , 37.3 , 44.7 et 52.0 cm .

La partie la plus externe du détecteur interne est le TRT. Il associe les techniques de rayonnement de transition et des chambres à dérive. Ce détecteur est composé de tubes en plastique de 4 mm de diamètre équipés d'un fil d'anode de $30\ \mu\text{m}$ plaqué en or-tungstène. Ces "pailles" sont regroupés en trois parties : un bloc central où les pailles sont orientées radialement, et deux blocs bouchons dans lesquels les pailles ont une orientation axiale. Un matériau radiateur (feuilles de polypropylène) est intercalé entre les pailles afin de produire des rayons X lors du passage de particules ultra-relativistes. Ces rayons sont ensuite absorbés par le gaz ($70\% Xe$, $20\% CF_4$, $10\% CO_2$) dans lequel baignent les pailles. Ce dispositif permet d'assurer un temps de dérive maximal de 42 ns ainsi que l'obtention d'une trace quasi-continue avec la production d'en moyenne 36 points de mesure.

2.2.3 Calorimètres

Le système de calorimétrie d'ATLAS est reporté sur la figure 2.5. Cet ensemble calorimétrique peut se décomposer en un calorimètre électromagnétique dédié à la mesure précise de l'énergie et de la position des électrons et des photons, et d'un calorimètre hadronique dédié à la mesure de l'énergie des jets et des hadrons isolés.

La calorimétrie électromagnétique utilise de l'Argon liquide comme milieu actif et couvre la région des pseudo-rapidités $|\eta| < 3.2$. Le calorimètre hadronique se décompose en un calorimètre à tuiles scintillantes dans la partie centrale ($|\eta| < 1.7$) et d'un calorimètre à Argon liquide ($1.5 < |\eta| < 3.2$) pour les bouchons. Enfin, le calorimètre vers l'avant met également en œuvre la technologie à Argon liquide pour assurer une bonne

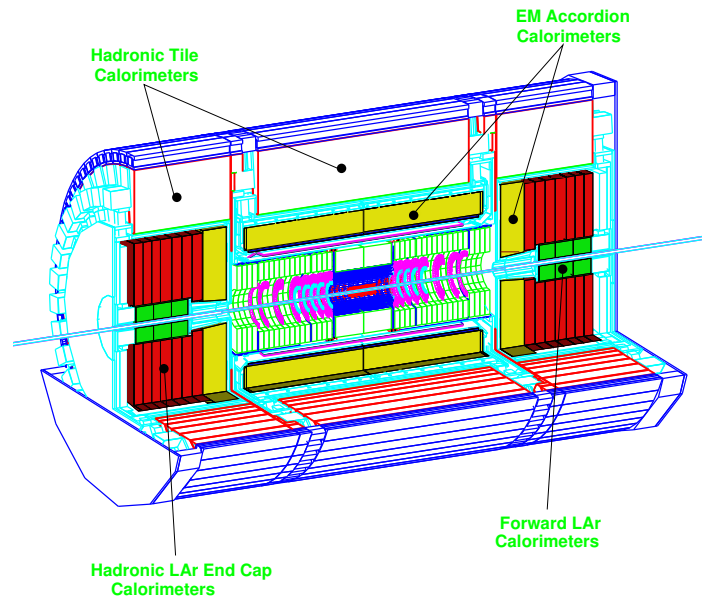


FIG. 2.5 – Système de calorimétrie d'ATLAS.

herméticité grâce à l'apport de sa couverture en pseudo-rapacité : $3.1 < |\eta| < 4.9$.

2.2.3.1 Calorimètres électromagnétiques

Le calorimètre électromagnétique (ECAL [14]) a pour objectif d'identifier les électrons et les photons et de mesurer leurs énergies et positions, avec comme contraintes les objectifs de physique évoqués précédemment. Pour réaliser cette tâche, la collaboration ATLAS a développé un calorimètre électromagnétique à échantillonnage. Ce type de dispositif consiste en l'alternance successive d'un milieu absorbeur dense qui permet le développement d'une gerbe, avec un milieu actif permettant quant à lui la détection des particules composant la gerbe. Le choix de la collaboration ATLAS s'est porté sur le plomb comme absorbeur. Sa faible longueur de radiation ($X_0(Pb) = 0.56 \text{ cm}$) permet d'obtenir un calorimètre compact. Afin d'assurer la cohérence mécanique de cet ensemble, des feuilles d'acier inoxydable sont accolées de part et d'autre de chaque plaque d'absorbeur. Le choix de l'Argon liquide comme milieu actif a quant à lui été motivé par son faible coût, sa plus grande pureté [14] par rapport au Krypton liquide et sa résistance aux radiations. Une électrode de lecture constituée de trois plaques de cuivre séparées par un isolant en Kapton, est placée entre deux plaques d'absorbeur. Les deux couches externes de cuivre de l'électrode connectées à la haute tension d'une part, et les feuilles d'acier inoxydable reliées à la masse d'autre part, permettent l'application d'un champ électrique sous l'action duquel le courant induit par la dérive des électrons d'ionisation pourra s'établir. Le temps de dérive est d'environ 450 ns .

Le calorimètre électromagnétique est divisé en deux parties : le tonneau central cou-

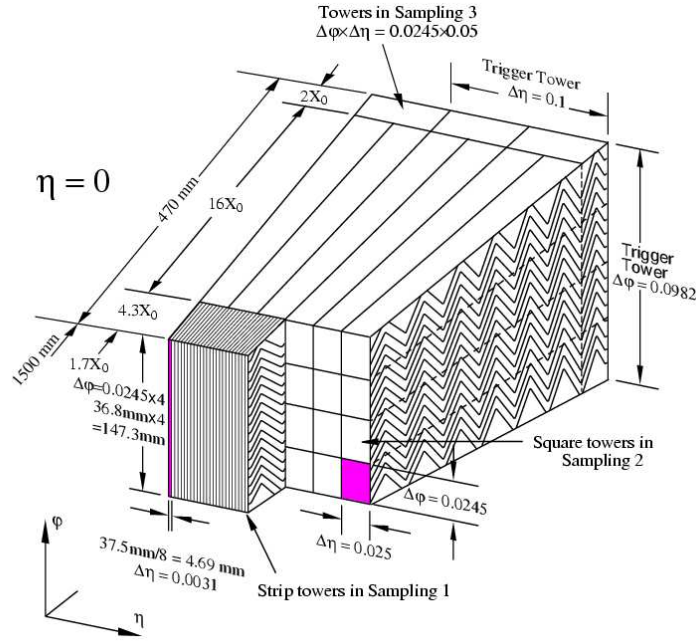


FIG. 2.6 – Schéma de la structure en accordéon du LArg d'ATLAS.

vant la gamme $|\eta| < 1.475$ en pseudo-rapacité, et deux bouchons ($1.375 < |\eta| < 3.2$). Précédée par une pré-échantillonneur pour corriger l'énergie perdue dans le détecteur interne, la partie centrale est enfermée dans le cryostat permettant également de refroidir le solénoïde. Les bouchons partagent leurs cryostats avec les calorimètres hadroniques à Argon liquide et les calorimètres avant. La géométrie originale du ECAL, en accordéon, permet d'assurer une couverture azimutale totale tout en conservant une distance constante entre absorbeur et électrodes. Cette géométrie en accordéon est représentée sur la figure 2.6. Le ECAL totalise 24 (26) longueurs de radiation dans la partie tonneau (bouchon). Il est segmenté en trois sections longitudinales dans la région des mesures de précision ($|\eta| < 2.5$, en regard du détecteur interne). Le premier compartiment, d'une granularité élevée en η ($\Delta\eta \simeq 0.003$) et d'une épaisseur constante d'environ $6 X_0$ ⁴, a été optimisé pour la séparation γ/π^0 , en agissant comme un détecteur pied de gerbe. Pour les pseudo-rapidités plus importantes ($2.5 < |\eta| < 3.2$), le calorimètre présente une segmentation en seulement deux compartiments et une granularité latérale en η moins importante. Cette géométrie satisfait toutefois les contraintes des objectifs de physique relatifs à cette région (reconstruction des jets et mesure de l'énergie transverse manquante). Formé de la partie tonneau et des deux parties bouchon, le ECAL compte partiellement 190000 voies de lecture.

Afin de n'être limité que par l'erreur statistique sur la détermination de la masse du boson de Higgs, l'échelle en énergie électromagnétique du ECAL doit être connue à 0.1% près. Une contrainte encore plus forte (0.02%) provient de la mesure de la masse du boson W à 15 MeV près. Pour atteindre de telles précisions sur l'échelle en énergie électromagnétique, il est nécessaire d'avoir recours à un étalonnage *in situ* du calo-

⁴en prenant en compte la matière en amont

rimètre. Cet étalonnage sera réalisé *via* des processus de physique, comme par exemple la contrainte en masse du boson Z^0 dans le cadre de la reconstruction du canal $Z^0 \rightarrow e^+e^-$, ou alors du rapport E/p des électrons provenant de la désintégration de bosons W ⁵.

Les performances attendues pour le calorimètre électromagnétique sont :

- résolution en énergie : $\sigma/E = 10\%/\sqrt{E} \oplus 0.7\%$, avec un terme de bruit de fond inférieur à 400 MeV,
- résolution sur l'angle azimutal ϕ : $4.6 \text{ mrad}/\sqrt{E}$,
- résolution sur l'angle polaire θ : $60 \text{ mrad}/\sqrt{E}$,
- taux de réjection de 1000, 500 et 3 sur, respectivement, les jets, électrons et π^0 pour des efficacités relatives sur les photons de 80%, 95% et 90%,
- taux de réjection de l'ordre de 10^5 sur les jets pour une efficacité sur les électrons de 70% environ.

Les énergies E sont en GeV.

2.2.3.2 Calorimètres hadroniques

Les calorimètres hadroniques [14, 15] ont pour fonction de mesurer les énergies et positions des jets d'une part, et de fournir une mesure de l'énergie manquante d'autre part. Ils peuvent également servir à l'identification des particules, notamment les muons de basse impulsion transverse.

La couverture angulaire des calorimètres hadroniques s'étend jusqu'à $\eta = 4.9$. Deux technologies sont utilisées dans ATLAS. Chacune est adaptée aux différents besoins physiques et aux différents niveaux de radiations. Profitant des taux de radiations relativement bas de la zone $|\eta| < 1.7$, le calorimètre hadronique de la partie centrale utilise des tuiles scintillantes. Pour les régions vers l'avant ($1.5 < |\eta| < 3.2$), le plus fort taux de radiations a favorisé l'utilisation de la technologie à Argon liquide.

Le système de calorimétrie hadronique présente une épaisseur d'environ 10λ ⁶. Cette épaisseur importante permet de garantir que les gerbes hadroniques soient bien contenues dans le calorimètre et par suite, de réduire au minimum le bruit de fond hadronique dans le spectromètre à muons. La combinaison de l'importante épaisseur en longueurs de radiations avec la large acceptance en pseudo-rapacité, assure de bonnes mesures de l'énergie transverse manquante.

Calorimètre à tuiles scintillantes : TileCal

Le Laboratoire de Physique Corpusculaire de Clermont-Ferrand est très fortement impliqué dans la construction, la certification et l'installation du TileCal. Situé dans la partie centrale, c'est un calorimètre à échantillonnage utilisant du fer comme absorbeur et des tuiles scintillantes comme milieu actif. Les dépôts d'énergie des particules incidentes sont convertis en lumière par les tuiles scintillantes. Le transport de cette lumière vers les photomultiplicateurs est assuré par des fibres optiques à décalage de longueur d'onde. L'originalité du TileCal réside dans l'orientation de ses tuiles scintillantes : ces tuiles sont placées perpendiculairement à l'axe faisceau et donc parallèles à la direction

⁵Ce rapport est calculé à partir de l'énergie mesurée par le ECAL et de la mesure de l'impulsion provenant du trajectographe interne

⁶La longueur d'interaction λ représente la perte d'énergie d'un hadron par interaction nucléaire. Elle correspond à la distance que doit parcourir un hadron pour perdre 63% de son énergie initiale

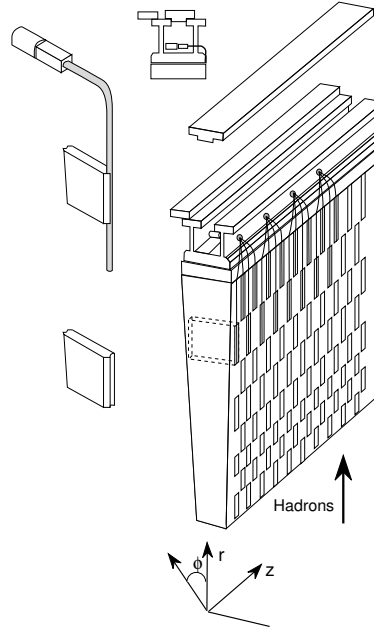


FIG. 2.7 – Coupe d'un module du TileCal. Nous pouvons noter l'orientation des tuiles selon la direction de propagation des hadrons. Ce design original facilite le routage des fibres optiques pour la lecture et permet également la redondance de cette information (les tuiles sont lues de chaque côté).

de propagation des hadrons (voir figure 2.7). Cette disposition permet par ailleurs un routage radial des fibres optiques, facilitant ainsi l'extraction de l'information lumineuse. Ce signal lumineux, extrait de chaque côté de la tuile (double lecture de chaque cellule), est ensuite converti en signal électrique par un photomultiplicateur.

Le TileCal s'inscrit dans un cylindre de 2.28 m de rayon interne et 4.25 m de rayon externe. Il se compose d'un tonneau central de 5.64 m de long ($|\eta| < 1.0$) et de deux parties dites tonneaux étendus de 2.91 m de long ($0.8 < |\eta| < 1.7$). Chacun de ces tonneaux est divisé en 64 modules azimutaux (voir figure 2.7). Ces modules sont radialement divisés en trois couches : A, BC et D. Elles correspondent respectivement à 1.4, 3.9 et 1.8 longueurs d'interaction à $\eta = 0$. Les cellules sont parfaitement projectives en ϕ (de par la géométrie en 64 modules) mais seulement pseudo-projectives en pseudo-rapacité, comme le montre la figure 2.8. Le TileCal contient ainsi près de 10000 photomultiplicateurs qui sont autant de voies de lecture. Une grande partie de l'électronique de lecture est située sur le détecteur. Elle est contenue dans un support mécanique, le super-tiroir, qui est inséré à l'intérieur des poutres métalliques soutenant les modules. Chaque super-tiroir est constitué d'un train de deux tiroirs. Les modules du tonneau central contiennent deux super-tiroirs, ceux du tonneau étendu n'en contenant qu'un seul.

Les performances du Tilecal obtenues en faisceau de pions en terme de résolutions sont :

$$\begin{aligned}
 - \frac{\sigma_E}{E} &= \left[\frac{(41.9 \pm 1.6) \text{ GeV}^{1/2} \%}{\sqrt{E}} + (1.8 \pm 0.1) \% \right] \oplus \frac{(1.8 \pm 0.1) \text{ GeV}}{E}, \\
 - \sigma_\phi &= \frac{(68.17 \pm 0.75) \text{ GeV}^{1/2} \text{ mrad}}{\sqrt{E}} + (0.90 \pm 0.11) \text{ mrad}, \\
 - \sigma_\theta &= \frac{(160.50 \pm 1.48) \text{ GeV}^{1/2} \text{ mrad}}{\sqrt{E}} + (8.15 \pm 0.20) \text{ mrad}.
 \end{aligned}$$

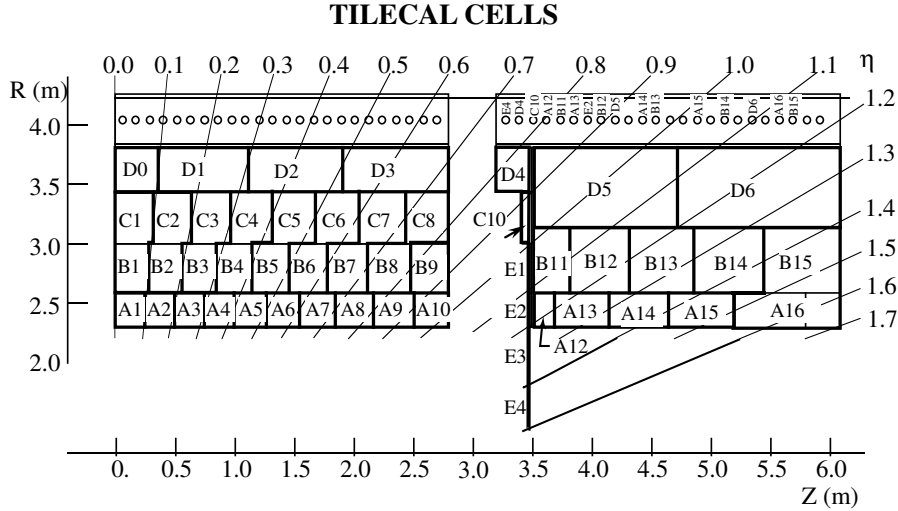


FIG. 2.8 – Les cellules du TileCal. Nous retrouvons le tonneau central sur la partie gauche du schéma, alors que la partie droite est consacrée au tonneau étendu. Les cellules A composent la première couche. Les cellules B et C (qui sont reliées) constituent la deuxième couche. La troisième et dernière couche est composée des cellules D.

En ce qui concerne les jets, la résolution en énergie estimée est :

$$\frac{\sigma_E}{E} = \frac{(53.9 \pm 1.3) \text{ GeV}^{1/2}\%}{\sqrt{E}} \oplus (1.28 \pm 0.20)\% \quad (2.1)$$

pour un cône de reconstruction $\Delta R = 0.4$. Cette estimation est compatible avec les objectifs de précisions de la collaboration ATLAS :

$$\frac{\sigma_E}{E} = \frac{50 \text{ GeV}^{1/2}\%}{\sqrt{E}} \oplus 3\% \quad (2.2)$$

Calorimètres hadroniques à Argon liquide

Les parties “bouchon” du calorimètre hadronique utilisent une technologie différente de celle utilisée par la partie centrale (TileCal). Dans cette région de pseudo-rapidité ($1.5 < |\eta| < 4.9$), l’utilisation d’un calorimètre à Argon liquide est en effet nécessaire à cause du niveau plus important de radiations. Cette partie du calorimètre hadronique est formée de deux roues de 2 m de diamètre, placées dans le même cryostat que les parties bouchon du ECAL. L’absorbeur en plomb du ECAL est ici remplacé par des plaques de cuivre disposées perpendiculairement au faisceau. Le temps de dérive obtenu avec ce dispositif est sensiblement le même que pour le calorimètre électromagnétique ($\sim 450 \text{ ns}$). Chaque roue est constituée de deux segments longitudinaux. Les cellules sont ici encore parfaitement projectives selon ϕ mais seulement pseudo-projectives suivant η .

Le calorimètre vers l’avant est situé à 4.7 m du point d’interaction et sera placé dans le même cryostat que les bouchons des calorimètres électromagnétique et hadronique. Compte tenu de son emplacement et des flux extrêmement intenses de particules et d’énergie, ce calorimètre doit être résistant aux radiations et rapide.

L’objectif principal de ce calorimètre est d’assurer une bonne herméticité pour une bonne mesure de l’énergie transverse, tout en complétant la mesure de l’énergie des jets.

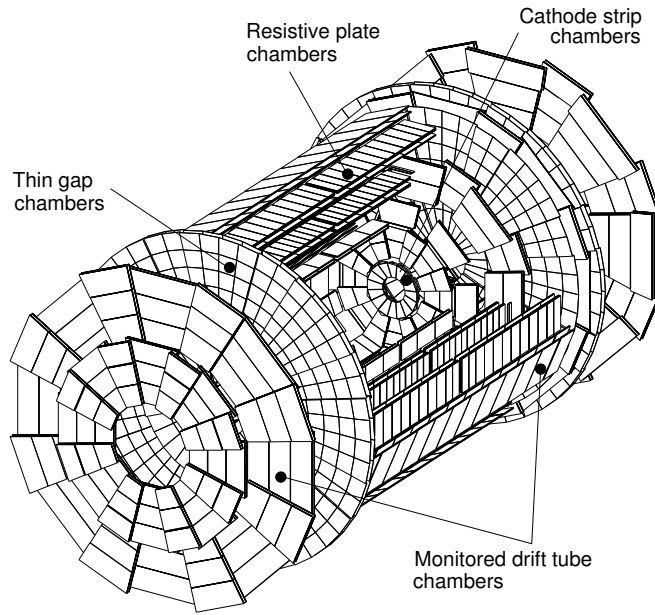


FIG. 2.9 – Vue d’ensemble du spectromètre à muons. Nous pouvons distinguer les quatre types de chambres à muons : TGC, RPC, CSC et MDT.

Il sera ainsi placé dans la région en pseudo-rapacité : $3.1 < |\eta| < 4.9$. Ce calorimètre sert également de protection pour les chambres à muons et pourra ainsi réduire le bruit de fond hadronique dans ces dernières. Le calorimètre vers l’avant (*Forward Calorimeter*, FCAL) est constitué de trois sections longitudinales. La première est en cuivre et les deux autres en tungstène, ceci afin de pouvoir contenir la totalité des gerbes hadroniques dans l’espace limité qui lui est alloué.

Les performances attendues pour les bouchons du calorimètre hadronique et le FCAL sont respectivement de :

$$\begin{aligned} \frac{\sigma_E}{E} &= \frac{50 \text{ GeV}^{1/2}\%}{\sqrt{E}} \oplus 3\% \\ \frac{\sigma_E}{E} &= \frac{100 \text{ GeV}^{1/2}\%}{\sqrt{E}} \oplus 10\% \end{aligned} \quad (2.3)$$

2.2.4 Spectromètre à muons

Le spectromètre à muons [16] est la dernière couche de détection d’ATLAS. Le principe de mesure de l’impulsion avec ce sous-détecteur repose sur la déflexion des muons dans le champ magnétique d’un grand toroïde à air. La mesure de la courbure des trajectoires des muons est réalisée par des chambres de déclenchement et des chambres de précision. Un ensemble de trois stations, sur l’ensemble de l’acceptance angulaire ($|\eta| < 2.7$), est placé de manière à couper la route des muons. Chaque station est un empilement d’une chambre de précision et d’une chambre de déclenchement.

Le déclenchement est réalisé par des RPC (*Resistive Plate Chambers*) dans la partie centrale et par des TGC (*Thin Gap Chambers*) dans les bouchons. Le système de

déclenchement est limité aux pseudo-rapidités $|\eta| < 2.4$. La reconstruction des traces est confiée aux chambres de précision : les MDT (*Muon Drift Tubes*) dans la partie centrale et les CSC (*Cathode Strip Chambers*) pour la région $2.0 < |\eta| < 2.7$. L'ensemble du spectromètre à muons s'inscrit dans un cylindre de 22 m de diamètre et 46 m de long. Nous pouvons également noter que le spectromètre à muons peut être utilisé suivant deux modes de fonctionnement : mode autonome et mode combiné. Ce dernier mode permet de combiner les informations provenant du système de trajectographie interne afin de reconstruire la trace d'un muon depuis le point d'interaction jusqu'aux dernières couches du spectromètre.

Les performances du spectromètre quant à la mesure de l'impulsion des muons sont fortement liées au positionnement relatif des différents modules de détection, et donc de notre connaissance de ce positionnement relatif. Cependant, il n'est pas possible de stabiliser la position de ces modules dans une enveloppe satisfaisant le niveau de précision demandé (50 μm). Cet effet provient de la grande taille du dispositif. Les corrections nécessaires se feront donc directement sur les données grâce à un système optique qui contrôlera les déplacements des stations. D'après des simulations [29], les performances du spectromètre sont :

- une résolution de 3% pour des impulsions transverses de 100 GeV, et entre 4 et 10% au-delà,
- une efficacité de reconstruction de l'ordre de 98%.

2.2.5 Système d'acquisition

Comme nous l'avons évoqué au début de ce chapitre, la fréquence de croisement des paquets de protons aux divers points d'interaction du LHC est de 40 MHz. A haute luminosité, le nombre moyen d'interactions proton-proton étant de 23 par croisement de paquets, cela correspond à environ 10^9 d'interactions inélastiques par seconde. A titre d'illustration, sur ce milliard d'interactions, seulement dix donneront lieu à une paire $t\bar{t}$.

Le niveau actuel de technologie ne permet pas de stocker tous les événements produits lors des collisions. De plus, comme le montre la figure 2.10, une très grande partie de ces collisions ne donne pas lieu à des événements permettant de découvrir ou d'étudier de nouveaux phénomènes physiques. Il faut donc recourir à une sélection en ligne qui puisse ne retenir que les événements intéressants et ramener ce taux à une fréquence compatible avec les ressources matérielles disponibles. Le système de déclenchement (*trigger*) doit ainsi fournir un facteur de réjection de 10^7 pour obtenir une fréquence d'enregistrement sur disque de l'ordre de la centaine de Hertz.

Pour ce faire, le système de déclenchement [17, 18] s'articule sur trois niveaux. Chaque niveau raffine les décisions prises au niveau précédent et applique des critères de sélections supplémentaires si besoin est. Il faut en effet non seulement rejeter le maximum d'événements de biais minimum mais également s'assurer une efficacité de sélection sur les processus physiques rares (et intéressants) la meilleure possible. La structure sur trois étages du système de déclenchement devrait permettre une grande flexibilité mais également une robustesse importante tout au long de la période de prise de données. Une vue schématique des différents niveaux de déclenchement est reportée sur la figure 2.11.

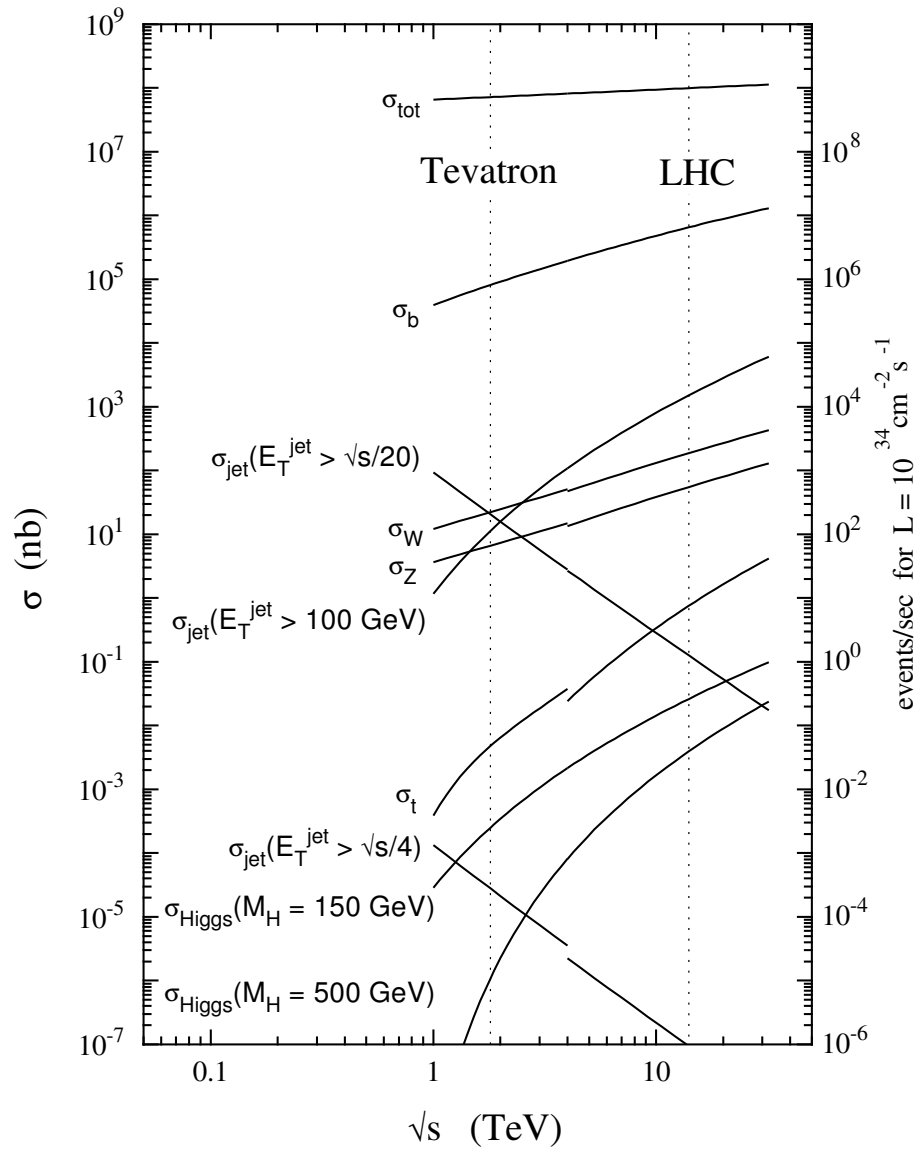


FIG. 2.10 – Sections efficaces et taux de productions de différents processus pour des collisions proton-(anti)proton en fonction de l'énergie disponible dans le centre de masse. Ces résultats ont été obtenus pour un fonctionnement à haute luminosité.

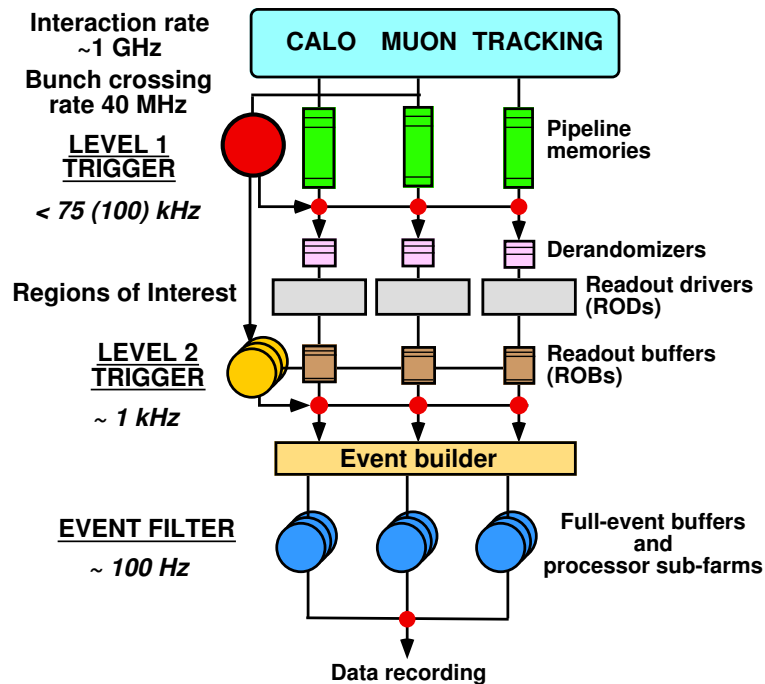


FIG. 2.11 – Vue schématique des différents niveaux de déclenchement constituant le système d'acquisition.

2.2.5.1 Niveau 1

Le premier niveau de déclenchement [19] utilise les informations des calorimètres et du spectromètre à muons. A ce stade du traitement de l'événement, la granularité accessible est réduite. Les muons de grande impulsion transverse sont recherchés avec pour seules informations, celles disponibles *via* les chambres de déclenchement. Les photons ou électrons de haut p_T , les jets et les désintégrations de τ sont identifiés par les calorimètres. L'énergie transverse manquante est également calculée uniquement à partir des tours de déclenchement des calorimètres. De plus, des coupures d'isolation sont appliquées sur les candidats photons, électrons et τ en hadrons.

Ces coupures (voir tableau 2.3) permettront de réduire le taux de déclenchement en sortie à environ 75 kHz ⁷ en sélectionnant rapidement les croisements de faisceau intéressants. De plus, ce premier niveau de déclenchement permet de définir des régions d'intérêt dans lesquelles les particules ayant participé au déclenchement, résident. Ces régions serviront de données d'entrée pour les niveaux suivants. Cependant, le niveau 1 de déclenchement a au maximum $2.5 \mu\text{s}$ pour donner sa réponse, temps correspondant à la capacité maximale des mémoires analogiques (*pipeline*) de chaque sous-détecteur. Grâce à des processeurs dédiés, le temps nécessaire à la décision de déclenchement a été ramené à $2 \mu\text{s}$. Une fois l'événement accepté, l'ensemble des données correspondantes est transféré vers des tampons de lecture (*ReadOut Buffer*, ROB), pour traitement par le niveau 2.

⁷Le taux maximum acceptable en sortie du niveau 1 est de 100 kHz .

Signature	Taux attendus à basse luminosité
$EM25I$	12.0 kHz
$2EM15I$	4.0 kHz
$MU20$	0.8 kHz
$2MU6$	0.2 kHz
$J200$	0.2 kHz
$3J90$	0.2 kHz
$4J65$	0.2 kHz
$J60 + XE60$	0.4 kHz
$TAU25I + XE30$	2.0 kHz
$MU10 + EM15I$	0.1 kHz
étalonnage, facteurs d'échelle	5.0 kHz
Total	~ 25.0 kHz

TAB. 2.3 – Tableau des principales signatures pour la sélection de niveau 1. [18]. Les menus sont de la forme $NoXXI$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet physique ($\mu, e/\gamma, \dots$) sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et I indique s'il faut ou non ajouter un critère d'isolation. XE représente l'énergie manquante dans le plan transverse. Notons qu'environ 20% de la bande passante est allouée pour les déclenchements de type étalonnage et autres. Le total de la bande passante inclut un facteur de sécurité effectif de 4 afin de pallier les incertitudes pouvant affecter les taux de déclenchement au niveau 1.

2.2.5.2 Niveau 2

Le déclenchement de niveau 2 [20] a accès aux régions d'intérêt définies par le niveau 1 ainsi qu'à toute la granularité des sous-détecteurs. Les objets muons sont ainsi raffinés grâce aux informations provenant des chambres de précision du spectromètre et leur impulsion transverse est mieux définie grâce au détecteur interne. Des critères d'isolation peuvent de même être appliqués en utilisant les informations des amas calorimétriques. L'accès à la granularité complète des calorimètres, associée aux informations sur les traces de haute impulsion transverse, permettent d'améliorer l'identification des électrons. Cependant, la haute probabilité de conversion des photons dans le détecteur interne ne permet pas d'envisager de *veto* sur les traces afin d'améliorer la réjection e/γ . Les τ sont sélectionnés en associant une trace de haute impulsion transverse dans le détecteur interne avec un dépôt d'énergie hadronique isolé et très collimé. Pour les jets, ce niveau 2 n'apporte quasiment aucun raffinement supplémentaire. Cependant, des études sont en cours afin d'identifier les jets de b en utilisant l'information sur les vertex déplacés mise à disposition par le détecteur interne. La mesure de l'énergie transverse manquante est améliorée d'une part en utilisant les cellules associées aux tours calorimétriques précédemment utilisées, et d'autre part grâce à la prise en compte des muons.

Grâce à ces informations complémentaires, le second niveau permet de ramener le taux de comptage au kHz . Le temps de latence nécessaire au niveau 2 pour donner sa décision varie de 1 à 10 ms . Selon que l'événement soit accepté ou non, les données du ROB sont transmises au niveau suivant ou abandonnées.

Menu	Taux attendus à basse luminosité
$e25i$	40 Hz
$2e15i$	< 1 Hz
$\gamma60i$	25 Hz
$2\gamma20i$	2 Hz
$\mu20i$	40 Hz
$2\mu10$	10 Hz
$j400$	10 Hz
$3j165$	10 Hz
$4j110$	10 Hz
$j70 + xE70$	20 Hz
$\tau35 + xE45$	5 Hz
$2\mu6$ avec coupures en masse étalonnage, facteurs d'échelle	10 Hz 20 Hz
Total	~ 200 Hz

TAB. 2.4 – Tableau des menus du système de déclenchement de dernier niveau (*High Level Trigger : HLT*) [18]. Les menus sont de la forme $NoXXi$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et i indique s'il faut ou non ajouter un critère d'isolation. xE représente l'énergie manquante dans le plan transverse.

2.2.5.3 Filtre d'événements

Le dernier niveau de déclenchement [20], appelé filtre d'événements (*Event Filter*, EF), a accès à la totalité des données. Les menus de déclenchement disponibles sont reportés dans le tableau 2.4. Il peut ainsi adapter les algorithmes développés dans le cadre de la reconstruction hors-ligne (*offline*) à un environnement en-ligne. Le taux de comptage attendu est alors de l'ordre de la centaine de Herz. Ainsi, après un budget temps de $1s$ pour la décision, le dernier niveau envoie les événements sélectionnés dans une base de données pour stockage définitif.

Chapitre 3

Environnement logiciel

Dans ce chapitre, nous allons nous intéresser à l'environnement logiciel ATHENA [22] qui est le support de travail de la collaboration ATLAS. Nous présenterons dans un premier temps les grandes lignes de la programmation objet en C++ puis nous détaillerons l'organisation interne d'ATHENA. Nous exposerons ensuite le flux de données qu'ATHENA doit gérer, depuis la génération jusqu'à l'analyse. Enfin nous présenterons les différents outils mis à disposition du physicien pour faciliter l'analyse de données ainsi que les outils que nous avons développés pour l'ensemble de la collaboration.

3.1 Éléments de C++

Depuis quelques années le domaine de la physique des particules subit une mutation au niveau des outils logiciels utilisés dans le cadre de la modélisation et la simulation des données de physique ainsi que leur analyse. Cette mutation se caractérise par l'abandon progressif des bibliothèques écrites en FORTRAN. Ce langage procédural étant sur le déclin, il a été remplacé par le C++, notamment, car il apporte la notion de programmation orientée objet (OO) mais également car il met à disposition l'expertise issue du monde de l'industrie qui a effectué cette transition une décennie plus tôt.

Le C++ est un langage informatique et à ce titre, il constitue un ensemble d'instructions qui servent à décrire à un ordinateur les actions que l'on veut effectuer. Un langage informatique est rigoureux. De ce fait, à chaque instruction décrite dans le vocabulaire du langage correspond une action du processeur. Cependant, comme tout type de langage, les langages informatiques ont évolué : chaque langage a apporté une nouvelle couche d'abstraction entre le développeur et la machine. Il y a eu tout d'abord l'assembleur qui est une très fine couche d'abstraction au-dessus de l'ordinateur. Ce langage permet en effet de parler directement au processeur *via* des noms de registres prédéfinis. Cependant, il n'existe pas un langage assembleur mais une multitude : un par type de processeur. Les langages appelés procéduraux ou impératifs, tels que le FORTRAN, le BASIC ou le C, apportèrent une nouvelle couche d'abstraction au-dessus de l'assembleur. Ils permettent dans une certaine mesure de rendre le programme indépendant du matériel sur lequel ce dernier s'exécute. En ce sens, ils constituent une formidable avancée mais ils requièrent néanmoins de raisonner en termes appropriés à la structure de l'ordinateur (fonctions, boucles conditionnelles, ...) plutôt qu'en ceux du problème à résoudre. Le travail de la modélisation d'un problème du monde réel en des termes compréhensibles par une

machine est donc laissé au programmeur. Cet effort additionnel résulte généralement en des programmes difficiles à écrire, déboguer et maintenir. Ce fossé est comblé par les langages OO qui donnent aux programmeurs les outils pour représenter le problème directement en termes appropriés au monde réel. Cette représentation est modélisée par des briques élémentaires appelées *objets*. C'est donc le programme qui s'adapte au problème (au moyen de nouveaux objets) plutôt que le programmeur qui adapte le problème en les termes compréhensibles par un ordinateur. La connection avec l'ordinateur est bien entendu toujours présente. Chaque objet ressemble à un ordinateur : il a un état et présente un ensemble d'opérations que l'on peut lui demander d'effectuer. L'*interface* d'un objet est donc l'ensemble des requêtes que l'on peut soumettre à un type d'objet. Ceci permet de différencier un objet par rapport à un autre. Les objets présentant la même interface sont alors regroupés dans une même classe ou en sous-classes d'une classe mère.

La programmation orientée objet peut schématiquement se résumer à l'élaboration d'un ensemble de classes d'objets modélisant au mieux le problème que l'on veut résoudre, et à la définition des messages que les différents objets peuvent s'envoyer et bien sûr la réaction à chacun de ces messages. Même si la construction de programmes peut alors se résumer à un assemblage astucieux de différents blocs ou à la réutilisation d'objets regroupés dans des bibliothèques généralistes ou dédiées, la création de nouveaux objets peut s'avérer fastidieuse et répétitive. Les environnements logiciels ("*framework*") permettent de pallier ce manque. Ils mettent à disposition un ensemble d'objets et d'outils qui factorisent les tâches répétitives ou complexes : gestion de la mémoire, classes généralistes de base, mécanismes de communication entre les objets ou avec le système sous-jacent, *etc.*...

3.2 Introduction à ATHENA

ATHENA est l'environnement logiciel ("*framework*") de la collaboration ATLAS. Les fondations de ce logiciel reposent sur Gaudi [23], développé initialement par la collaboration LHCb et maintenant amélioré conjointement par ATLAS et LHCb. Lors de la prise de données, ATHENA doit être capable de traiter les données provenant du système de déclenchement d'ATLAS ainsi que du système d'acquisition de données. De plus, ATHENA doit être en mesure de fournir les données précédemment traitées aux physiciens d'ATLAS ainsi que les outils d'analyse nécessaires à la production de résultats de physique.

De par la complexité du détecteur ATLAS ainsi que par sa taille, il est nécessaire d'avoir à disposition un logiciel qui soit modulaire, robuste tout en étant assez flexible pour être en mesure de satisfaire les besoins de l'expérience. Le langage retenu permettant de satisfaire au mieux toutes ces exigences s'est avéré être le C++, couplé à la programmation orientée objet.

Le cœur d'ATHENA est composé essentiellement de modules ("*packages*") écrits en C++, mais certains composants externes sont en FORTRAN (essentiellement les générateurs Monte-Carlo) ou en Java (*ex* : Atlantis, un visualisateur graphique d'événements pour ATLAS).

L'utilisation d'un langage orienté objet tel que le C++ permet, entre autre, de définir un ensemble d'interfaces de base qui définit comment les différents acteurs (les objets) vont communiquer tout en ayant la possibilité de changer leur comportement en utilisant différentes implémentations de ces interfaces. Ainsi, si les interfaces sont bien conçues,

les changements nécessaires qui interviendront au cours de la prise de données d'ATLAS se traduiront simplement par le changement d'une implémentation particulière, sans interférer avec les autres modules d'ATHENA.

Dans le paragraphe suivant, nous allons détailler l'architecture interne d'ATHENA en rappelant les principes clefs qui ont été respectés lors de son développement puis nous énumérerons les différents modules qui ont été identifiés lors de l'écriture d'ATHENA.

3.3 Organisation d'ATHENA

Le principe de base de la programmation orientée objet est le concept d'interface. Ainsi, l'utilisation intensive d'interfaces abstraites dans le code d'ATHENA permet de factoriser une grande quantité de code et *a priori* minimise le nombre de bogues.

De plus, le code d'ATHENA, réparti en différents modules, permet une nette séparation entre les classes représentant les données et les classes manipulant ces données (les algorithmes). Cette séparation algorithmes/données permet d'ailleurs de réutiliser un maximum d'algorithmes dans différents endroits d'ATHENA. En effet si quelqu'un développe un algorithme d'ajustement de trace contraint à la masse, et que cet algorithme repose uniquement sur l'interface définissant un quadri-vecteur, alors cet algorithme pourra être greffé dans n'importe quel code utilisateur et pourra ainsi être appliqué sur tout objet satisfaisant l'interface "quadri-vecteur". D'un point de vue plus technique, cette séparation permet aussi de ne charger que les bibliothèques relatives à un certain type d'objets ou d'algorithmes pour un travail donné. Il est en effet inutile de charger l'intégralité des bibliothèques nécessaires à la reconstruction des dépôts d'énergie dans les calorimètres lorsque l'on est seulement intéressé par la reconstruction de traces. On réduit ainsi le temps nécessaire au démarrage proprement dit du travail de reconstruction (moins de bibliothèques à charger) ainsi que l'empreinte mémoire de l'application.

Par ailleurs, il existe une séparation claire entre les données destinées à l'enregistrement sur disque (les données persistentes) et celles dont le domaine de validité n'excède pas celui de la mémoire volatile. Ceci permet de ne pas lier les différents algorithmes de reconstruction et d'analyse à une technologie particulière pour enregistrer les données sur bande. En effet, l'expérience ATLAS devant prendre des données durant une dizaine d'années, plusieurs technologies de persistance des données sont susceptibles de se succéder ¹.

Nous avons donc survolé les différentes figures de mérite de l'environnement logiciel ATHENA. Ces principes clefs sous-tendent l'organisation en modules d'ATHENA. En effet, ATHENA peut être découpée de manière cohérente par sous-détecteurs (détecteur interne, calorimètres puis spectromètre à muons), par le type des données manipulées (génération d'événements, simulation du détecteur, reconstruction puis analyse) ou bien par le type d'interface qui nous est présenté (interfaces pour les opérations basiques, interfaces pour la génération ou la simulation et interfaces pour la sélection d'événements, leur reconstruction puis leur analyse). Les deux dernières classifications semblent plus pertinentes puisqu'elles peuvent être retrouvées à l'intérieur des modules de presque chaque sous-détecteur. Ainsi dans la suite de ce paragraphe, nous détaillerons les différents types

¹Avant même le début de la campagne de prise de données, la collaboration ATLAS a déjà migré depuis le format de fichiers ZEBRA [24] vers POOL [25]

d'objets (définis par leur interface) qu'ATHENA met à disposition pour l'élaboration de programmes de reconstruction ou d'analyse, le découpage en terme de fonctionnalités (génération, simulation, reconstruction et analyse) sera abordé dans le prochain paragraphe.

3.3.1 Objets de base d'ATHENA

L'objet central dans tout type de travail, tant au niveau génération qu'au niveau analyse de physique, est l'algorithme. Cette particularité se retrouve dans l'implémentation d'ATHENA puisque nous retrouvons un objet appelé `Algorithm` au cœur même de l'architecture de l'environnement logiciel d'ATLAS (figure 3.1). Chaque utilisateur voulant créer son algorithme devra créer une classe dérivant de la classe de base `Algorithm` puis surcharger les méthodes `initialize`, `execute` et `finalize`. La méthode `execute` sera ainsi appelée pour le traitement de chaque événement alors que les méthodes `initialize` et `finalize` ne seront appelées qu'une et une seule fois, respectivement lors du lancement de l'application et juste avant son arrêt.

L'`Algorithm` peut ainsi utiliser les différents services d'ATHENA pour accéder aux histogrammes, aux données relatives au détecteur et celles relatives à l'événement courant. De plus grâce au "*JobOptions Service*", l'utilisateur peut modifier les valeurs d'un certain jeu de paramètres de l'algorithme sans avoir à recompiler. Cela se fait grâce à un fichier écrit en Python [26], langage de scripts orienté objet, dans lequel l'utilisateur peut modifier les couples clef-valeur pour chaque instance d'un algorithme.

En outre, et toujours dans l'esprit de la plus grande réutilisation de code, il est d'usage de déléguer au maximum le travail d'un algorithme à un jeu d'`AlgTools`. Comme les `Algorithms`, les `AlgTools` ont une méthode d'initialisation et de finalisation. Cependant, la méthode principale d'un `AlgTool` pourra être appelée plusieurs fois au cours d'un même événement. Les algorithmes peuvent ainsi se réduire à un intelligent agencement d'`AlgTools` pour accomplir leur tâche. De plus, les `AlgTools` pourront ainsi être réutilisés par d'autres algorithmes avec d'autres paramètres de fonctionnement, réduisant de fait la quantité de code à réécrire.

D'autre part, il faut bien noter que n'importe quel travail d'analyse de physique doit mettre en œuvre un certain nombre d'algorithmes : il faut donc que ces différents algorithmes puissent se communiquer les données à analyser. Cette tâche est déléguée au `StoreGateSvc`.

3.3.2 Accès aux données : `StoreGateSvc`

Afin de minimiser le nombre de copies de données lors du passage d'un algorithme A à un algorithme B, ATHENA utilise le modèle du tableau noir (figure 3.2). Ce concept consiste en une base de données stockée en mémoire enregistrant des références vers les différents objets produits au cours de l'exécution de l'application. Ainsi, au lieu de copier un objet A d'un algorithme vers un autre, seule la référence vers l'objet est copiée.

Au-delà de l'aspect optimisation des performances, ce modèle réduit grandement le couplage entre les algorithmes. En effet, l'algorithme B n'a besoin de connaître ni l'identité (le type ou la classe) de l'algorithme qui a produit les données qu'il est en train de traiter, ni la manière (le protocole ou l'interface) à utiliser pour récupérer ces données. Le seul acteur que chaque algorithme a besoin de connaître pour publier ou récupérer des

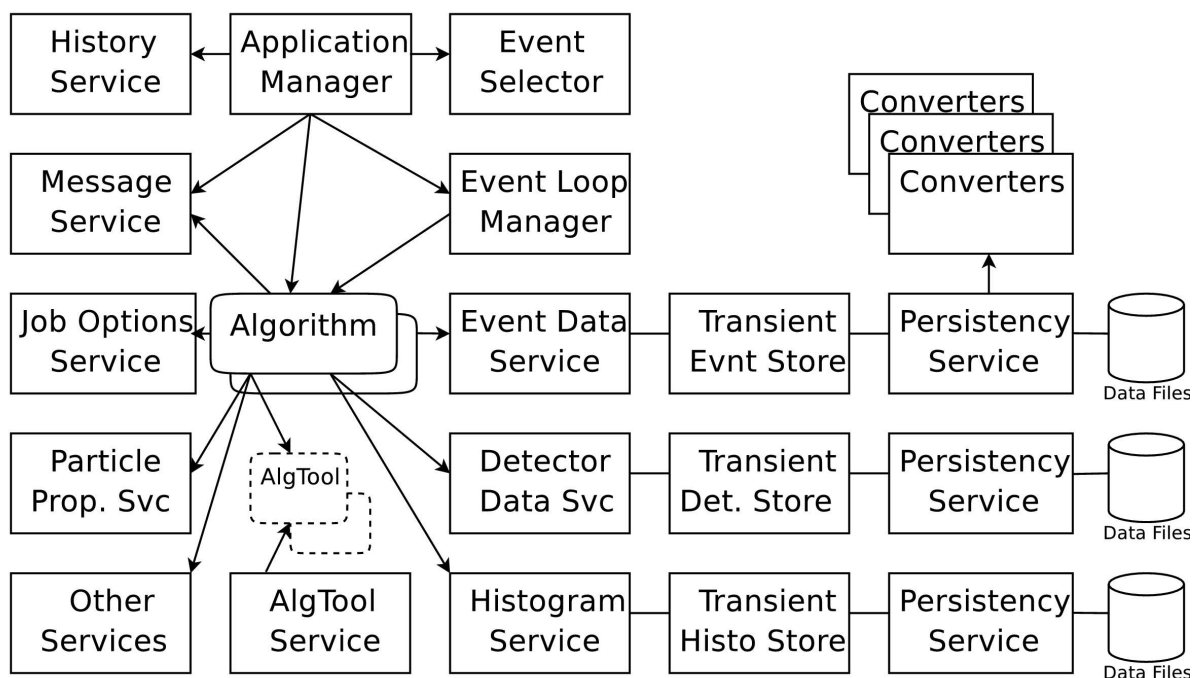


FIG. 3.1 – Les différents objets de base qu'ATHENA met à disposition et en relation. On retrouve au centre de cette architecture la classe **Algorithm**. Les instances de cette classe sont chargées par le gestionnaire d'application (*Application Manager*) et appelées à chaque nouvel événement par le contrôleur de boucle (*Event Loop Manager*). Chaque instance d'une classe dérivant d'**Algorithm** a ensuite accès aux différents services mis à disposition par Gaudi/ATHENA : le *Message Service* pour contrôler le niveau de verbosité de l'algorithme, le *JobOptions Service* pour configurer les paramètres d'exécution de cet algorithme et le service de gestion des histogrammes. Les instances de chaque **Algorithm** peuvent également récupérer des pointeurs vers des **AlgTools** via le service des outils. Ces **AlgTools** peuvent être appelés plusieurs fois par événement, à la discrétion de l'implémentation de chaque algorithme (au contraire des instances de classes dérivant de **Algorithm** qui ne sont appelées qu'une et une seule fois par le contrôleur de boucle d'événements).

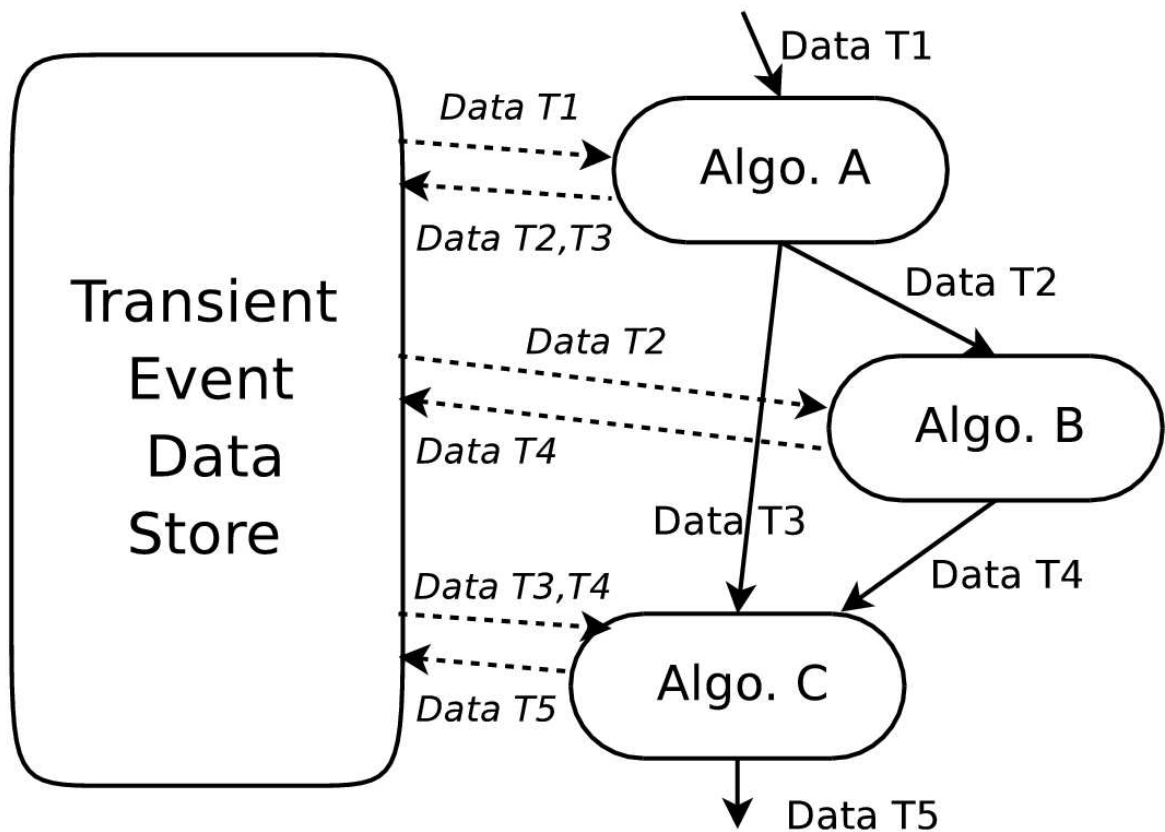


FIG. 3.2 – Modèle du tableau noir (*StoreGateSvc*) : flux apparent et flux réel des données entre algorithmes. L'algorithme A récupère les données T1 depuis *StoreGate* et enregistre les données T2 et T3 à la fin de son exécution. Puis l'algorithme B récupère les données T2 et publie les données T4 après traitement. Vient ensuite l'algorithme C dans l'ordre d'exécution qui a besoin des données T3 et T4 afin de produire les données T5. Vu de l'extérieur, tout se passe comme si l'algorithme C demandait les données T3 à l'algorithme A et les données T4 à l'algorithme B mais le seul vrai interlocuteur de chaque algorithme est *StoreGate*. On peut ainsi remplacer l'algorithme B par un algorithme B' qui publiera d'autres données T4 sans pour cela modifier toute la chaîne d'exécution en aval.

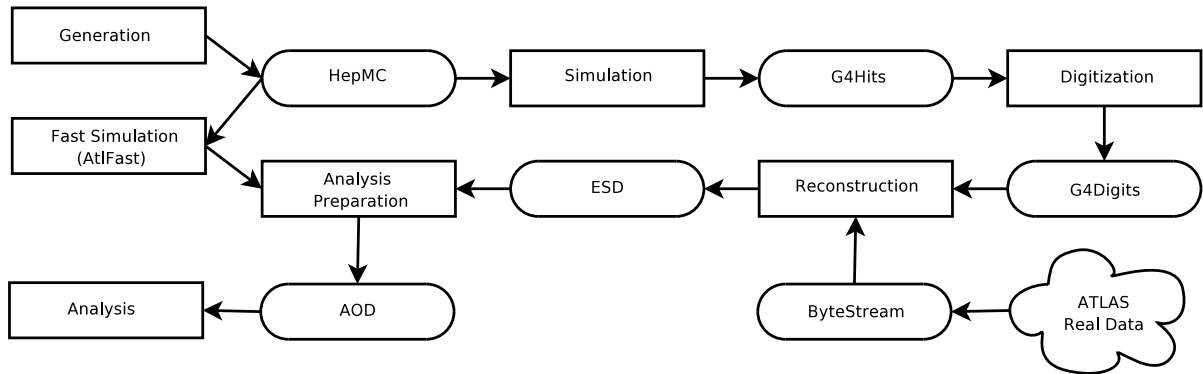


FIG. 3.3 – Schéma simplifié des différents flux de données lors de la simulation rapide (*At1Fast* [35]) et de la simulation complète (*Geant4* [33]). Les rectangles représentent les différentes étapes du processus de gestion du flux de données. Les ovoides représentent les types d’objets échangés entre les différentes étapes.

données, est `StoreGateSvc`. Ainsi, chaque algorithme peut enregistrer un objet ou une liste d’objets en mémoire *via* `StoreGate` au moyen d’une simple chaîne de caractères. Une fois l’objet enregistré dans `StoreGate`, c’est le tableau noir qui en prend la responsabilité et qui s’assure que cet objet sera vidé de la mémoire à la fin du traitement de chaque événement. D’autre part, `StoreGate` permet d’enregistrer les liens qui peuvent exister entre objets enregistrés dans `StoreGate`, l’utilisateur peut donc créer dans un algorithme de reconstruction une association entre une trace et un dépôt calorimétrique par exemple, puis récupérer cette association dans un algorithme dédié à la calibration ou l’analyse.

L’ensemble des briques de base que nous venons d’évoquer est ensuite réutilisé à chaque étape de l’analyse *“offline”* : génération, simulation, reconstruction et analyse de physique.

3.4 Flux de données

Les modules d’ATHENA peuvent être regroupés en modules traitant de la génération d’événements Monte-Carlo, ceux traitant de la simulation du détecteur, ceux qui gèrent la reconstruction des événements dans le détecteur et enfin les modules qui sont le point de départ des analyses de physique. Les données ainsi que leur trajet entre les différentes étapes du processus de modélisation d’un phénomène physique sont reportés sur la figure 3.3. Nous détaillerons dans la suite l’étape de génération des événements qui alimente soit la simulation complète du détecteur et les processus de numérisation des données et leur reconstruction, soit la simulation (et reconstruction) rapide avec le programme *At1Fast*, puis la préparation des données en un format adapté aux analyses de physique (AOD : *Analysis Object Data*). Lorsqu’ATLAS sera en mesure de prendre des données, un troisième chemin viendra se greffer : les parties génération et simulation seront court-circuitées et les données arriveront directement à l’étape reconstruction.

3.4.1 Génération

Les générateurs Monte-Carlo sont les outils qui permettent de simuler les interactions de particules qui ont lieu à l'intérieur des collisionneurs de particules. Ces logiciels fournissent des événements de physique qui sont ensuite comparés aux données réelles lorsque les collisionneurs sont en fonctionnement afin de vérifier la cohérence de notre compréhension du détecteur et de la physique. Lors de l'élaboration d'un nouveau détecteur, ils permettent d'extraire les contraintes pour un bon fonctionnement lors de la prise de données ainsi que d'estimer son potentiel de découverte. De plus, ces programmes permettent d'élaborer les stratégies d'analyse et d'estimer les corrections à apporter aux calculs d'acceptance. Enfin ils illustrent également les incertitudes quant à notre modélisation des processus physiques.

Actuellement, la plupart des générateurs Monte-Carlo sont en FORTRAN puisque les premières versions ont été écrites dans ce langage. Mais pour des raisons de maintenance de ces logiciels, les auteurs se tournent vers le C++. En effet, *Pythia8* [27] et *Herwig++* [28] sont en cours d'écriture et ne présentent donc pas toutes les fonctionnalités de leurs prédécesseurs. Seul *Sherpa* a été écrit dès le début en C++. Ces nouveaux programmes n'ont pas été suffisamment testés pour être utilisés en production², ce qui implique de se reposer sur les bibliothèques FORTRAN, au moins jusqu'au démarrage du LHC.

La génération d'un événement Monte-Carlo se décompose comme suit. La collision primaire, proton-proton dans le cas du LHC, est tout d'abord simulée en prenant en compte la modélisation du processus dur parton-parton, la modélisation des interactions des partons spectateurs de chacun des protons et la modélisation des événements d'empilement (interaction entre d'autres protons des paquets). Les éventuels partons issus de la collision subissent le processus d'hadronisation dû à l'interaction de couleur qui peut se décomposer en deux étapes. Il y a tout d'abord la génération d'une cascade de quarks et de gluons lorsque l'énergie des partons est telle qu'ils peuvent être considérés comme libres ($\sim 1 \text{ GeV}$) : ce processus est décrit de manière perturbative par QCD (pQCD). On a ensuite recourt à des modèles phénoménologiques tels que le modèle des cordes de Lund [31]. Les états finals des partons sont donc un ensemble de particules que l'on peut rassembler dans une enveloppe ayant la forme d'un cône. Les leptons sont soumis à la même méthodologie (mais *via* la modélisation de QED) afin d'arriver aux états finals.

L'interaction de ces particules avec le détecteur est traitée lors de l'étape de la simulation.

3.4.2 Simulation

La simulation est l'étape qui permet d'obtenir la réponse du détecteur à un événement de physique produit lors de la génération, connaissant la géométrie du détecteur ainsi que les caractéristiques détaillées des matériaux le composant. Ce processus peut se décomposer en deux étapes : la simulation du détecteur et la numérisation de sa réponse.

La simulation du détecteur ATLAS avait été essentiellement basée sur une modélisation *via* le programme écrit en FORTRAN *GEANT3* [32]. Avec le développement de *Geant4* [33,

²Il est à noter cependant que *Sherpa* a été utilisé lors de la production des données pour l'atelier de Physique de Rome [30], ce qui constitue la première étape de la validation.

34] depuis l'an 2000, la collaboration ATLAS a commencé à migrer vers cet environnement de simulation orienté objet.

L'étape simulation (de même que la reconstruction) est très gourmande en puissance de calcul. Pour satisfaire les analyses de physique ayant besoin d'une grande statistique en événements, un programme de simulation rapide du détecteur, `At1Fast` [35, 36], a été développé. On pourra noter qu'`At1Fast` regroupe les étapes de simulation du détecteur et de reconstruction des données. Ainsi, on a directement accès à des objets de physique des particules (jets, électrons, ...) en sortie de ce programme. La grande statistique accessible grâce à `At1Fast` se paye au prix d'une modélisation moins fine et complète de la géométrie du détecteur et de sa résolution.

3.4.2.1 Simulation complète : Geant4

`Geant4` fait partie du projet "*LCG Application*" [37], développé à travers le monde et coordonné par le CERN. Tout comme `Gaudi` et `ATHENA`, `Geant4` est un environnement logiciel (*Framework*). Il est cependant orienté vers la simulation de détecteurs en physique des particules et à cette fin, il met à disposition un ensemble d'outils pour décrire la géométrie du détecteur³ et la propagation des particules à travers les différents éléments du détecteur. La description des matériaux de ces éléments ainsi que la modélisation des processus physiques sont ensuite prises en compte afin de rendre la simulation la plus réaliste possible.

La figure 3.4 schématise les différentes étapes de la simulation ainsi que les objets créés lors de ce processus. Les objets issus de la génération sont tout d'abord récupérés pour être ensuite utilisés comme données d'entrée pour simuler le trajet des particules dans le détecteur. Le résultat de ce processus est la production d'une collection d'objets appelés *hits* et contenant l'information des interactions de chaque particule avec les différentes parties du détecteur (position, dépôt d'énergie, ...). Si l'utilisateur a activé la simulation de l'empilement d'événement, des fichiers de bruit de fond sont alors lus à partir desquels des événements sont aléatoirement extraits pour en tirer une collection de *hits* supplémentaire. Cette collection de *hits* est ensuite intégrée à la précédente. Le processus de numérisation prend en compte le caractère fini de l'intervalle en gamme de lecture des différents composants de l'électronique d'acquisition. La simulation de ce processus passe donc par la création d'objets RDO (*Raw Data Object*) qui sont donc la modélisation de ce que l'on aura en sortie du détecteur ATLAS. En parallèle, le pendant Monte-Carlo de ces objets (SDO pour *Simulation Data Object*) est créée et l'association entre ces objets est réalisée. Lors de la prise de données, l'ensemble de cette chaîne de simulation sera remplacée par la lecture et la transformation du flux de données (*ByteStream*) provenant du détecteur. Comme le laisse supposer la figure 3.4, le nombre d'étapes rencontrées lors du processus de simulation complète rend cette partie très gourmande en ressources informatiques : un événement $H(130 \text{ GeV}/c^2) \rightarrow ZZ^* \rightarrow 4l$ est simulé en $\sim 2000s$ et occupe $\sim 2Mb$ d'espace disque [38]. Ces performances sont en deçà du point de fonctionnement prévu pour la prise de données [40], cependant plusieurs améliorations

³Il est important de noter que la géométrie d'ATLAS lors de la simulation est générée à partir du code décrivant celle de la reconstruction et assure donc que les géométries soient identiques. En effet, les objets traités étant différents au cours de ces étapes, ce mécanisme de génération d'une géométrie à partir de l'autre permet de s'assurer de la cohérence des résultats sans avoir à dupliquer du code inutilement.

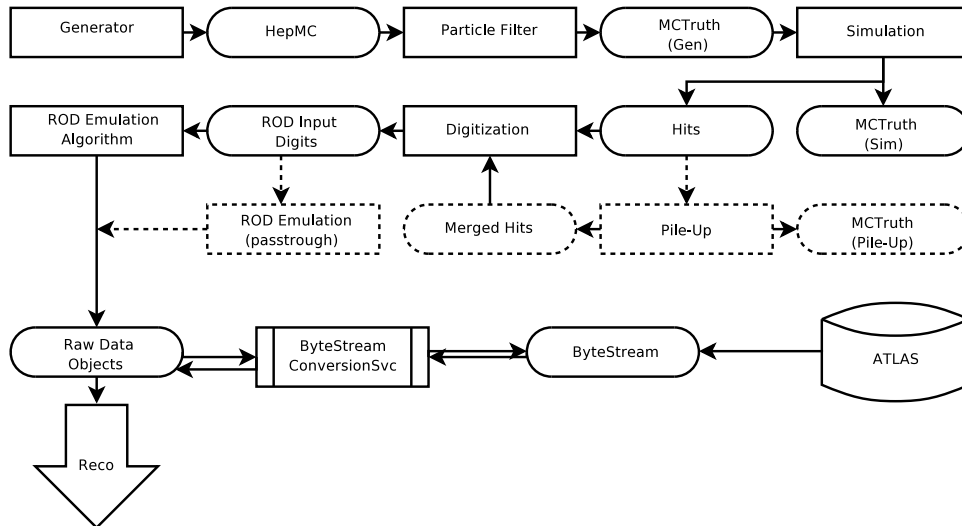


FIG. 3.4 – Flot des données lors de la simulation complète du détecteur. Les différentes étapes de la simulation sont représentées par les rectangles et les objets produits au cours de chaque étape, par des ovoïdes. Les rectangles en pointillés sont des étapes optionnelles de la simulation. La première ligne décrit l'étape génération qui fournit les événements de physique.

et optimisations sont plausibles. Néanmoins, il est possible d'utiliser la simulation rapide pour les analyses nécessitant une statistique importante.

3.4.2.2 Simulation rapide : AtlFast

AtlFast, le programme de simulation rapide du détecteur ATLAS, simule des événements de physique en prenant en compte les effets dus à la réponse du détecteur ainsi que les effets dus à la chaîne de reconstruction logicielle. Cette simulation est réalisée à partir des collections de quadri-vecteurs de particules stables mises à disposition par les générateurs Monte-Carlo. Afin de gagner en temps de calcul, tous les détails de la simulation du détecteur ainsi que toutes les étapes de la reconstruction ne sont pas reproduites. Ainsi, toutes les interactions des particules avec les éléments du détecteur ne sont pas simulées. Les calorimètres sont regroupés en une seule entité et découpés en cellules transversales uniquement. L'énergie d'une cellule est alors la somme des particules la traversant et la résolution sur l'énergie des jets est modélisée (à basse luminosité et pour le tonneau) par cette fonction de résolution [35] :

$$\frac{\delta E}{E} = \frac{0.50}{\sqrt{E}} \oplus 0.03$$

Les jets sont ensuite construits à partir de *clusters* (amas de cellules) puis labellés suivant un critère de proximité par rapport aux particules Monte-Carlo ($\Delta R < 0.2$ pour les jets de b et c , et $\Delta R < 0.3$ pour les jets de τ).

Un algorithme optionnel, **AtlfastB**, peut ensuite être exécuté pour appliquer différentes configurations d'efficacité d'étiquetage de jets de b et appliquer une fonction d'étalonnage paramétrisée en fonction de l'impulsion transverse du jet initial. Les différentes configurations utilisées dans le chapitre suivant sont reportées sur la figure 3.5.

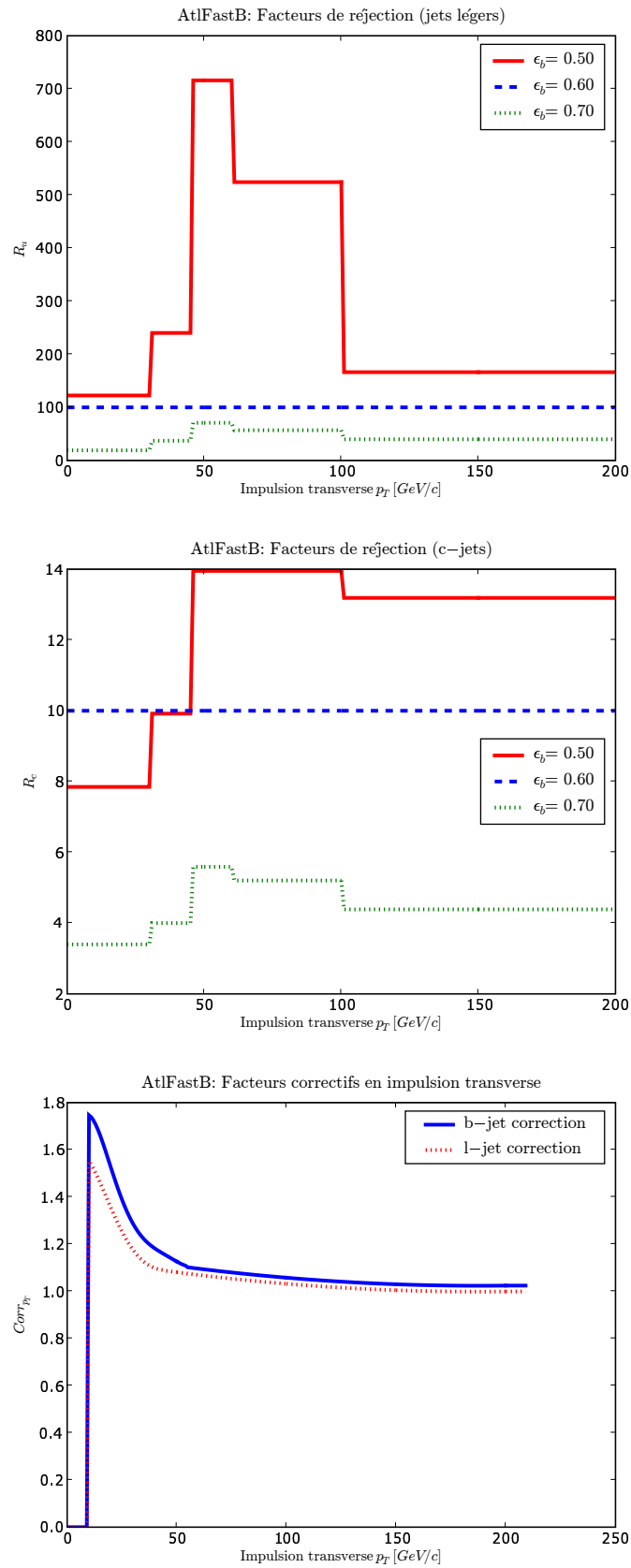


FIG. 3.5 – Facteurs de réjection d'At1FastB (en haut) et fonctions d'étalonnage d'At1FastB (en bas) pour différentes valeurs d'efficacités d'étiquetage des jets de b .

La validation des paramètres utilisés pour la simulation rapide se fait grâce aux résultats obtenus en simulation et reconstruction complètes. Le développement d'un outil de comparaison [39] a récemment été démarré. Grâce à **AtlFast**, une étude peut se décomposer en deux temps : une première analyse avec **AtlFast** permettant d'estimer la faisabilité d'extraction d'une variable tout en prototypant le corps de l'algorithme d'analyse. Si les résultats sont probants, la deuxième étape peut être abordée en utilisant la simulation et reconstruction complètes.

3.4.3 Reconstruction

Comme illustré sur la figure 3.4, une fois le processus de simulation du détecteur terminé vient ensuite l'étape de la reconstruction des données, qu'elles soient simulées ou bien directement issues de l'électronique d'acquisition. Le débit estimé [40] étant de 3 *Pb* (petabytes ⁴) de données brutes par an, il n'est pas possible de distribuer ce flux tel quel à l'ensemble de la collaboration ATLAS. Les données en sortie du système de déclenchement sont donc converties en objets C++ (*Raw Data Objects*) puis traitées par la chaîne de reconstruction et combinées en un fichier ESD (*Event Summary Data*) qui contient toutes les informations nécessaires pour mener à bien n'importe quelle analyse de physique. La taille prévue pour ce fichier est de 500 *kb* par événement [40]. Il faudra donc optimiser les objets ESD puisque la taille actuelle ⁵ est d'environ 1 *Mb* par événement (les informations Monte-Carlo contribuant à hauteur de moins de 150 *kb*).

Le processus de reconstruction peut se subdiviser en deux sous-étapes : la reconstruction des données pour chaque sous-détecteur puis la reconstruction combinée. Cette dernière utilise les informations de tous les détecteurs afin de reconstruire les quadri-vecteurs le plus fidèlement possible et identifier les particules, pour toute la gamme en impulsion, la totalité de l'acceptance et ce pour toutes les luminosités. Les traces et les dépôts d'énergie étant reconstruits, les briques de bases nécessaires à la reconstruction des particules et leur identification sont disponibles. Cette étape est alors effectuée de manière lâche afin de biaiser le moins possible les analyses en aval. Elles pourront ainsi établir leur jeu de coupures optimal qui leur est spécifique.

Le travail d'un algorithme de reconstruction typique se décompose comme suit. Il récupère d'abord une ou plusieurs collections de données brutes (RDO) depuis **StoreGate** puis appelle séquentiellement un jeu d'outils (usuellement des **AlgTools** mais cela peut aussi être des sous-algorithmes) puis publie sur **StoreGate** une collection de données reconstruites (ESD). Comme mentionné précédemment, les outils de reconstruction sont réutilisés entre détecteurs de même type (trajectographes et calorimètres) grâce aux interfaces abstraites pour minimiser les dépendances.

3.4.3.1 Reconstruction des traces

La caractéristique fondamentale [41] qui a été prise en compte lors de l'élaboration de l'architecture des classes de données pour la trajectographie est le support des différents systèmes de trajectographie tout en ayant un socle commun de code. Cela se traduit bien

⁴Un petabyte correspond à 10^6 gigabytes.

⁵Taille d'un fichier ESD d'événements $t\bar{t}$ produits avec la version 10.0.1 d'ATHENA à l'occasion de l'atelier de physique de Rome [30].

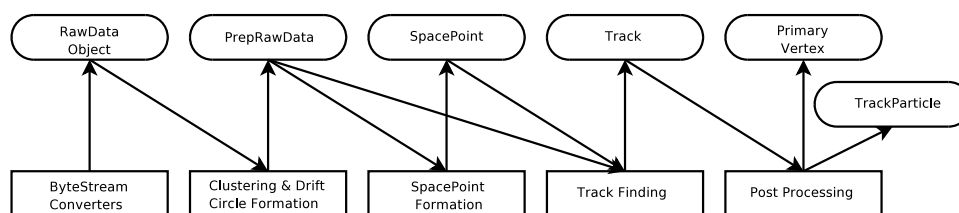


FIG. 3.6 – Chaîne de reconstruction des traces. Les boîtes de la ligne du haut représentent les objets de données, celles du bas étant les algorithmes. Les flèches représentent la direction du flux de données.

sûr par une classe commune représentant les traces mais aussi par la définition d'un jeu de paramètres commun décrivant ce qu'est une trace ainsi qu'une interface cohérente avec les différentes surfaces qu'une trace est amenée à traverser.

Une vue schématique du flux de données lors de la reconstruction des traces est reportée sur la figure 3.6. Les objets RDO sont analysés afin de construire des objets dédiés à la reconstruction, labellés *PrepRawData* (aussi appelés RIO pour *Reconstruction Input Objects*), qui contiennent généralement deux informations : l'identifiant du canal de lecture et la quantité mesurée associée.

3.4.3.2 Reconstruction des dépôts d'énergie

Toujours dans l'optique d'avoir une base de code qui puisse être partagée entre le maximum de détecteurs, il existe une classe de cellules, `CaloCell`, pour tous les calorimètres, même si le format des données brutes est spécifique à chaque calorimètre. Comme illustré par la figure 3.7, un adaptateur permet de créer des cellules indifféremment depuis la simulation ou bien les données brutes. Une fois arrivé à ce stade, tous les outils mis en jeu pour construire des amas de cellules ne requièrent que l'interface relative aux quadri-moments, ce qui simplifie considérablement la reconstruction des dépôts d'énergie (ainsi que la réutilisation de ces outils par différentes mises en œuvre d'algorithmes de regroupement de cellules). Les cellules sont donc regroupées en colonnes de cellules puis en amas d'énergie (`EnergyCluster`). Ces amas peuvent eux-mêmes être constitués d'amas : c'est une des grandes forces du système de navigation [42] qui permet d'agréger des objets pour en former un composite. En outre, ce système permet également de récupérer la liste des objets d'un certain type le composant : il est par exemple possible de demander à l'objet `Jet` de fournir l'intégralité des cellules ou des tours de cellules le composant.

Une fois que chaque sous-détecteur a publié les données reconstruites sur `StoreGate`, ces informations sont combinées pour faciliter le travail d'identification des particules.

3.4.3.3 Reconstruction combinée

La combinaison des différentes informations reconstruites par chaque sous-détecteur permet d'établir un certain nombre d'hypothèses de travail quant à l'identité des particules ayant inter-agi.

Les électrons et les photons sont modélisés par le même objet à cette étape de la reconstruction complète : `egamma`. Ces particules peuvent être reconstruites *via* deux algorithmes de reconstruction différents. L'algorithme de recherche d'électrons de haut

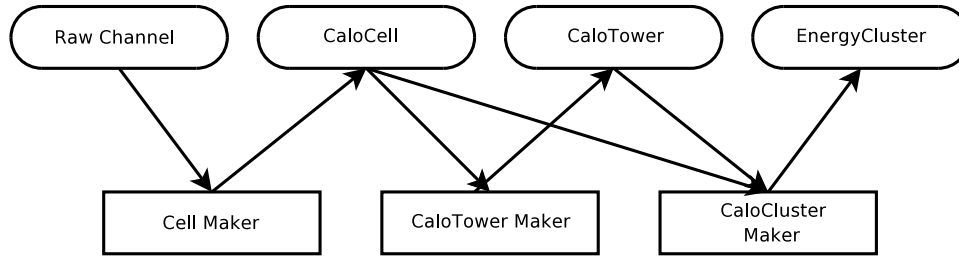


FIG. 3.7 – Chaîne de reconstruction des dépôts d'énergie. Les boîtes de la ligne du haut représentent les objets de données, celles du bas étant les algorithmes. Les flèches représentent la direction du flux de données.

p_T associe des traces à des amas d'énergie et construit les variables de profil de gerbe (fractions d'énergies déposées dans les différents compartiments du calorimètre, énergie hors-cône, énergie transverse contenue dans plusieurs couronnes en $\Delta R, \dots$). Les photons de haut p_T sont identifiés [30] de manière similaire, le seul critère additionnel étant un veto sur les traces (sauf pour les photons issus de conversions). Les électrons de bas p_T ($\leq 10 \text{ GeV}/c$) sont reconstruits à partir de l'extrapolation d'une trace chargée depuis le détecteur interne vers le système calorimétrique et du dépôt autour de cette trace. Cet algorithme est particulièrement intéressant dans le cadre de l'identification d'électrons à l'intérieur de jets, leptons “mous” qui sont ensuite utilisés par les algorithmes d'étiquetage des jets de b ⁶.

L'identification des muons est optimisée en fonction des différentes gammes en impulsion transverse. Les muons de haut p_T ($> 100 \text{ GeV}/c$) sont identifiés en extrapolant les paramètres des traces dans le spectromètre à muons jusqu'aux calorimètres, puis vers le trajectographe interne et le point d'interaction. Il est de plus possible de réaliser la combinaison entre la trace provenant du spectromètre avec celle du détecteur interne. Les muons d'impulsions transverses intermédiaires ($6 \text{ GeV}/c \leq p_T \leq 100 \text{ GeV}/c$) sont identifiés par les deux systèmes de trajectographie, alors que pour les muons de bas p_T , les traces sont reconstruites dans le détecteur interne. En effet, les muons correspondants à cette gamme d'impulsion perdent la majeure partie de leur énergie dans les calorimètres et ne traversent donc pas la totalité du spectromètre. Selon l'algorithme utilisé, l'efficacité d'identification varie de $\sim 88\%$ à 94% [30].

Le processus de reconstruction des jets peut être appliqué sur différentes sources : à partir des dépôts d'énergie dans les cellules du calorimètre ou bien directement à partir des particules Monte-Carlo. Dans le cas des jets Monte-Carlo il faut au préalable adapter les données au format attendu par les algorithmes de reconstruction de jets, ce qui permet ensuite de tester la réponse des différents algorithmes puisque la chaîne logicielle est la même. Ainsi, plusieurs types d'algorithmes sont mis à disposition par ATHENA : l'algorithme du Cône [43] (avec ou sans graine) et l'algorithme de K_T [44, 45]. Le processus de reconstruction est très modulaire puisqu'il est mis en œuvre par un simple squelette sur lequel viennent se greffer les différentes stratégies de reconstruction (les AlgTools). La reconstruction de ces jets repose essentiellement sur le système calorimétrique, cependant

⁶Notons que cette méthode d'étiquetage des jets de b est moins efficace que celle utilisant la reconstruction de vertex secondaires.

le très grand nombre de voies de lecture mis en jeu empêche de prendre les cellules comme briques de base pour les algorithmes de reconstruction. Les cellules sont donc regroupées en colonnes de cellules (`CaloTower`) et en amas (`CaloCluster`). Les tours de cellules sont un ensemble de cellules, de taille fixe en η et ϕ (généralement $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$) alors que les amas sont un regroupement de cellules qui est basé sur des critères de corrélation entre les différents dépôts d'énergie et de proximité des cellules. Le résultat de chaque algorithme de reconstruction de jets est une collection d'objets `Jets` qui, sur le même modèle que les objets `CaloEnergyCluster`, sont des objets composés (au sens de la navigation) de collections d'objets satisfaisant les interfaces "quadri-vecteur" et "navigation".

Enfin, l'énergie manquante est reconstruite à partir des muons reconstruits et de l'énergie déposée dans les cellules de tous les calorimètres. Cette mesure de l'énergie manquante est ensuite corrigée par l'énergie perdue par les particules par interaction avec le cryostat situé entre les calorimètres électromagnétique et hadronique.

Une fois les premières identifications de particules effectuées, ces données sont collectées et enregistrées dans le fichier ESD pour être ensuite utilisées par l'étape de la préparation à l'analyse qui les reformate et les étiquette.

3.4.4 Analyse

La taille des fichiers ESD⁷ étant trop importante pour que ces derniers soient distribués à grande échelle, l'étape de la préparation à l'analyse nécessite la création d'un fichier AOD ("*Analysis Object Data*") correspondant à 20 % de la taille des fichiers issus de la reconstruction. Le contenu de ce fichier AOD doit être suffisant pour la majorité des analyses de physiques. Les objets AOD sont donc construits à partir des objets ESD mais en ne gardant pas les informations détaillées des entités physiques : les cellules composant un jet restent dans le fichier ESD, seul le quadri-vecteur total caractérisant le jet est gardé au niveau AOD. Cependant si une analyse de physique a besoin d'informations plus détaillées, il est possible de remonter au niveau ESD de manière totalement transparente grâce au lien qui est sauvegardé entre ces deux fichiers ("*Back Navigation*") : chaque objet AOD garde le lien vers l'objet ESD duquel il est issu. De plus les fichiers AOD peuvent aussi bien être construits à partir des objets issus de la simulation complète (ESD) qu'à partir des objets issus de la simulation rapide (`AtlFast`), comme illustré sur la figure 3.3. Ceci permet donc de transposer facilement une analyse préliminaire réalisée grâce à `AtlFast`, dans le contexte plus détaillé de la simulation complète sans aucune modification des algorithmes de l'analyse puisque reposant sur les mêmes objets.

Les fichiers ESD sont filtrés pour construire des collections de particules identifiées à partir des objets issus de la reconstruction combinée. De nouveaux objets, spécifiques à la préparation à l'analyse, tels que les jet étiquetés (`JetTag`), sont également ajoutés aux fichiers AODs.

Ainsi, la collection d'objets Monte-Carlo est filtrée pour ne garder que les particules dites spéciales (bosons, hadrons, . . .), supposées être intéressantes dans le cadre d'analyses de physique, et rejeter toutes les autres particules produites lors des étapes intermédiaires

⁷La taille d'un fichier ESD contenant 50 événements $t\bar{t}$ produits avec la *release* 10.0.1 pour l'atelier de Physique de Rome [30] est d'environ 50 Mb.

de la génération. Chaque collection d'objets de l'étape reconstruction est convertie en une ou plusieurs collections d'objets de préparation à l'analyse (ex : les objets `egamma` sont convertis en objets `Electron` ou `Photon`). À ce stade, les définitions des différents objets sont encore très lâches afin de laisser la plus grande marge de manœuvre aux analyses en aval : seuls les électrons et photons dont la fraction d'énergie hadronique est inférieure à 20% sont gardés. Au contraire, tous les taus issus de la reconstruction combinée sont sélectionnés et convertis en taus au format AOD (`TauJet`⁸). De même chaque jet, et ce pour toutes les collections de jets produites par les différents algorithmes de reconstruction ($\Delta R = 0.4, 0.7$ et K_T), est converti en un jet au format AOD (`ParticleJet`). Ces `ParticleJets` sont ensuite analysés par l'algorithme d'étiquetage des jets de b et sont donc convertis en jets étiquetés (`JetTag`). Cette étape d'étiquetage nécessite l'information des traces pour reconstruire un éventuel vertex déplacé dû au relativement long temps de vie des hadrons b ainsi que les quadri-moments des leptons non isolés contenus dans le jet (qui pourraient être issus de désintégrations semileptoniques).

Ce filtrage d'objets ayant été réalisé, les collections de particules identifiées sont enregistrées et consistent la base standard des différentes analyses de physique. Cependant, de par l'organisation de la reconstruction combinée (exécution d'une liste d'algorithmes ayant une liste d'objets à analyser), les objets présents dans les fichiers AOD peuvent être redondants et se superposer et donc être ambigus. La reconstruction des `egamma` est effectuée par l'algorithme `egammaBuilder` qui appelle un ensemble d'`AlgTools` sur chaque amas de cellules que les algorithmes en amont ont reconstruit. La reconstruction des `Jet` opère de manière similaire : un dépôt d'énergie calorimétrique peut alors être reconstruit une première fois comme un électron puis une deuxième fois comme un jet. Un certain nombre de ces doubles comptages peut être résolu lors de l'écriture sur disque sans biaiser l'échantillon (superposition électron/photon). Néanmoins il est nécessaire de mettre à disposition une panoplie d'outils standards pour gérer ces ambiguïtés.

3.5 Outils pour l'analyse de données

L'ensemble des outils et classes développés par le groupe "*Physics Analysis Tools*" [46] (PAT) est disponible sur le CVS d'ATLAS sous le répertoire `PhysicsAnalysis`. Le groupe PAT se tient à la charnière entre la reconstruction et l'analyse à base de n-tuple. Son objectif est donc de développer, à l'intérieur du *framework* `ATHENA`, un environnement dans lequel les outils développés dans le cadre de la reconstruction sont accessibles tout en garantissant que les outils d'analyses basés sur les n-tuples (qui peuvent être produits à partir des fichiers AOD) soient également utilisables. Ces outils doivent de plus être disponibles tant dans un environnement de travail en *batch* que lors d'analyses en interactif. L'environnement pour l'analyse dans le cadre d'`ATHENA` repose sur une hiérarchie de classes représentant les données de physique qui sont ensuite manipulées par un ensemble d'algorithmes dont le travail est facilité par un jeu d'outils. En tant que membre du groupe PAT, j'ai développé un ensemble de classes facilitant l'analyse, et qui seront décrites dans la suite.

⁸La caractérisation du tau (désintégration en 1 ou 3 "*prongs*") est laissée à la charge des analyses en aval.

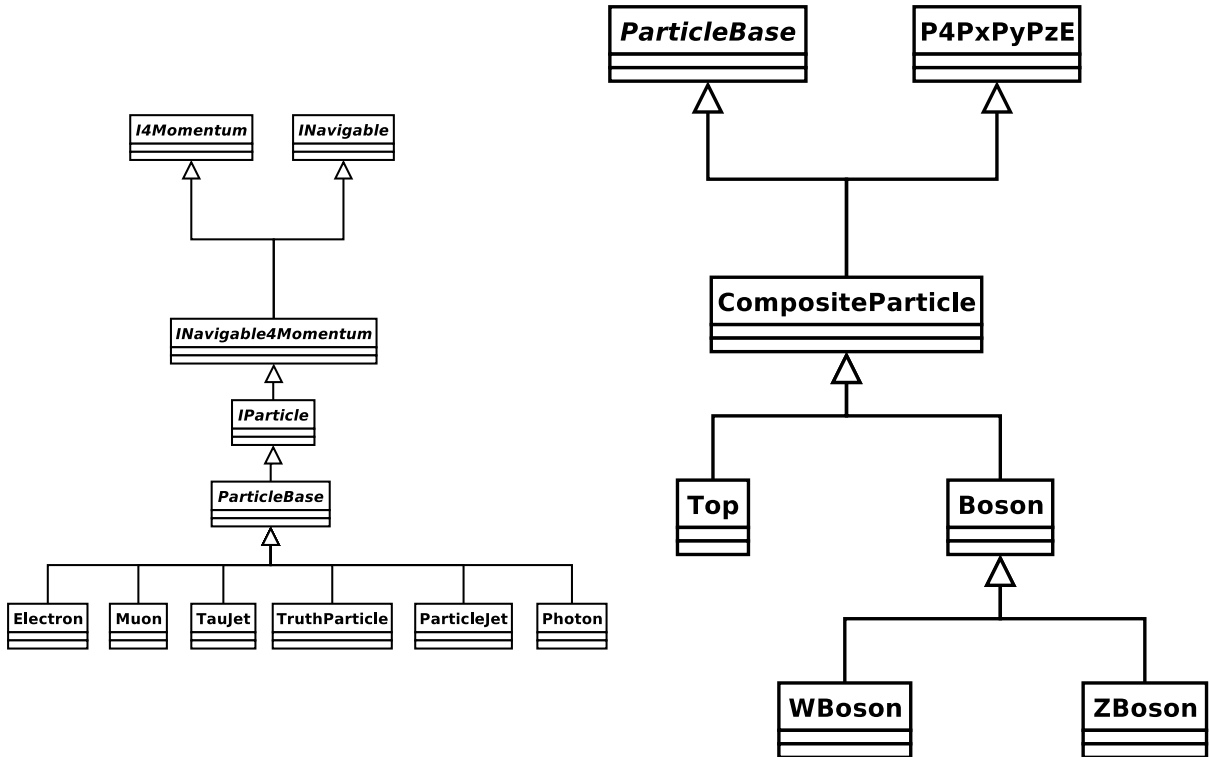


FIG. 3.8 – Diagrammes UML de l'arbre d'héritage des classes d'analyse : à gauche les classes de base et à droite les classes dédiées à la reconstruction de chaînes de désintégration. Les classes dont le nom est en italique sont abstraites et celles dont le nom commence par un I sont virtuelles [21].

3.5.1 Classes pour l'analyse

Les classes manipulées au niveau de l'analyse héritent toutes d'une classe commune : `INavigable4Momentum`. Ceci permet d'exploiter le polymorphisme du C++ tant dans les classes de données que les algorithmes. En effet, comme illustré sur la partie gauche de la figure 3.8, la classe interface `INavigable4Momentum` hérite de deux classes interfaces : `I4Momentum` et `INavigable`. La classe `I4Momentum` permet de déclarer l'interface relative à l'accès aux données cinématiques tout en cachant son implémentation, ce qui permet d'utiliser le jeu de composantes le plus approprié ($E, \eta, \phi, m; p_x, p_y, p_z, e; \dots$) en fonction du type d'objet (`Electron, ParticleJet, \dots`) considéré. Quant à l'interface `INavigable`, elle permet de modéliser le fait que les objets reconstruits sont des compositions d'objets plus fondamentaux du détecteur dont le type est *a priori* inconnu (ou pouvant varier d'un objet reconstruit à l'autre). Cette interface permet donc de composer un objet à partir de plusieurs objets et de leur attribuer un poids qui sera pris en compte lors du calcul du quadri-moment. De plus chaque constituant d'un objet héritant de `INavigable` peut lui même être composé d'autres objets (un `Electron` est composé de `Clusters` et de `Tracks`, les `Clusters` étant eux-même des amas de `Cells`). Enfin, `INavigable` permet de récupérer la liste d'objets d'un certain type (`Cell`) constituant un objet honorant cette interface (`Electron`).

Une autre interface est insérée entre les classes d'analyse et l'interface pour la navigation : `IParticle`. Cette interface ajoute la gestion des informations caractérisant une

particule : son identification (telle que définie par le *Particle Data Group*), sa charge, son vertex de production et le type de données à partir desquelles elle a été reconstruite (simulation rapide, simulation complète, Monte-Carlo, ...). La classe `ParticleBase`, que j'ai introduite dans le cadre du travail effectué au cours du portage du code de l'analyse (voir chapitre suivant) au sein du *framework* `ATHENA`, factorise un certain nombre de méthodes dont l'implémentation est la même quel que soit le type d'objet considéré.

D'autre part, la classe `CompositeParticle` modélise une particule reconstruite à partir de ses produits de désintégration. Plusieurs spécialisations de cette classe ont été implémentées dans le paquet `CompositeParticleEvent` afin de faciliter l'analyse à partir des AODs au sein du groupe `Top`. Comme illustré sur la partie droite de la figure 3.8, on retrouve donc un objet `Top` ainsi que les classes `Boson`, `ZBoson` et `WBoson` que j'ai introduites dans le cadre de cette analyse.

3.5.2 Outils pour l'analyse

Les outils pour l'analyse des données au niveau AOD sont regroupés dans les paquets `AnalysisUtils` et `AnalysisTools`. J'ai écrit une grande partie de ces classes. Le paquet `AnalysisTools` est essentiellement l'emballage sous forme d'`AlgTool` de la plupart des fonctionnalités mises à disposition par le paquet `AnalysisUtils`. Ces paquets regroupent donc des outils pour combiner les éléments d'une collection et former une nouvelle collection, pour réaliser toutes les permutations possibles à l'intérieur d'une collection donnée et d'autres outils pour classer les éléments d'une collection en fonction de leur énergie, de leur η ou de leur impulsion transverse.

Il existe également une architecture de classes de filtres afin de réaliser des coupures sur différentes propriétés accessibles *via* l'interface `IParticle` (voir paragraphe 3.6).

3.5.3 Algorithmes d'analyse

Il existe deux classes d'algorithmes mis à disposition par le groupe PAT. Il y a tout d'abord les algorithmes utilisés lors de la création des fichiers AOD qui sont eux-mêmes sous-divisés en deux catégories : ceux utilisés dans le contexte de la reconstruction complète et ceux utilisés dans le contexte de la simulation rapide `AtlFast`. Enfin, un certain nombre d'algorithmes d'exemples ont été écrits et regroupés dans le paquet `AnalysisExamples`, contenant ainsi mes propres exemples relatifs aux classes que j'ai développées. De plus, plusieurs groupes de physique ont également écrit leurs propres algorithmes d'intérêt général.

3.6 Outils de filtrage

Toutes les analyses de physique ont en commun de nécessiter de filtrer et sélectionner certains éléments parmi des collections d'objets. J'ai donc développé des classes facilitant cette sélection. Ce travail de filtrage s'effectue sur la base de critères cinématiques mais peut également être réalisé sur la base de chaînes de désintégrations au niveau Monte-Carlo ou reconstruit.

3.6.1 Prédicats de sélection

La figure 3.9 représente l'arbre d'héritage des classes de filtrage mises à disposition au niveau AOD. La classe mère `IFilterCuts` (cf ma présentation [48]) est une simple classe permettant aux classes filles de copier le jeu de paramètres de coupures entre différentes implémentations de classes de filtrage, grâce à la présence d'une fonction purement virtuelle. L'arbre d'héritage se scinde alors en deux branches, l'une pour les classes purement Monte-Carlo et l'autre pour les données reconstruites. Cependant, toutes les classes de sélection implémentent la méthode `isAccepted(const T*)` qui retourne la décision du filtre et peuvent alors être réutilisées par l'environnement logiciel de sélection, de combinaison et de permutation qui se trouve dans le paquet `AnalysisUtils`. De plus, la surcharge de l'opérateur "operator()" permet d'utiliser ces filtres comme prédicat dans tous les algorithmes de la STL ("Standard Template Library" [51]) du langage C++, ce qui autorise une syntaxe plus concise et plus expressive.

Les classes de filtrage dédiées aux données reconstruites héritent toutes de la classe `IFilter<T>`. La généricité introduite par cette classe (l'argument "template") permet aux classes filles de définir des critères de sélection qui s'appuient sur l'interface de l'argument générique sans avoir à convertir son argument depuis une classe mère vers la classe fille. On peut ainsi écrire un filtre dédié aux électrons et qui s'appuie donc sur les données membres spécifiques à cette classe sans avoir à convertir un pointeur vers une `IParticle` en un pointeur vers un `Electron`, opération qui peut pénaliser les performances du filtre.

On retrouve donc quatre classes qui héritent de `IFilter<T>`. `MomentumFilter<T>` permet de définir des coupures minimum et/ou maximum sur toutes les données cinématiques d'une particule (impulsion, masse, énergie, ...). `PdgIdFilter<T>` permet de sélectionner les particules *via* leur numéro d'identification comme défini par le *Particle Data Group* [47]. `ChargeFilter<T>` permet de sélectionner les particules par rapport à leur charge mesurée. `CategoryIdFilter<T>` permet quant à elle de sélectionner les particules par rapport à une catégorie de particules (hadrons, baryons, bosons, leptons, ...). Par soucis de commodités, une classe (`IIParticleFilter<T>`) hérite de ces quatre classes de sélection et combine, *via* un OU logique, la décision de chacune de ses classes mères. Enfin, la classe `IParticleFilter` résoud l'argument de généricité au type `IParticle`, ce qui permet à la fois de réduire l'exposition directe aux "templates" mais aussi d'avoir un outil de filtrage polyvalent au niveau de l'interface `IParticle`.

L'extrait de code de la figure 3.10 illustre la simplicité avec laquelle un filtre peut être configuré. Ce bout de code qui copie les particules contenant un quark *b* ou étant étiquetées *b* ayant une énergie d'au moins 40 GeV peut être appliqué indifféremment sur des objets reconstruits ou Monte-Carlo.

3.6.2 Filtrage de vertex

La classe `McVtxFilter` met en œuvre plusieurs classes outils afin d'exprimer des critères de sélection de vertex. Cette classe a été présentée au cours de différents ateliers "Software" [48–50]. Ce critère de sélection se construit au moyen d'une chaîne de caractères passée en argument décrivant le motif (*pattern*) de vertex recherché. Ainsi les chaînes "6->" et "->-11+11" permettent respectivement de sélectionner les vertex de désintégration de quark top (labellé par 6) et les vertex de désintégration en paire d'électrons (l'électron et son anti-particule étant respectivement labellés par -11 et 11).

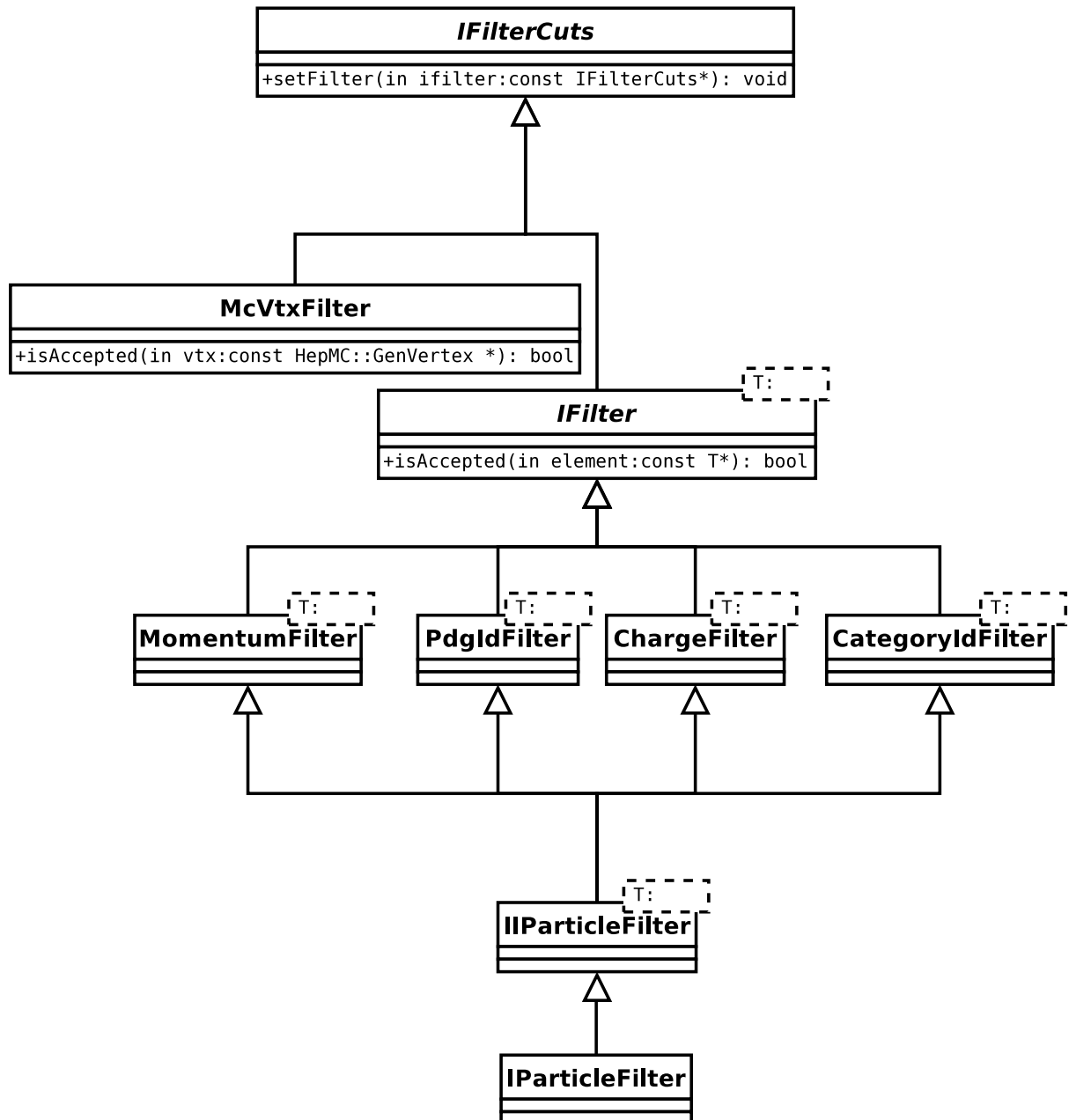


FIG. 3.9 – Diagramme UML des classes de filtrages. Toutes les classes implémentent la méthode `isAccepted()` qui réalise l'opération de filtrage.

```

1 #include "AnalysisUtils/IParticleFilter.h"
2 IParticleFilter filter;
3 filter.setCategory( PDGCat::BFlavour );
4 filter.setEneMin( 40.*GeV );
5 std::vector<const IParticle*> myBs;
6
7 AnalysisUtils::copy_if( parts.begin(), parts.end(),
8                         filter,
9                         std::back_inserter(myBs) );

```

FIG. 3.10 – Exemple de code illustrant la simplicité des filtres de particules : seuls les éléments du vecteur `parts` contenant un quark b avec une énergie d'au moins 40 GeV seront copiés dans le vecteur `myBs`. Ce filtre pourra être exécuté indifféremment sur des jets ou des particules Monte-Carlo.

En interne, `McVtxFilter` regarde si chaque demi-vertex (entrant ou sortant) d'un objet `HepMC::GenVertex` satisfait aux conditions exprimées par le motif de désintégration. Plusieurs modes de fonctionnement sont alors disponibles. L'utilisateur peut demander un strict appariement du nombre de pattes présentes au vertex : un vertex $Z^0 \rightarrow e^+e^-\gamma$ ne serait alors pas accepté par le deuxième motif donné précédemment en exemple. Le filtre peut aussi être configuré de manière à ce que la distinction entre particules et anti-particules ne soit pas faite : le premier motif permettrait alors de sélectionner les désintégrations de quarks top et anti-top.

La classe `McVtxFilter` est ensuite utilisée par la classe `McVtxFilterTool`. Cet `AlgTool` contient une liste de ces filtres qui peuvent être configurés indépendamment les uns des autres au niveau du `jobOption`. Ceci permet donc de remplir facilement un conteneur de particules Monte-Carlo satisfaisant des motifs de désintégration sans avoir à compiler une seule ligne de code : il suffit que l'utilisateur intègre à son algorithme un pointeur vers cet `AlgTool` pour avoir accès à ces possibilités de filtrage. L'extrait de `jobOption` de la figure 3.11 illustre la configuration d'un tel algorithme.

3.7 Outils d'associations

Le paquet `AnalysisAssociation` que j'ai développé regroupe les classes et outils nécessaires à la représentation d'une association entre deux objets ainsi qu'à l'établissement de cette association. La classe `AodAssocs` permet ainsi d'établir et d'enregistrer un lien entre deux objets dérivant de la classe d'interface `INavigable4Momentum`. Ceci permet donc d'associer un objet (ou une collection d'objets) `TrackParticle` à un objet `ParticleJet`, mais il est aussi possible d'associer objets reconstruits avec objets Monte-Carlo (`TruthParticle`). Ce conteneur d'associations devrait à terme être utilisé lors de la construction des AOD à partir des ESD afin d'étiqueter les éventuelles superpositions d'objets (electron-jet, electron-tau,...) pour en faciliter la gestion en aval par les algorithmes d'analyse. `AodAssocs` fonctionne à la manière d'une `std::map<key,value>` : une collection de paires clefs-valeurs (qui sont des pointeurs persistants entre `INavigable4Momentum`) est enregistrée. Il est ainsi possible de récupérer tous les objets associés à un

```

1 # Filter which looks for a Z0 decaying into a b and a bbar
2 MyAlg.McVtxFilterTool.DecayPatterns = [ "23 -> -5 + 5" ]
3
4 # Filter looking for a Z0 decaying into an e+e- or muon pair
5 MyAlg.McVtxFilterTool.DecayPatterns += [ "23 -> -11|-13 +
6     11|13" ]
7
8 # Filter looking for a Z0 decaying either into anything
9 MyAlg.McVtxFilterTool.DecayPatterns += [ "23 ->" ]
10
11 # Filter looking for anything decaying into an e+e- pair
12 MyAlg.McVtxFilterTool.DecayPatterns += [ "-> -11 + 11" ]
13
14 # this is the decay pattern for a t->W+ b
15 MyAlg.McVtxFilterTool.DecayPatterns += [ "6 ->24 + -5" ]

```

FIG. 3.11 – Exemple de *jobOption* illustrant la configuration d'un filtre Monte-Carlo sélectionnant différentes particules satisfaisant des motifs de désintégration. Des *patterns* peuvent être rajoutés ou enlevés sans nécessiter de recompilation du code d'analyse.

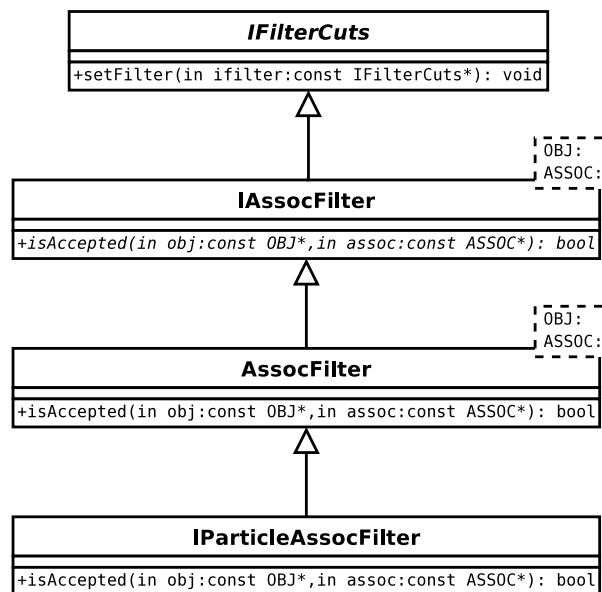


FIG. 3.12 – Diagramme UML des classes de filtrage d'associations. Toutes les classes implémentent la méthode `isAccepted()` qui réalise l'opération de filtrage.

objet donné.

Les classes de sélection d'associations sont construites sur le même modèle que les prédicats de filtrage. Elles héritent toutes de la classe `IFilterCuts` mais prennent deux arguments génériques lors de l'appel à la fonction de sélection : `isAccepted(const*OBJ, const*ASSOC)`. Divers critères de cinématique, sur le même modèle que `MomentumFilter`, peuvent être exprimés : plages en ΔR , ΔP_T , ...

3.8 Réorganisation de l'implémentation du Monte Carlo

Les différents outils développés par le groupe PAT ont tous été écrits pour des objets héritant de la classe `INavigable4Momentum`. Ceci a permis de se concentrer sur les fonctionnalités des outils plutôt que de s'attarder sur les spécificités de chaque objet que les outils auraient eu à traiter. Ainsi la classe modélisant les particules Monte-Carlo au niveau AOD, `TruthParticle`, suit le même schéma d'héritage. Cependant, la vérité Monte-Carlo est modélisée par le *package* `HepMC` dans le reste de l'environnement logiciel `ATHENA`. J'ai pu résoudre cette tension entre les deux modélisations, en construisant la classe `TruthParticle` autour de la classe `HepMC::GenParticle` qui est le format de la plupart des générateurs Monte-Carlo. Le code développé a été introduit dans le code officiel. La classe AOD devient alors une simple coquille autour de `HepMC::GenParticle` et permet alors de réutiliser toutes les méthodes du paquet `HepMC` tout en présentant l'interface typique des objets AOD. Le résultat net de cette opération est donc d'offrir le meilleur des deux implémentations aux utilisateurs des classes AOD.

Nous avons donc fait le tour des outils développés par le groupe PAT ainsi que de l'architecture des classes mises à disposition pour l'analyse. A cette étape de l'analyse, l'utilisateur a le choix de créer son n-tuple de données reconstruites et de continuer son analyse sur ce support (et donc de sortir de l'environnement `ATHENA`) ou bien de rester dans le *framework*.

3.9 Au-delà de l'analyse avec des n-tuples

L'objectif du groupe PAT est de fournir des outils afin de faciliter le maniement ainsi que l'analyse des données de physique. Le modèle en vigueur d'analyse des données en physique des particules consiste généralement à extraire des fichiers issus de la reconstruction complète, un échantillon réduit d'événements intéressants pour une analyse donnée puis de définir un jeu de variables qui rempliront des n-tuples. Le groupe PAT est en train de tester un nouveau concept et une nouvelle étape dans le flot d'analyse orienté objet : l'`EventView` [52]. J'ai participé aux prémices de cette classe [53, 54].

Cet objet a pour but de présenter à l'utilisateur une vue cohérente et exhaustive d'une liste d'objets physiques mutuellement exclusifs [52]. Ceci peut se traduire par le fait qu'à ce stade de l'analyse, toutes les ambiguïtés seraient levées et l'identification des particules figées et enregistrées dans l'`EventView`. Cet objet constitue l'élément de base d'une analyse, dans lequel l'utilisateur peut stocker des variables relatives à la globalité de l'événement (sphéricité, balancement en impulsion transverse, ...) mais également d'étiqueter des particules reconstruites ou composites afin d'en faciliter l'accès au cours de l'analyse. Cette classe permet alors d'enregistrer l'état de l'analyse et de facilement passer cet état d'algorithmes en algorithmes.

Cet objet a été mis en production [55] pour l'atelier de Physique de Rome et le retour des utilisateurs a été très nettement favorable [56].

3.10 Grilles de calcul

L'architecture du modèle de traitement des données d'ATLAS est très fortement basé sur l'utilisation des grilles de calcul. Ces grilles constituent en effet un moyen de mettre à disposition de l'ensemble de la collaboration ATLAS, les ressources de calcul disséminées dans les différents laboratoires de la collaboration. Ces ressources de calcul seront utilisées pour le traitement des données brutes (construction des fichiers ESD et AOD) et leur stockage. Cependant, il est également prévu que les physiciens d'ATLAS puissent soumettre des exécutables de leur analyse sur ces grilles.

Ce système de grilles a été testé dans le cadre de la production des données pour l'atelier de physique de Rome. J'ai ensuite développé un ensemble de scripts [57] qui a permis de retraiter [58] la totalité du lot 4100 afin de récupérer les informations Monte-Carlo nécessaires à l'étalonnage des jets légers (les quarks directement issus du W , avant toute radiation) et par la même occasion, d'ajouter le conteneur de jets étiquetés b reconstruits avec l'algorithme de cône ($\Delta R = 0.4$). En effet, ces informations manquaient dans l'échantillon de fichiers AOD.

3.11 Conclusion

Au cours de ce chapitre, nous avons présenté la plateforme logicielle de la collaboration ATLAS, appelée ATHENA. C'est à la fois un système permettant de gérer le flux de données provenant du système d'acquisition mais également l'environnement à l'intérieur et à l'aide duquel les physiciens analysent les données simulées ou réelles (faisceaux tests pour l'instant). L'analyse présentée dans le chapitre suivant utilise cette plateforme logicielle, les outils que j'ai développés au sein d'ATHENA ainsi que les scripts écrits pour faciliter l'envoi de codes de calculs sur la grille.

Chapitre 4

Etalonnage en énergie des jets

La connaissance de l'échelle absolue en énergie des jets est une des principales problématiques des expériences auprès des collisionneurs. Comme la plupart des canaux étudiés comportera un ou plusieurs jets dans l'état final, la précision sur la mesure de l'énergie des jets déterminera la précision attendue sur les paramètres fondamentaux du Modèle Standard. La mesure de la masse du quark *top* en est un exemple flagrant. La connaissance précise de cette observable permettra de tester la cohérence du Modèle Standard et ainsi d'appliquer des contraintes supplémentaires sur ses autres paramètres libres (*ex* : la masse du boson de Higgs). La collaboration ATLAS s'est donnée pour objectif d'étalonner l'échelle en énergie des jets à une précision de l'ordre du pourcent.

Les quarks et gluons produits lors de collisions, appelés dans la suite partons initiaux, ne seront pas observables directement puisque soumis au confinement de l'interaction forte. Ce sont les hadrons issus de la fragmentation des partons puis de leur hadronisation, que l'on détecte sous forme de jets. La mesure de l'énergie du parton initial est réalisée grâce aux dépôts d'énergie des hadrons dans les calorimètres. Cependant, de nombreux effets sont à prendre en compte pour déterminer l'énergie d'un parton initial à partir des jets reconstruits avec les amas calorimétriques. Il y a tout d'abord les effets physiques. Lors du processus de fragmentation du parton initial, l'information sur l'énergie va être diluée parmi tous les partons de la cascade. Il faudrait donc apparier et combiner tous ces partons (sans inclure les contributions d'autres cascades de partons) pour remonter à l'énergie initiale. Il est également possible que le parton initial émette un gluon de haute énergie (*Final State Radiation*) qui donnera lui-même naissance à un second jet. Il faut aussi prendre en compte l'effet des événements sous-jacents (*Underlying Event*) ainsi que les interactions multiples qui vont rendre plus difficile la compréhension de l'événement. A cela s'ajoutent les effets d'instrumentation. Une fois formé et collimé par la poussée de Lorentz, le jet de hadrons va interagir avec le calorimètre. Le contenu en particules de ce jet explique le fait que la gerbe développée présente une composante électromagnétique (*via* les pions neutres qui se désintègrent en paires de photons) et une composante hadronique. Cette particularité constitue un écueil pour les calorimètres non compensés (tel que celui d'ATLAS) qui ne répondent pas de la même manière à chacune de ces composantes. Seules les fractions *e* et *h* de l'énergie des composantes électromagnétique et hadronique seront mesurées par le calorimètre dont la réponse peut être exprimée par :

$$R = e \times E_e + h \times E_h \quad (4.1)$$

Le rapport *e/h* caractérise ce que l'on appelle la compensation. Lorsque ce rapport est

différent de 1, le calorimètre est dit non-compensé. Pour le Tilecal, la valeur trouvée en faisceau test [59] est de $\frac{\epsilon}{h} = 1.37 \pm 0.03$. Cependant, cette valeur doit être réexaminée car de nouveaux résultats obtenus avec le test combinée de 2004 [60] sont en désaccord.

De plus, non seulement la composante hadronique fluctue de gerbe à gerbe mais une fraction de cette énergie ne sera pas détectée. Cette énergie invisible provient de la désexcitation du milieu actif du calorimètre (énergie de recul des noyaux ayant interagi avec la particule incidente, neutrinos émis lors de la désexcitation des noyaux, ...). Par ailleurs, les zones non-instrumentées du calorimètre, les effets de non-linéarité ou les fuites d'énergie (si la gerbe n'est pas totalement contenue) vont dégrader la résolution sur l'énergie mesurée. Enfin, les algorithmes de reconstruction des jets peuvent également introduire un biais en réalisant un mauvais appariement des cellules en amas. Il faut ainsi appliquer des coefficients de correction aux dépôts d'énergie dans les cellules et amas du calorimètre pour prendre en compte ces effets. Cette tâche est remplie par les algorithmes d'étalonnage, après avoir appliqué les facteurs de correction obtenus en faisceau test et avec les cosmiques (phase dite de "*commissioning*" [61]).

Le but de l'étalonnage en énergie des jets est donc, à partir de la mesure d'un certain dépôt en énergie, de remonter à l'énergie du quark initial. Il faut ici différencier l'étalonnage relatif, c'est à dire la dispersion de la mesure autour de la valeur moyenne, de l'étalonnage absolu, c'est à dire la valeur du pic par rapport à la valeur attendue. La mesure de l'échelle absolue en énergie est habituellement obtenue en faisceau test en déterminant la réponse en énergie du détecteur à un faisceau monoénergétique de particules. Cependant cette méthode ne peut pas être transposée aux jets car il n'y a pas de dispositif permettant d'obtenir des jets en faisceau test. D'autre part cette méthode n'inclut pas les effets des événements sous-jacents et des empilements d'événements. Une stratégie possible est donc d'utiliser les simulations numériques telles que **Geant4**, qui devront être ajustées aux données réelles lorsqu'elles seront disponibles. Cela nécessite une très bonne modélisation tant des processus physiques (modèles phénoménologiques de fragmentation des partons tels que le modèle des cordes de Lund [31]) que de la réponse du détecteur (nécessité d'une simulation correcte des gerbes hadroniques). Une autre approche est donc d'étalonner la mesure de l'échelle en énergie des jets directement avec les données. Dans cette optique, on recherche les dépôts d'énergie de particules identifiées dans le calorimètre que l'on compare avec l'énergie attendue. Cette énergie attendue est fournie soit par un autre détecteur plus précis ou bien par une contrainte, en masse ou en impulsion.

Par exemple, l'échelle électromagnétique est déterminée en utilisant la désintégration du J/Ψ [62] ou du boson Z^0 [63] en une paire e^+e^- . Les coefficients d'étalonnage sont alors obtenus en demandant que la masse invariante de la paire d'électrons soit égale à la masse de la résonance. La précision sur l'échelle électromagnétique est limitée dans ce cas par la largeur de la résonance. Dans le cas d'un jet, la particule initiale est un parton et le processus $W \rightarrow qq'$ peut être utilisé pour réaliser l'étalonnage ¹.

Par ailleurs, les deux grands types d'effets, ceux directement dûs à la physique et ceux provenant de l'instrumentation, donnent lieu à deux références pour l'énergie des

¹Le processus $Z \rightarrow q\bar{q}$ pourrait également être utilisé. Les principaux problèmes résident en l'obtention d'un échantillon pur de ce type de désintégrations ainsi qu'en l'établissement d'une stratégie de déclenchement.

jets. On peut en effet prendre comme référence soit l'énergie du parton initial et donc la fonction d'étalonnage que l'on calculera intègrera la chaîne complète du parton au jet, soit l'énergie combinée des hadrons constituant le jet ². Dans ce cas, on ne corrigera que des effets purement instrumentaux précédemment évoqués.

4.1 Différentes méthodes d'étalonnage

Il existe plusieurs méthodes d'étalonnage permettant de corriger des biais rencontrés lors de la mesure. Chacune essaye de recouvrir la gamme la plus large en énergie afin d'obtenir un étalonnage sur tout le spectre. De plus, il est aussi nécessaire de faire varier au maximum les hypothèses de départ des diverses méthodes pour estimer les systématiques dont elles sont entâchées.

4.1.1 Étalonage en faisceau test

La collaboration ATLAS a engagé un programme intensif de tests en utilisant le faisceau *H8* au CERN. Ces tests permettent de valider les performances attendues des différents sous-détecteurs d'ATLAS avec des données réelles. Ainsi, environ 1/8 du calorimètre à Argon liquide et du calorimètre à tuiles scintillantes a été soumis à des faisceaux d'électrons et de pions d'énergie allant jusqu'à 300 *GeV* et de muons de 20 à 150 *GeV*. Le dernier test combiné s'est déroulé en 2004 [64–66].

La reconstruction de ces objets permet d'estimer la résolution et la linéarité de la réponse des détecteurs attendues lors du démarrage d'ATLAS et pendant la prise de données. Deux méthodes ont été étudiées pour reconstruire l'énergie des particules incidentes. La méthode *Benchmark* [67, 68] a été utilisée pour reconstruire l'énergie de faisceau de pions en corrigeant l'énergie mesurée par un jeu de facteurs indépendants de l'énergie. Une autre méthode, inspirée par la collaboration H1 [69], repose sur le fait que les gerbes électromagnétiques sont plus denses que les gerbes hadroniques. De ce fait, les cellules correspondant à un faible signal sont corrigées individuellement de manière à égaler celles contenant un signal plus important (dépôt de la partie électromagnétique de la gerbe). Les échelles en énergies obtenues avec ces deux méthodes avec un faisceau de pions de 300 *GeV* ont été étudiées dans [67]. Les rapports énergie incidente sur énergie mesurée après étalonnage sont alors de 5% et 2% pour la méthode *Benchmark* et l'étalonnage à la H1, respectivement.

Cependant, l'énergie des lignes de faisceau ne permet pas d'atteindre la gamme en énergie des jets lors de la prise de données. La seule quantité que l'on puisse extraire de ces mesures est une échelle en énergie hadronique. Seules les méthodes *in situ* peuvent donner des informations pour la gamme en énergie des jets auprès du LHC.

4.1.2 Étalonage *in situ*

Le principe de l'étalonnage *in situ* est d'utiliser des événements enregistrés pendant la prise de données pour calculer les poids à appliquer sur les dépôts d'énergie pour remonter au plus près de l'énergie réelle.

²Cette entité est appelée jet Monte-Carlo.

4.1.2.1 E/p

La méthode E/p est basée sur la détection de hadrons chargés énergétiques et isolés (π^\pm et K^\pm). Le rapport de l'énergie mesurée par les calorimètres sur l'impulsion mesurée par les trajectographes internes doit être égal à 1. Cette méthode permet d'inter-étalonner les sous-détecteurs de technologies différentes et d'obtenir l'échelle absolue en énergie des hadrons avec une précision au niveau du pourcent [70]. De plus, les coefficients d'étalonnage obtenus en faisceau test peuvent être directement appliqués lors de la prise de données et ainsi transférer les échelles d'énergie obtenues en faisceau test à la configuration *à la ATLAS*. Cependant cette méthode ne peut fournir que l'échelle en énergie des hadrons chargés isolés et il faut recourir à une autre technique afin d'extraire celle des jets.

4.1.2.2 Balance en impulsion transverse

Cette méthode s'appuie sur la production de deux objets dos-à-dos. Dans un collisionneur, l'impulsion transverse dans l'état initial est en moyenne nulle³. Ainsi, nous pouvons utiliser la mesure précise du quadri-moment d'une des deux particules pour déduire les facteurs de correction (les constantes d'étalonnage) à appliquer sur les composantes de la seconde particule afin de rétablir la balance en p_T . Le canal de prédilection dans ATLAS est celui des événements $Z^0 + jet$ dans lesquels le Z^0 se désintègre en électrons ou en muons :

$$p + p \rightarrow Z^0 + q(g) + X \rightarrow l^+l^- + jet + X$$

Le déclenchement du système d'acquisition se fait alors grâce à la paire de leptons de haute impulsion transverse dont la masse invariante reproduit celle du boson Z^0 . L'étalonnage [71, 72] est alors réalisé en demandant que :

$$p_T^{jet} = p_T^{Z^0} \quad (4.2)$$

Il est intéressant de noter qu'*a priori* ce canal permet d'étalonner n'importe quel type de jet (issu de quark b ou de quark léger) avec une précision sur l'échelle en énergie de 1% [72]. Il est également possible d'utiliser un photon à la place des leptons issus de la désintégration du Z^0 . Cependant, malgré la plus grande statistique disponible, ce canal est moins prometteur car souffrant de bruits de fond mal contrôlés (π^0 et photons isolés parmi le bruit de fond QCD).

4.1.2.3 Résonance

Le principe de cette méthode d'étalonnage repose sur la reconstruction d'une résonance dont on connaît suffisamment bien les propriétés intrinsèques. Une fois la résonance reconstruite à partir de ses produits de désintégration, on peut appliquer les facteurs de correction *ad hoc* sur les quadri-moments des états finals pour recalibrer le pic de la résonance à la valeur attendue.

³Dans un collisionneur de protons tel que le LHC, l'impulsion transverse dans l'état initial n'est pas strictement nulle car les partons interagissant sont soumis au mouvement de Fermi à l'intérieur du nucléon. Le référentiel du centre de masse n'est donc pas celui du laboratoire.

Dans la suite nous allons nous intéresser plus particulièrement à l'étalonnage de l'échelle en énergie des jets légers en utilisant la désintégration du boson W en une paire de jets et à l'étalonnage des jets de b grâce à la désintégration du Z^0 en une paire de jets de b .

4.2 Etalonnage des jets légers *via* la résonance du boson W

Ce paragraphe décrit la démarche utilisée afin d'extraire les coefficients d'étalonnage en énergie des jets légers. Nous verrons dans un premier temps comment extraire un lot très pur de bosons W à partir du processus $t\bar{t}$. La méthode générale d'extraction des coefficients de correction sera ensuite présentée.

4.2.1 Sélection des événements

La méthode d'étalonnage que nous décrirons dans la suite, repose sur la reconstruction du boson W . Il faut donc avoir à disposition un lot très pur de W afin de ne pas biaiser la fonction d'étalonnage.

Cette méthode d'étalonnage utilise la production de paires $t\bar{t}$, dont des diagrammes de Feynman sont représentés sur la figure 4.1, se désintégrant dans le canal semi-leptonique suivant :

$$t\bar{t} \rightarrow Wb + W\bar{b} \rightarrow l\nu_l b + j\bar{j}\bar{b}$$

où l'un des bosons W se désintègre leptoniquement et l'autre en une paire de quarks. La section efficace de ce processus calculée avec *Pythia* 6.203 [76, 77] est de 34.8 ± 0.1 pb. Cependant un calcul à l'ordre suivant [78] donne $\sigma_{t\bar{t}} = 833 * 29.6\% = 246$ pb. Ceci laisse espérer la production de presque 2.5 millions d'événements $t\bar{t}$ (soit environ une dizaine de paires par minute) dans le canal semi-leptonique après un an de prise de données à basse luminosité ($\mathcal{L} = 10^{33} \text{cm}^{-2} \text{s}^{-1}$) correspondant à une luminosité intégrée de 10fb^{-1} .

Ce canal a été choisi parmi les autres possibilités (canaux tout leptonique et tout hadronique) car il permet de reconstruire directement le top du côté hadronique tout en gardant une signature claire grâce au côté leptonique sur lequel le système de déclenchement va s'appuyer (lepton isolé de haute impulsion transverse). L'extraction du canal tout hadronique est en effet rendue difficile par le bruit de fond QCD attendu, le rapport signal sur bruit étant d'environ 10^{-9} à la production [73]. Le canal tout leptonique est bien entendu inutile pour l'étalonnage des jets légers. Le canal semi-leptonique a alors été retenu pour cette étude. Les bruits de fond suivants :

- $b\bar{b} \rightarrow l\nu + jets$
- $W + jets \rightarrow l\nu + jets$
- $Z + jets \rightarrow l^+l^- + jets$
- $W + W \rightarrow l\nu + jets$
- $W + Z \rightarrow l\nu + jets$
- $Z + Z \rightarrow l^+l^- + jets$
- $Wb\bar{b} \rightarrow l\nu + b\bar{b}$

ont été étudiés dans un travail précédent [74] dans lequel il a été estimé que le rapport signal sur bruit pouvait être amené de $S/B \simeq 10^{-5}$ à 65 après application du jeu de

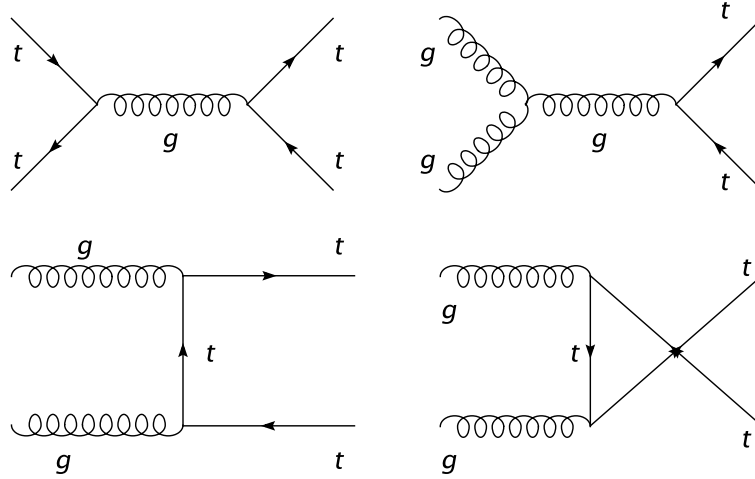


FIG. 4.1 – Diagrammes de Feynman de production de paires $t\bar{t}$. Au LHC, environ 90% des paires $t\bar{t}$ seront produites par fusion de gluons.

coupages adéquat. Nous ne nous intéresserons donc dans la suite de ce document qu'au processus $t\bar{t}$, le bruit de fond étant considéré comme négligeable après ces sélections. Il convient également de noter que le système de déclenchement n'a pas été inclus dans cette analyse car les menus du dernier niveau de déclenchement (les menus dédiés à la physique), reportés dans le tableau 4.1, laissent une confortable marge de manœuvre par rapport aux coupures utilisées dans la suite. En effet, les événements sont sélectionnés en demandant :

- un lepton isolé (e^\pm, μ^\pm) de $p_T > 20 \text{ GeV}/c$ et $|\eta| < 2.5$ (Sel1),
- une énergie transverse manquante $E_T > 20 \text{ GeV}$ (Sel2),
- au moins quatre jets de $p_T > 40 \text{ GeV}/c$ et $|\eta| < 2.5$ (Sel3),
- parmi ces quatre jets, un ou deux b -jet avec une impulsion transverse $p_T > 40 \text{ GeV}/c$ (resp. Sel4 ou Sel5).

Les efficacités obtenues avec la simulation rapide sont reportées dans le tableau 4.2. Ces résultats ont été obtenus à partir d'un lot de 400000 événements, simulés et reconstruits avec `At1Fast`. Les jets ont été reconstruits avec l'algorithme du cône ($\Delta R = 0.4$ ⁴) et une efficacité d'étiquetage $\epsilon_b = 50\%$ pour les jets issus de b .

Le boson W candidat est ensuite reconstruit en combinant toutes les paires de jets légers (parmi les jets ayant un $p_T > 40 \text{ GeV}/c$ et dans l'acceptance $|\eta| < 2.5$). La masse invariante de toutes les paires jet-jet est représentée sur la figure 4.2. Avec les événements Monte-Carlo, il est possible de caractériser la contribution du bruit de fond combinatoire qui représente la population des paires de jets pour lesquelles au moins un jet de la paire n'est pas originaire d'un des quarks provenant du W Monte-Carlo. Une association quark-jet est construite en essayant d'apparier à chaque quark du W , un jet reconstruit distant d'au plus 0.4 en ΔR . Nous avons veillé à ne pas introduire de double comptage : un jet est associé une et une seule fois. Les paires de jets provenant de ce bruit de fond combinatoire sont représentées par la zone hachurée en bleu sur la figure 4.2. Ce graphe

⁴Le rayon d'un cône est défini comme la norme euclidienne entre deux points : $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$

Menu	Taux attendus à basse luminosité
$e25i, 2e15i$	$\sim 40 Hz$
$\gamma60i, 2\gamma20i$	$\sim 25 Hz$
$\mu20i, 2\mu$	$\sim 40 Hz$
$j400, 3j165, 4j110$	$\sim 40 Hz$
$j70 + xE70$	$\sim 20 Hz$
$\tau35 + xE45$	$\sim 5 Hz$
$2\mu6$ avec coupures en masse	$\sim 10 Hz$
étalonnage, facteurs d'échelle	$\sim 20 Hz$

TAB. 4.1 – Tableau des menus du système de déclenchement de dernier niveau (*High Level Trigger : HLT*). Les menus sont de la forme $NoXXi$ où N est le nombre d'objets requis pour déclencher, o est le type d'objet sur lequel le déclenchement s'effectuera, XX est l'énergie transverse requise en GeV et i indique s'il faut ou non ajouter un critère d'isolation. xE représente l'énergie manquante dans le plan transverse.

Coupure	Efficacité (At1Fast)	Nbre évts pour $\mathcal{L} = 10fb^{-1}$
Sel1	$67.4 \pm 0.2\%$	1.7×10^6
Sel2	$60.4 \pm 0.2\%$	1.5×10^6
Sel3	$18.9 \pm 0.3\%$	0.4×10^6
Sel4	$12.4 \pm 0.3\%$	0.3×10^6
Sel5	$3.5 \pm 0.3\%$	0.09×10^6

TAB. 4.2 – Tableau des efficacités séquentielles des différentes coupures de présélection pour les événements $t\bar{t}$ générés avec **At1Fast**. La dernière colonne donne le nombre d'événements attendus pour un an de prise de données à basse luminosité. Chaque ligne donne donc le nombre d'événements passant la totalité des coupures pour une étape donnée de la présélection.

Processus	Section efficace (pb)	Efficacité totale (%)
$t\bar{t}$	250	3.5
$b\bar{b} \rightarrow l\nu + jets$	2.2×10^6	3×10^{-8}
$W + jets \rightarrow l\nu + jets$	7.8×10^3	2×10^{-4}
$Z + jets \rightarrow l^+l^- + jets$	1.2×10^3	6×10^{-5}
$WW \rightarrow l\nu + jets$	17.1	7×10^{-3}
$WZ \rightarrow l\nu + jets$	3.4	1×10^{-2}
$ZZ \rightarrow l^+l^- + jets$	9.2	3×10^{-3}

TAB. 4.3 – Tableau des sections efficaces de production et des efficacités totales de sélection pour le signal et les différents bruits de fond. Tableau extrait de [75].

illustre le fait que même si parmi toutes ces combinaisons, nous avons effectivement la bonne paire de jets (si chacun des jets a pu être reconstruit et que chacun passe les coupures de présélection), la pureté ⁵ en paires jet-jet provenant effectivement du boson W est assez faible : $p = (7.5 \pm 0.1)\%$. Il est nécessaire de faire un choix parmi toutes ces paires jet-jet pour améliorer cette pureté.

La combinaison retenue est celle la plus proche du pic de la masse invariante du W avant étalonnage. Cette combinaison est reportée sur la figure 4.3 en traits pleins rouges (haut et bas). Les distributions en traits pointillés bleus sont le bruit de fond combinatoire (figure du haut) et la masse invariante de la bonne paire de jets (figure du bas), c'est-à-dire la paire formée des jets associés aux quarks provenant de la désintégration du boson W . Ce critère de sélection permet alors d'enrichir notre lot de W : $p = (45.0 \pm 0.3)\%$. Nous pouvons noter par ailleurs que les ajustements gaussiens des masses invariantes des paires jj et des paires de jets ayant été associées aux quarks provenant du W Monte-Carlo sont compatibles : $\langle m_{jj}^{Rec} \rangle = 79. \text{ GeV}/c^2$ et $\sigma_{m_{jj}} = 7 \text{ GeV}$. Nous pouvons d'ores et déjà noter le décalage du pic en masse du W reconstruit (par l'algorithme de reconstruction et celui d'association). La procédure d'étalonnage va trouver les poids à appliquer aux quadri-vecteurs afin d'ajuster ce pic à la masse générée avec Pythia 6.203 ($m_W^{Mc} = 80.42 \text{ GeV}/c^2$ et $\Gamma_W = 2.124 \text{ GeV}$)

Une autre sélection est alors effectuée afin d'augmenter cette pureté. Les candidats W ne sont en effet retenus que s'ils se trouvent à l'intérieur d'une fenêtre de $20 \text{ GeV}/c^2$ [73,74] autour du pic en masse ($80 \text{ GeV}/c^2$ avec At1Fast-01-00-91). Ceci permet de porter la pureté à $p = (65.7 \pm 0.3)\%$. La figure 4.4 illustre la contamination des différents types de bruits de fonds combinatoires pour les événements possédant au moins un jet de b (Sel4, figure du haut) ou deux jets de b (Sel5, figure du bas). Nous pouvons en effet dégager deux types de bruits de fonds : celui correspondant aux événements pour lesquels la bonne combinaison n'existe pas car au moins un des jets provenant du W ne satisfait pas les coupures de présélection (zone hachurée par des traits inclinés sur la gauche). L'autre type de bruit de fond provient des événements pour lesquels une mauvaise combinaison est choisie car se situant plus près du pic en masse que la bonne combinaison (zone hachurée par des traits inclinés sur la droite). Les différentes pollutions sont reportées dans le tableau 4.4. Pour réduire cette pollution et ainsi augmenter la pureté de l'échantillon en vraies paires de W , il est possible de ne sélectionner que les événements pour lesquels exactement 4 jets sont reconstruits.

Les candidats top hadroniques sont ensuite reconstruits en associant le W précédemment reconstruit avec tous les jets labellés b . Le candidat retenu est celui de plus haute impulsion transverse. Cette sélection est illustrée par la figure 4.5. L'ajustement par une gaussienne donne $\langle m_{jbb} \rangle = 173.0 \text{ GeV}/c^2$ et $\sigma_{jbb} = 15.3 \text{ GeV}/c^2$. Ces résultats sont à comparer aux caractéristiques du top généré avec Pythia 6.203 de masse et de largeur : 175.0 et 1.398, respectivement. Une dernière coupure sur la masse du top reconstruite est appliquée : nous ne gardons pour étalonnage que les événements satisfaisant $|m_{jbb} - m_{jbb}^{Pic}| < 15 \text{ GeV}/c^2$. Cette coupure permet d'avoir un lot en W pur à $(71.2 \pm 0.3)\%$ sur l'ensemble du spectre en masse du W et une pureté de $(80.5 \pm 0.3)\%$ dans la fenêtre

⁵La pureté est définie comme étant le nombre de paires jet-jet pour lesquelles chacun des jets a pu être associé à un quark issu du W avant tout processus de radiation dans l'état final. Ce nombre est alors rapporté au nombre total de W hadroniques.

	Sel4	Sel5
Pollution type I	23.5 ± 0.3	18.3 ± 0.5
Pollution type II	19.2 ± 0.2	13.0 ± 0.4
Pureté en W_{jj}	65.7 ± 0.3	74.8 ± 0.6

TAB. 4.4 – Pollutions des bruits de fonds de types I et II (respectivement : jet hors sélection et meilleure combinatoire) pour les événements satisfaisant la coupure Sel4 ou la coupure Sel5. La pureté en W pour chacune de ces classes d'événements est également reportée.

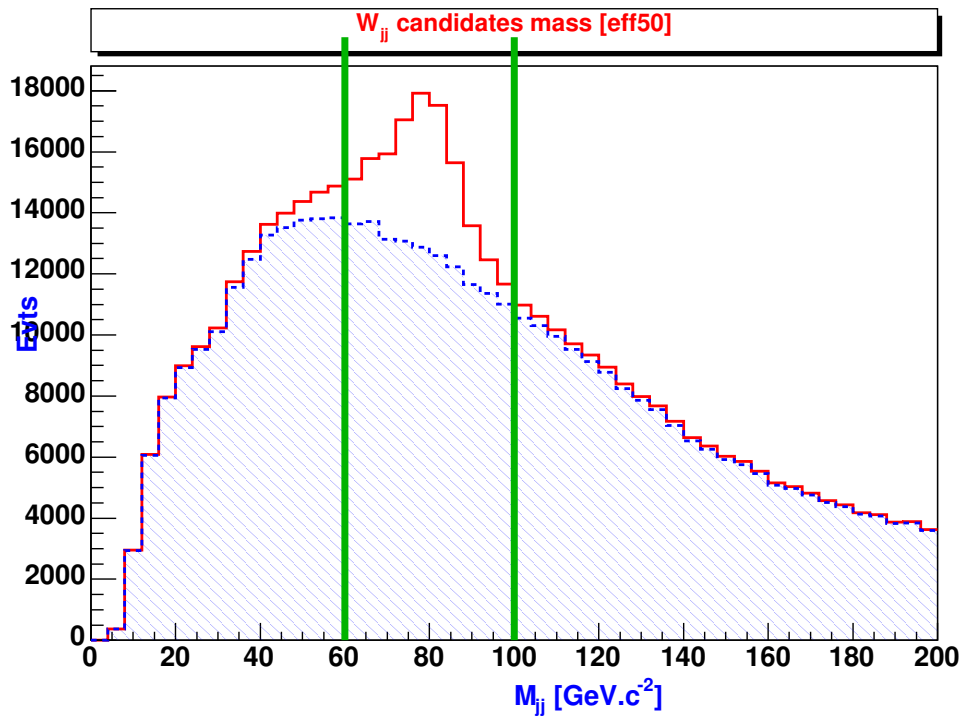


FIG. 4.2 – Masse invariante de tous les candidats W en traits pleins rouges (haut et bas). Le bruit de fond combinatoire est représenté en pointillés bleus coiffant la zone hachurée (figure du haut). Dans la figure du bas, les traits pointillés bleus représentent la masse invariante des jets associés aux quarks originaires du W .

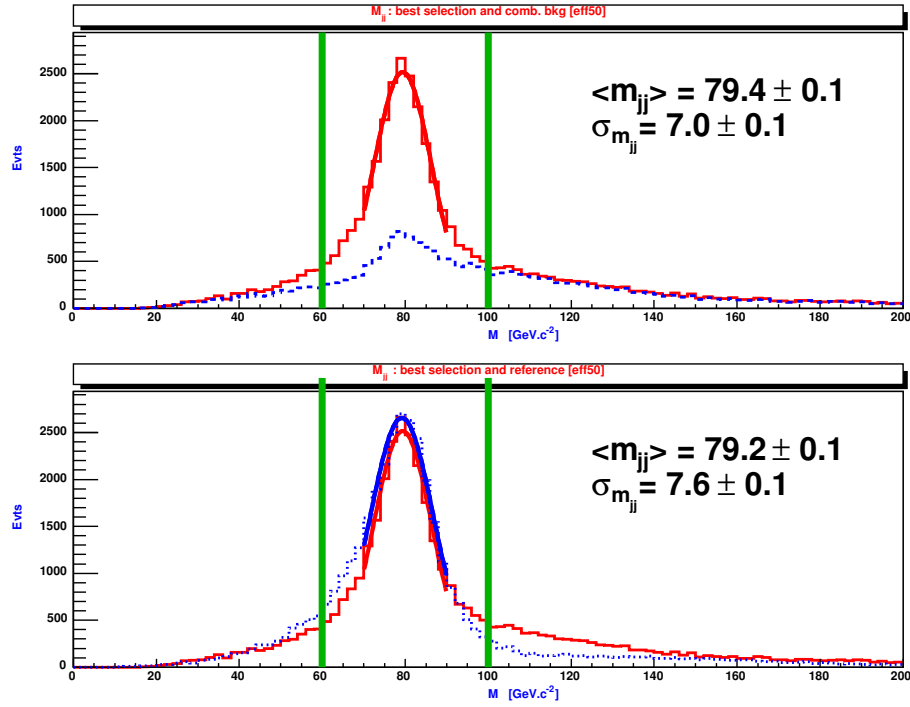


FIG. 4.3 – Masse invariante des paires de jets réalisant la meilleure combinaison par rapport au pic en masse. Les traits pleins rouges représentent la meilleure combinaison par rapport au pic en masse (haut et bas). Le résultat de son ajustement par une gaussienne est donné sur la figure du haut. Le bruit de fond combinatoire est représenté en pointillés bleus sur la figure du haut. Les pointillés bleus de la figure du bas représentent la masse invariante reconstruite à partir des jets qui ont été associés aux quarks provenant du W Monte-Carlo. Le résultat de son ajustement par une gaussienne est donné sur la figure du bas.

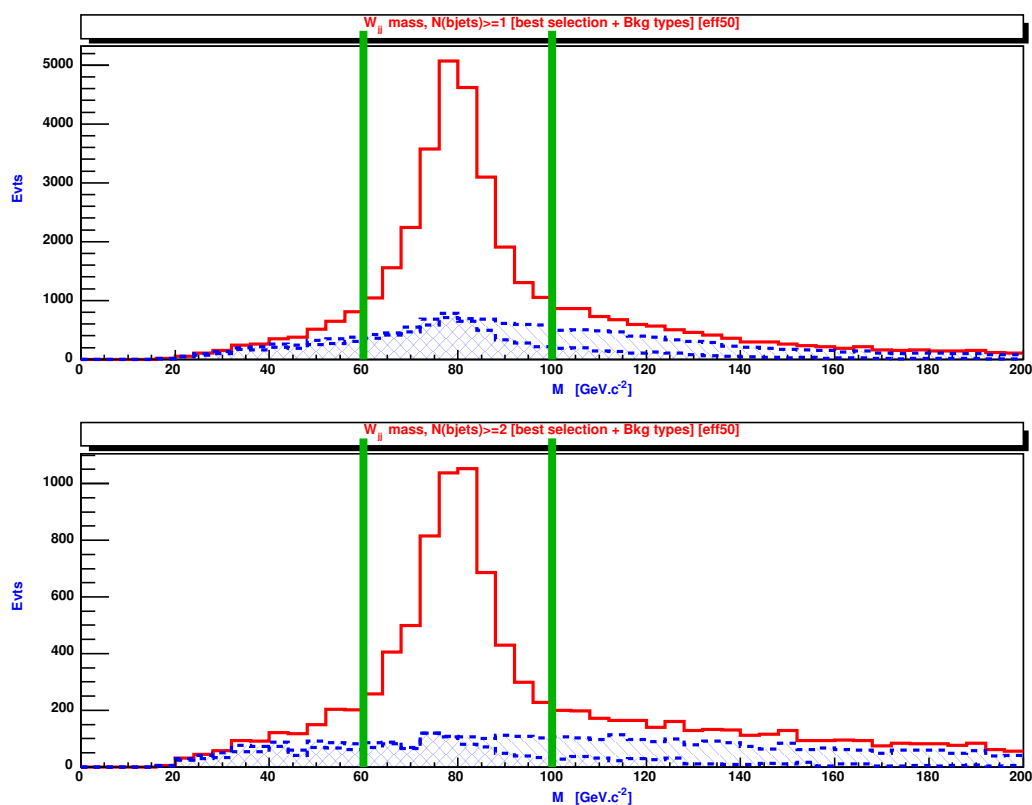


FIG. 4.4 – Masse invariante des paires de jets réalisant la meilleure combinaison par rapport au pic en masse, pour la coupe avec au moins un jet de b dans l'événement (haut) et avec deux jets de b (bas). Les différents types de bruits de fonds combinatoires sont représentés par les zones hachurées en bleu. Les hachures inclinées sur la droite correspondent aux événements pour lesquels une meilleure combinaison jet-jet est trouvée, supplantant celle qu'il aurait fallu choisir. Les hachures inclinées sur la gauche correspondent aux événements pour lesquels au moins un des jets de la bonne combinaison ne passe pas les coupures de présélection.

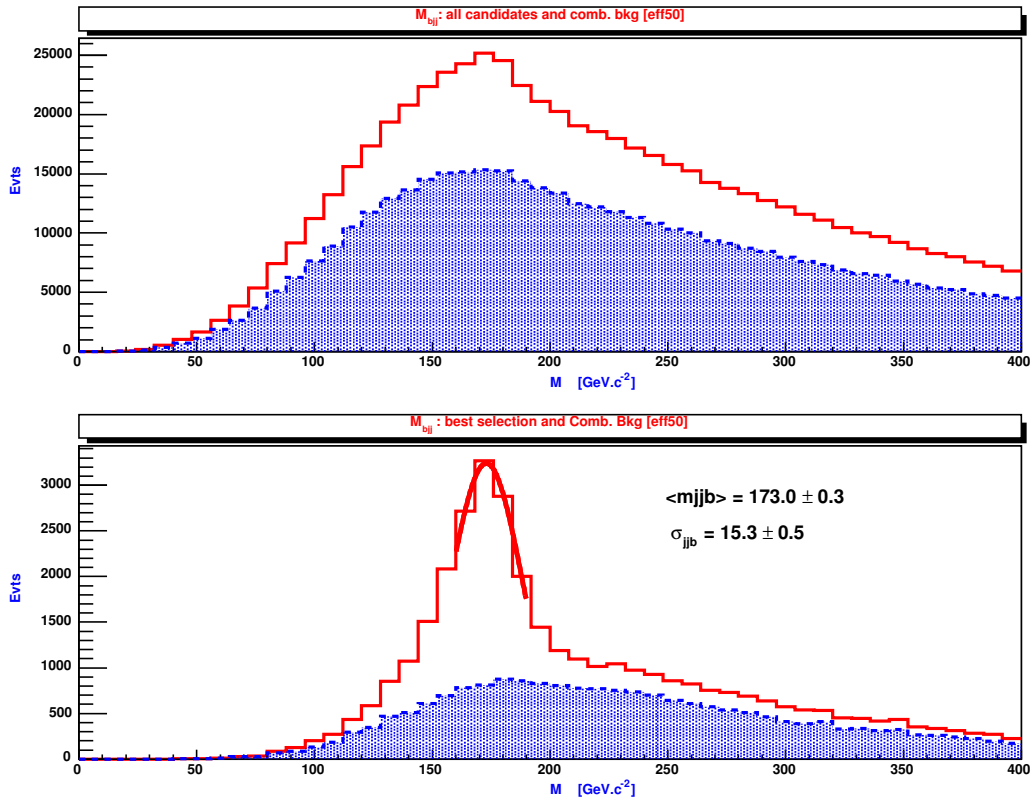


FIG. 4.5 – Masses invariantes (en traits pleins rouges) de toutes les combinaisons jtb (figure du haut) et masse invariante de la combinaison maximisant l’impulsion transverse du top ainsi reconstruit (figure du bas). Le bruit de fond combinatoire est représenté par la zone hachuré en pointillés bleus.

en masse de 20 GeV autour du pic.

A cette étape de l'analyse, nous avons à disposition un lot suffisamment pur de W sur lequel nous pouvons ensuite travailler pour extraire les coefficients d'étalonnage ($N_W \sim 22000$ pour $\mathcal{L} = 10 fb^{-1}$). La méthode générale utilisée pour obtenir ces coefficients est donnée dans le paragraphe suivant.

4.2.2 Méthode d'étalonnage

A ce stade de l'analyse, nous avons à disposition les quadri-vecteurs des deux jets constituant le W reconstruit. A des fins de vérification, nous avons également gardé les partons les ayant initiés. En utilisant les quadri-vecteurs reconstruits, nous pouvons alors calculer les poids à appliquer sur ces derniers afin de satisfaire la contrainte en masse du W . L'énergie E , la pseudo-rapidité η et l'angle azimutal ϕ sont ajustés événement par événement. En effet, la masse invariante reconstruite du candidat W est :

$$m_{jj} = \sqrt{2E_{j_1}E_{j_2}(1 - \cos\theta_{j_1j_2})} \quad (4.3)$$

sous l'approximation d'une masse nulle pour les jets légers. Ainsi, une mauvaise reconstruction de l'angle des jets influe sur la qualité de la reconstruction de la masse invariante. De plus, comme décrit dans [73], l'angle qui est reconstruit est systématiquement plus petit que l'angle entre les partons⁶. Ceci introduit donc un biais dans la procédure d'extraction des poids et empêche d'obtenir l'étalonnage avec la précision souhaitée de 1%, si nous nous contentons d'un ajustement seulement en énergie. De plus, cet effet augmente avec le recouvrement des jets. Pour corriger ce biais, les coefficients d'étalonnages obtenus événement par événement sont donc le résultat d'un ajustement en énergie et en direction, avec m_W comme contrainte : α^E, α^η et α^ϕ . Aucune hypothèse sur la forme ou bien la dépendance en énergie n'est requise afin d'obtenir ces facteurs correctifs.

Ces coefficients d'étalonnage sur les jets de W peuvent être utilisés tels quels dans le cadre de la mesure de la masse du *top*. Cependant, nous allons extraire de cette collection de coefficients une fonction dépendant de l'énergie qui pourra être utilisée dans d'autres canaux de physique.

Nous regroupons alors les coefficients par tranches en énergie, chaque intervalle contenant la même statistique. Chacune des distributions ainsi construites est alors ajustée par une distribution gaussienne sur l'intervalle $[-2\sigma; +\sigma]$ centré sur le pic. Ceci permet de s'affranchir des effets des queues de distribution à basse énergie, dûs aux radiations dans l'état final. Nous pouvons ensuite ajuster les valeurs moyennes de ces gaussiennes par la fonction suivante :

$$F(E) = a + \frac{b}{E} \quad (4.4)$$

qui donnera le facteur correctif voulu pour une énergie de jet avant étalonnage donnée. D'autres expressions analytiques sont envisageables pour décrire l'évolution des facteurs de corrections en énergie. Il n'y a en effet pas d'hypothèse *a priori* sur la forme de cette fonction d'ajustement. Nous utilisons la fonction qui reproduit le mieux l'évolution observée des facteurs de correction. En effet, lorsque l'énergie incidente augmente, le

⁶Avec les AOD produits pour l'atelier de physique de Rome, l'algorithme de reconstruction des jets attribue une masse aux jets ($E \neq p$). Il est alors possible de déterminer un $\cos\theta$ reproduisant le $\cos\theta^{MC}$.

terme en b/E diminue et le poids à appliquer tend vers a . Ceci correspond d'une part au fait que la précision sur la mesure du dépôt d'énergie dans un calorimètre s'améliore avec l'énergie incidente (effet de la résolution intrinsèque du calorimètre) et d'autre part, à haute énergie, les particules constituant un jet étant plus collimées, il y a moins de pertes d'énergies dûes à l'algorithme de cône. Au contraire, à basse énergie, la résolution en énergie se dégrade et il faut alors appliquer une correction plus grande pour remonter à l'énergie du parton incident, et l'effet des pertes d'énergies en dehors du cône est plus important. L'énergie corrigée sera alors :

$$E_{Corr} = F(E_{Raw}) \times E_{Raw} \quad (4.5)$$

Cette méthode permet donc, à partir de notre paire de quadri-vecteurs initiale, de mettre en relation pour chaque jet :

- E_{Raw} , l'énergie du jet avant tout traitement,
- E_{Fit} , l'énergie obtenue événement par événement grâce à l'ajustement *via* la contrainte en masse invariante, et
- E_{Corr} , l'énergie du jet après application de la fonction d'étalonnage obtenue après paramétrisation.

Par définition, les facteurs de corrections associés à chacune des énergies E_{Fit} et E_{Corr} s'expriment de la manière suivante :

$$\alpha_{Fit}^E = \frac{E_{Fit}}{E_{Raw}}; \quad \alpha_{Corr}^E = \frac{E_{Corr}}{E_{Raw}} \quad (4.6)$$

Nous pouvons alors comparer chacun de ces coefficients aux poids Monte-Carlo construits à partir du rapport :

$$\alpha_{Mc} = \frac{E_{Mc}}{E_{Raw}} \quad (4.7)$$

Ce sont les poids qu'il faudrait appliquer pour remonter à l'énergie du parton initial. Les poids α_{Mc} peuvent être regroupés dans les mêmes tranches en énergies que pour les jets puis ajustés par la fonction de l'équation 4.4. Cette paramétrisation sera alors considérée comme la fonction Monte-Carlo d'ajustement.

4.2.3 Extraction des coefficients d'étalonnage

L'objet de ce paragraphe est de présenter les deux méthodes d'extraction des coefficients d'étalonnage qui ont été utilisées. La première méthode est tirée de la référence [73] et la deuxième utilise les multiplicateurs de Lagrange afin de mieux tirer parti de la contrainte de la masse invariante des deux jets. Ces deux méthodes ont été testées sur un lot de $\sim 400k$ événements reconstruits avec `AtlFast-01-00-91`, dans le cadre d'`ATHENA-10.0.1`.

4.2.3.1 Méthode du χ^2

Description

Dans le cadre de cette méthode, les coefficients d'étalonnage pour les jets légers sont trouvés par l'intermédiaire d'un ajustement d'un χ^2 . Cette procédure doit permettre de

retrouver la masse m_W^{PDG} en modifiant le quadri-vecteur des jets, de manière indépendante. Le χ^2 est défini comme suit :

$$\chi^2 = \left(\frac{m_{jj} - M_W^{PDG}}{\Gamma_W} \right)^2 + \sum_{i=1,2}^{jets} \left(\frac{E_i^m - \alpha_i^E E_i^m}{\sigma_{E_i}} \right)^2 + \sum_{i=1,2}^{jets} \left(\frac{\eta_i^m - \alpha_i^\eta \eta_i^m}{\sigma_{\eta_i}} \right)^2 + \sum_{i=1,2}^{jets} \left(\frac{\phi_i^m - \alpha_i^\phi \phi_i^m}{\sigma_{\phi_i}} \right)^2 \quad (4.8)$$

où Γ_W est la largeur naturelle du boson W ($\Gamma_W^{PDG} = 2.118 \pm 0.042$ GeV [47]), les $\alpha_i^{E,\eta,\phi}$ sont les coefficients d'étalonnage en énergie, η et ϕ , et les X_i^m (avec $X = E, \eta, \phi$) sont les quantités mesurées pour chaque jet.

Les résolutions en énergie et en direction sont obtenues grâce aux associations quark-jet précédemment construites. Pour différents intervalles en énergie de jet, les distributions $(E_q - E_j)/E_j$ sont construites puis ajustées par des gaussiennes. La variation de l'écart type de ces gaussiennes en fonction de l'énergie des jets est ensuite ajustée par la fonction :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \quad (4.9)$$

pour la résolution en énergie, et par la fonction suivante pour les résolutions en direction :

$$\sigma_{X=\eta,\phi} = a + \frac{b}{E^2} \quad (4.10)$$

Les fonctions obtenues ainsi pour les résolutions en énergie et en direction et pour des jets reconstruits avec l'algorithme de Cône ($\Delta R = 0.4$) sont données dans la figure 4.6. Nous obtenons pour les résolutions :

$$\begin{aligned} \frac{\sigma_E}{E} &= \frac{(83 \pm 2)\%}{\sqrt{E}} \oplus (5.5 \pm 0.2)\% \\ \sigma_\eta &= (0.025 \pm 0.001) + \frac{58 \pm 2}{E^2} \\ \sigma_\phi &= (0.025 \pm 0.001) + \frac{85 \pm 3}{E^2} \end{aligned} \quad (4.11)$$

Le terme stochastique (c'est-à-dire a dans l'équation 4.9) gouverne le flot d'énergie (*energy flow*). Nous pouvons remarquer que cette quantité calculée avec **At1Fast** est moins bonne que la résolution intrinsèque du calorimètre dans son ensemble [29]. Ce résultat prend en compte l'effet des événements sous-jacents (*Underlying event*), des radiations de gluons et des algorithmes de reconstruction des jets. De plus, nous pouvons noter l'effet de la géométrie pseudo-projective en η qui dégrade la résolution, par rapport à celle que nous obtenons pour l'angle azimutal.

Résultats

Les coefficients d'étalonnage (obtenus événement par événement) en fonction de l'énergie sont reportés sur la figure 4.7. Le W ajusté étant composé de 2 jets, il y a donc 2

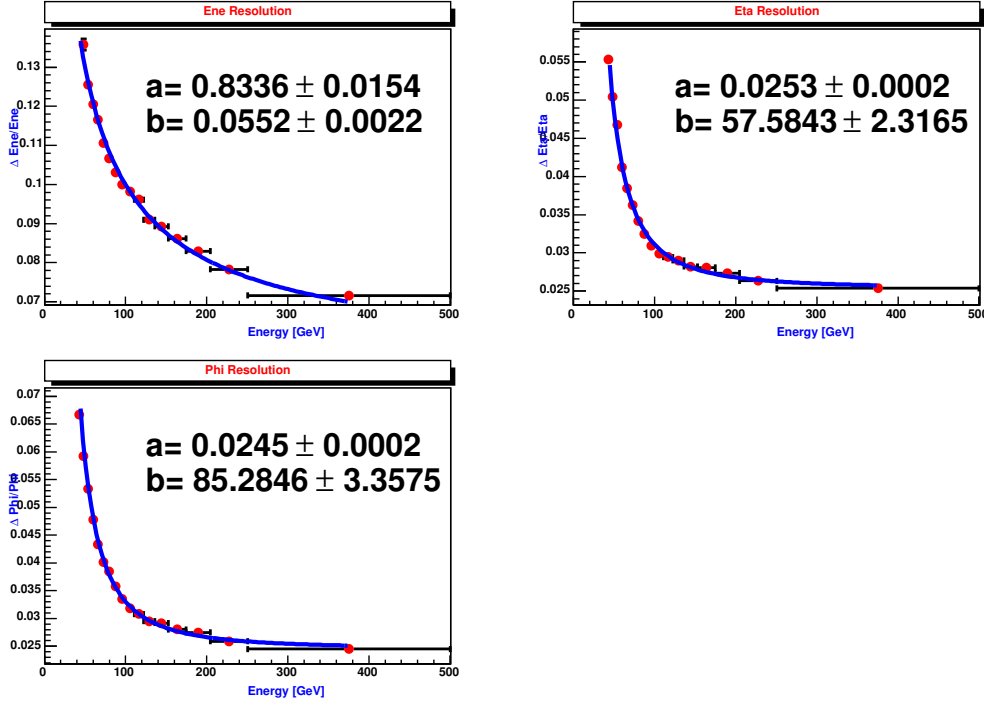


FIG. 4.6 – Résolutions en E , η et ϕ obtenues pour les événements $t\bar{t}$ simulés et reconstruits avec At1Fast-01-00-91.

poids par événement qui sont utilisés pour construire cette distribution. Comme expliqué précédemment, le plan des poids en fonction de l'énergie des jets est découpé en intervalles en énergie, chaque intervalle ayant la même statistique. La distribution obtenue pour l'intervalle $75 \text{ GeV} \leq E_{Raw} < 80 \text{ GeV}$ est donnée en exemple sur la figure 4.8.

L'ajustement du χ^2 a été réalisé en utilisant la réimplémentation en C++ du programme MINUIT [79]. Après avoir regroupé les énergies des jets en intervalles de même statistique, les rapports E_{Parton}/E_{jet} sont calculés et reportés sur la figure 4.9, pour différentes étapes de la procédure d'étalonnage. Les ronds verts vides sont les rapports entre l'énergie du quark et l'énergie du jet telle que donnée par At1Fast. Contrairement aux versions précédentes, l'énergie des jets est en moyenne surestimée. Ceci se retrouve dans la fonction de paramétrisation des poids Monte-Carlo :

$$F_{MC}(E) = (0.984 \pm 0.001) + \frac{0.1 \pm 0.2}{E} \quad (4.12)$$

Le terme constant est en effet inférieur à 1. Nous pouvons également remarquer que cette nouvelle calibration a nettement amélioré la reconstruction de l'énergie à basse énergie [73]. Les ronds rouges pleins représentent les rapports en énergie où l'énergie du jet est corrigée par la constante d'étalonnage calculée événement par événement à l'aide de l'ajustement contraint. Nous remarquons qu'à ce stade, les rapports en énergie obtenus avec l'ajustement en énergie et en direction, sont à l'intérieur des 1% de précision. Il en va de même pour les rapports corrigés avec la paramétrisation des poids, représentés par des triangles bleus pleins. La fonction de paramétrisation ajustée sur les poids calculés

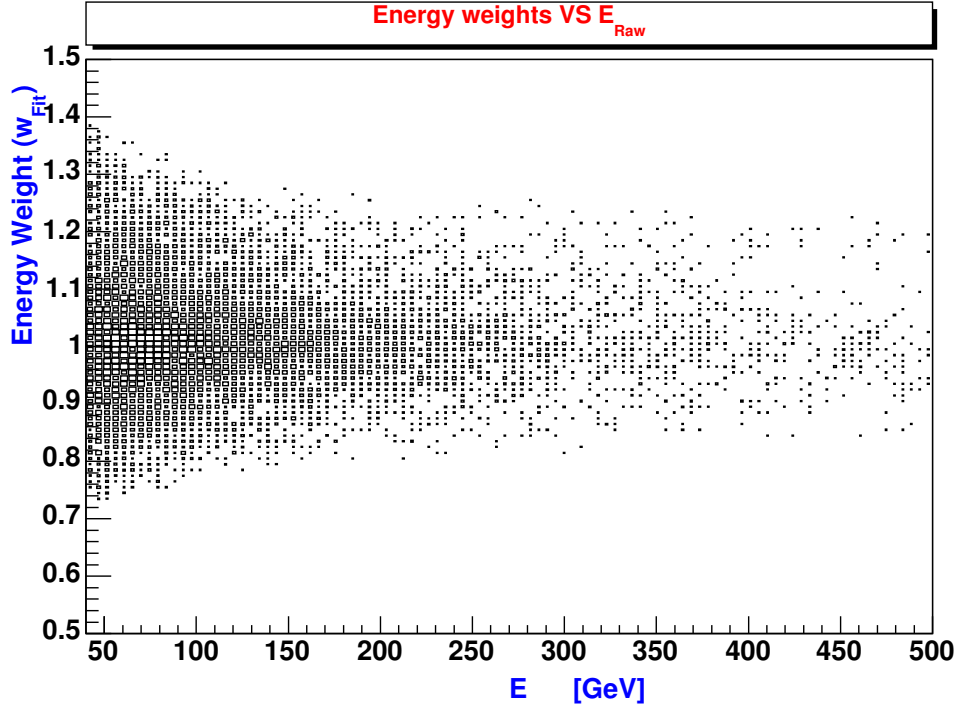


FIG. 4.7 – Distribution des coefficients de correction en énergie obtenus événement par événement *via* un ajustement par MINUIT. Il y a 2 entrées par événement puisque l’ajustement d’un W porte sur les 2 jets le constituant.

événement par événement est :

$$F(E) = (0.985 \pm 0.001) + \frac{0.3 \pm 0.1}{E} \quad (4.13)$$

Le terme constant est en accord avec celui trouvé pour l’ajustement Monte-Carlo. Le terme en énergie introduira par contre un biais pour l’étalonnage des jets de basse énergie. Cependant, d’après la figure 4.9 et la distribution des triangles bleus, ce biais reste sous contrôle au niveau du 1%.

La masse invariante de la paire de jets formant le W reconstruit est représentée sur la figure 4.10. L’application des facteurs de correction sur les quadri-vecteurs des jets a bien décalé le pic en masse sur la valeur requise lors de l’ajustement. Ceci prouve que l’ajustement du χ^2 fonctionne de la manière attendue.

4.2.3.2 Multiplicateurs de Lagrange

Description

Cette technique est utilisée pour ajuster un échantillon de données par un ensemble de paramètres dont la valeur doit satisfaire une ou plusieurs équations de contrainte. C’est une manière de contraindre l’ajustement à se comporter de manière *physique* et donc de donner un résultat qui se conforme aux lois de la physique sous-jacente. Étant donnée l’équation de contrainte suivante :

$$f(E_{1,2}, \eta_{1,2}, \phi_{1,2}) = m_{12}^2 - m_{PDG}^2 \doteq 0 \quad (4.14)$$

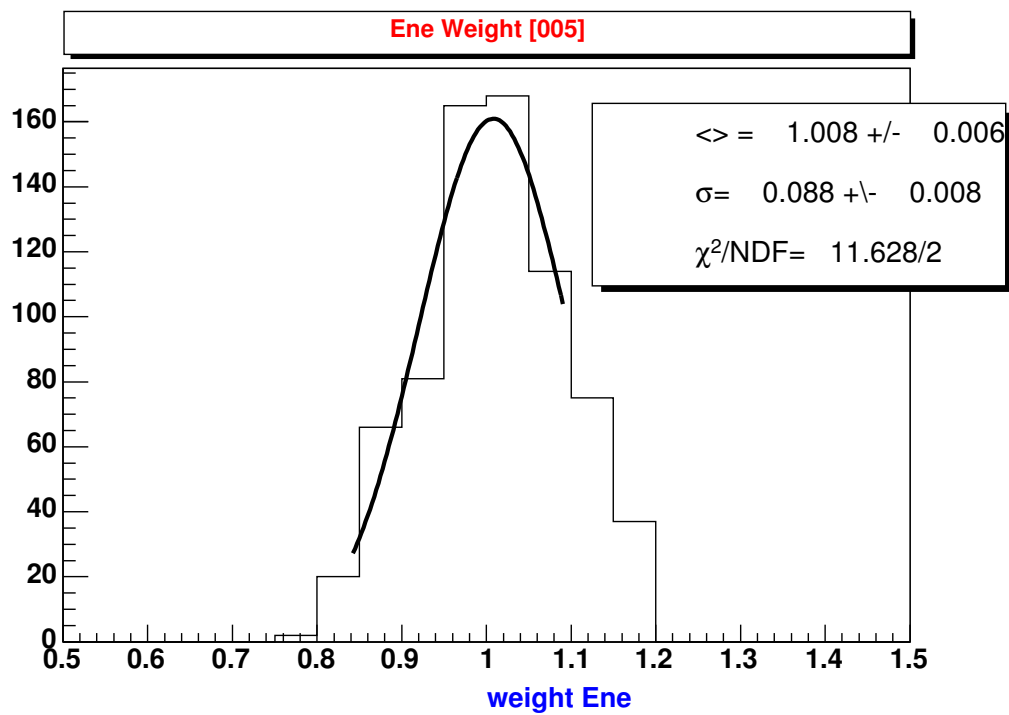


FIG. 4.8 – Distribution des coefficients de correction pour l'intervalle en énergie [75 GeV; 80 GeV]. Ces coefficients de corrections sont ajustés par une gaussienne sur l'intervalle asymétrique [$pic - 2 \times \sigma$; $pic + \sigma$] afin de limiter les effets dûs aux radiations de gluons. La moyenne de la gaussienne sera ensuite utilisée pour paramétriser l'évolution des constantes de correction.

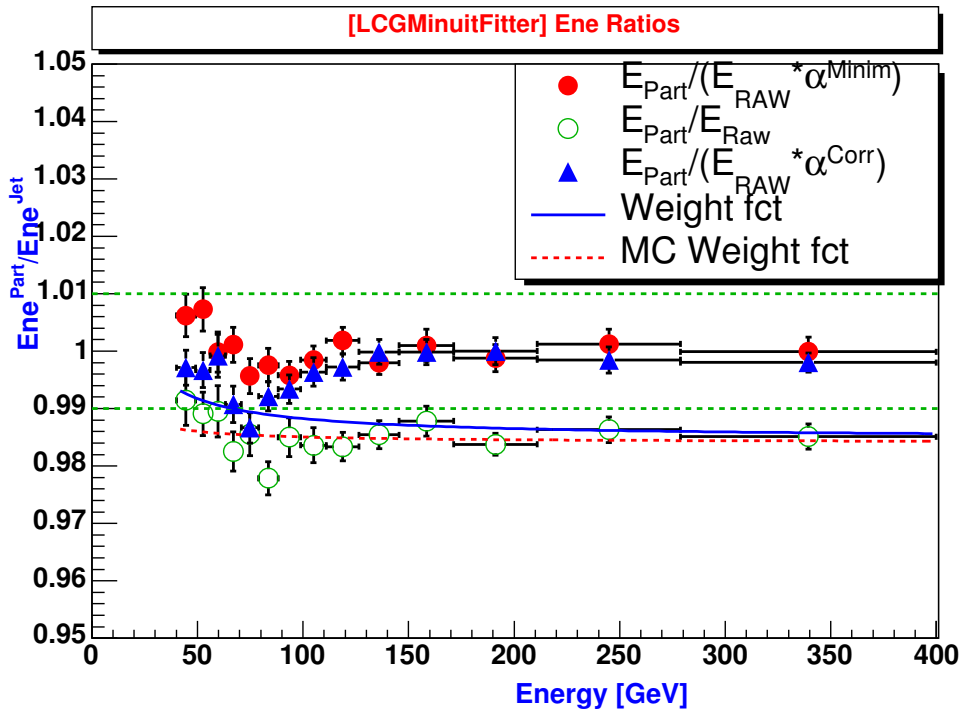


FIG. 4.9 – Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec l'ajustement par MINUIT. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4.

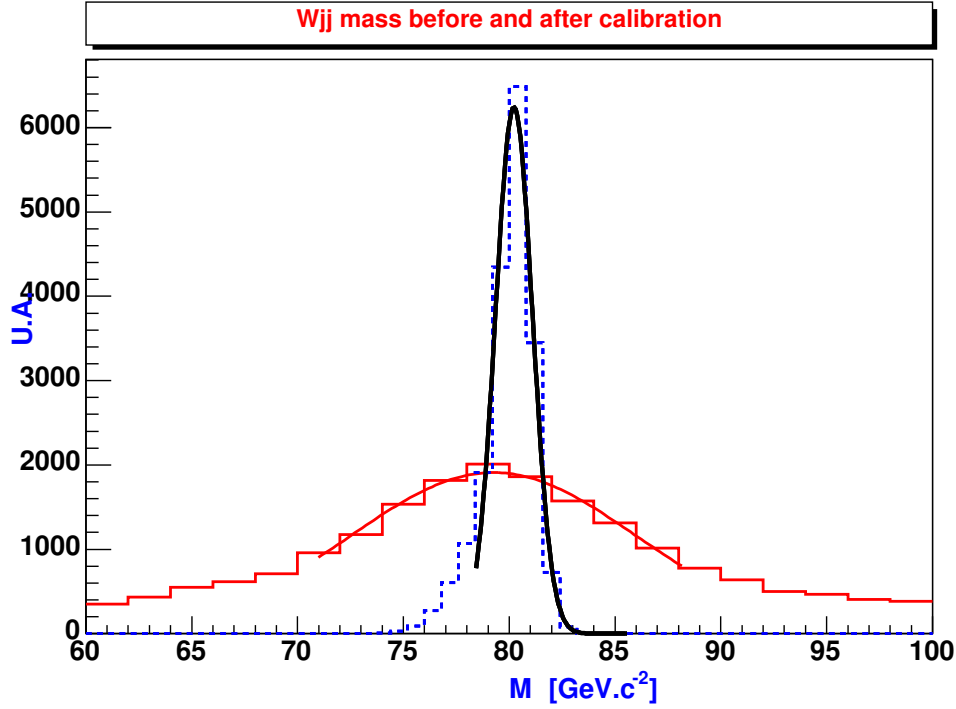


FIG. 4.10 – Distribution de la masse invariante m_{jj} avant (en traits pleins rouges) et après application des facteurs de correction événement par événement (en traits pointillés bleus).

la méthode des multiplicateurs de Lagrange [80] consiste en l'introduction de variables additionnelles dans la fonction à ajuster qui devient alors :

$$\chi^2 = (\vec{x} - \vec{x}_0)^T V_x^{-1} (\vec{x} - \vec{x}_0) + 2\vec{\lambda}^T \vec{f}(E_{1,2}; \eta_{1,2}; \phi_{1,2}) \quad (4.15)$$

où $\vec{\lambda}$ est le vecteur contenant les multiplicateurs de Lagrange, V_x est la matrice de covariance et \vec{x} le vecteur des paramètres à ajuster. Il est important de noter que la principale hypothèse sur laquelle repose l'ajustement par les multiplicateurs de Lagrange est la distribution normale des variables contenues dans le vecteur \vec{x} . Nous pouvons obtenir directement les valeurs de \vec{x} qui minimisent le χ^2 si les équations de contrainte sont linéaires. Dans notre cas, il faut linéariser l'équation de contrainte en réalisant un développement en série de Taylor autour d'un point \vec{x}_A :

$$f(\vec{x}) \sim f(\vec{x}_A) + \left(\frac{\partial f}{\partial \vec{x}} \right)_{\vec{x}=\vec{x}_A} (\vec{x} - \vec{x}_A) \doteq 0 \quad (4.16)$$

Ce qui donne en utilisant la notation condensée matricielle :

$$B_{kl} = \left(\frac{\partial f_k}{\partial x_l} \right); \quad cstr_k = f_k(\vec{x}_0); \quad \vec{\delta} = \vec{x} - \vec{x}_0; \quad (4.17)$$

$$\Rightarrow B\vec{\delta} + cstr = \vec{0} \quad (4.18)$$

L'ajustement du χ^2 revient alors à résoudre le système :

$$\begin{cases} \frac{\partial \chi^2}{\partial x_l} = 0 \\ \frac{\partial \chi^2}{\partial \lambda_k} = 0 \end{cases} \quad (4.19)$$

Il convient de noter à ce stade que la contrainte introduite ne prend pas encore en compte la largeur de la résonance que nous essayons d'ajuster. Afin d'avoir un comportement correct de l'ajustement, la fonction de contrainte est modifiée de la manière suivante :

$$f(\vec{x}) = m_{12}^2 - (\alpha \cdot m_{PDG}^2)^2 \doteq 0 \quad (4.20)$$

où α est une variable aléatoire qui suit une loi de distribution gaussienne ⁷.

Dans le cas qui nous intéresse, la fonction de contrainte peut se développer de la manière suivante :

$$f(x) = m_1^2 + m_2^2 + 2E_1E_2 - 2\beta\sqrt{E_1^2 - m_1^2}\sqrt{E_2^2 - m_2^2} - m_{PDG}^2 \quad (4.21)$$

où :

$$\beta = \sin \theta_1 \sin \theta_2 \cos(\phi_1 - \phi_2) + \cos \theta_1 \cos \theta_2 \quad (4.22)$$

Le système à résoudre devient alors :

$$\begin{cases} \frac{\partial \chi^2}{\partial E_i} = V_{x_0}^{-1}(x - x_0) + \frac{\partial f}{\partial E_i} \cdot \lambda = 0 \\ \frac{\partial \chi^2}{\partial \eta_i} = V_{x_0}^{-1}(x - x_0) + \frac{\partial f}{\partial \eta_i} \cdot \lambda = 0 \\ \frac{\partial \chi^2}{\partial \phi_i} = V_{x_0}^{-1}(x - x_0) + \frac{\partial f}{\partial \phi_i} \cdot \lambda = 0 \\ \frac{\partial \chi^2}{\partial \lambda} = \left(\frac{\partial f}{\partial E_1}, \frac{\partial f}{\partial \eta_1}, \dots, \frac{\partial f}{\partial \phi_2} \right) + \frac{\partial f}{\partial E_i} \cdot \lambda = 0 \end{cases} \quad (4.23)$$

Résultats

De même que précédemment, les rapports en énergie sont regroupés par intervalles en énergie de même statistique. Le résultat de la procédure d'ajustement pour les multiplicateurs de Lagrange est reporté sur la figure 4.11. La fonction de paramétrisation obtenue pour les poids est la suivante :

$$F(E) = (0.986 \pm 0.001) + \frac{0.2 \pm 0.1}{E} \quad (4.24)$$

Les corrections des énergies des jets par les constantes d'étalonnage obtenues événement par événement et celles *via* la paramétrisation par l'équation 4.4, sont dans la bande des 1%. Nous pouvons également remarquer que les résultats obtenus dans les deux méthodes sont compatibles. La fonction de paramétrisation des poids est d'ailleurs sensiblement la même, seule la vitesse de convergence diffère : l'ajustement par les multiplicateurs de Lagrange ne nécessite qu'une dizaine d'itérations, alors que l'ajustement par χ^2 en requiert un ordre de grandeur plus important.

⁷L'écart-type de cette loi normale sera identifié à la largeur du W .

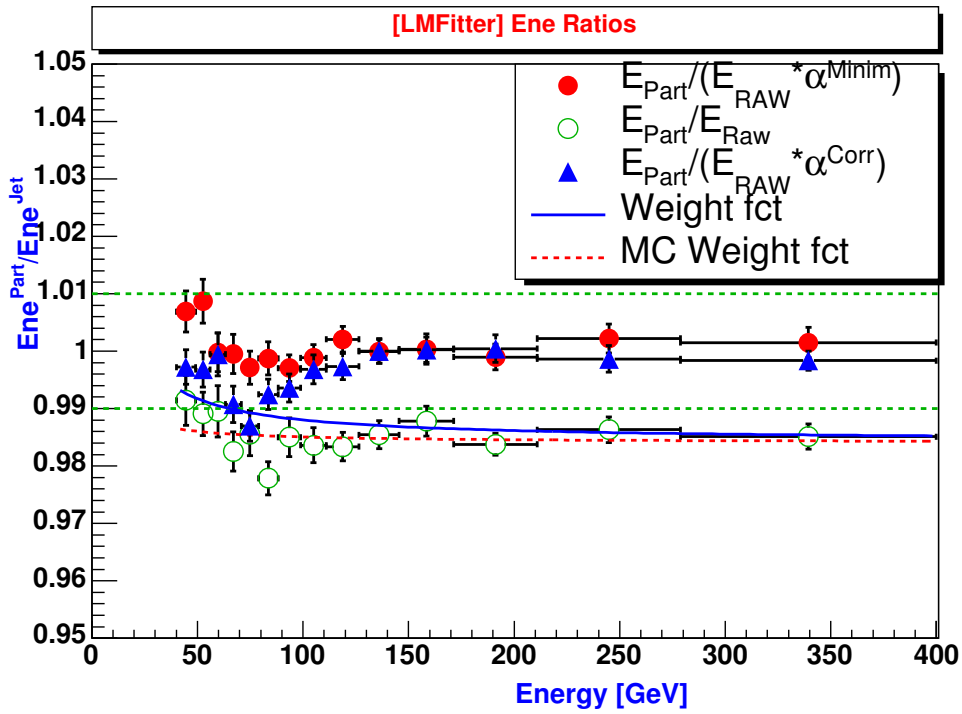


FIG. 4.11 – Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec les multiplicateurs de Lagrange. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4.

4.2.4 Comparaison des résultats

Nous avons vu que les deux méthodes donnent des résultats similaires. En effet, comme le montre la figure 4.12, les masses de W obtenue après ajustement sont sensiblement les mêmes. Cet effet est également visible sur la distribution (figure 4.13) de la différence entre les facteurs de corrections en énergie obtenus événement par événement pour chacune des méthodes. Nous avons également étudié l'impact du bruit de fond combinatoire sur la détermination des facteurs de correction. Nous avons appliqué l'algorithme d'étalonnage sur les événements pour lesquels la paire de jets a pu être associée aux quarks provenant du W Monte-Carlo. Les fonctions de paramétrisation obtenues pour l'évolution des poids sont alors :

$$F(E) = (0.984 \pm 0.001) + \frac{0.4 \pm 0.1}{E} \quad (4.25)$$

$$F(E) = (0.984 \pm 0.001) + \frac{0.4 \pm 0.1}{E} \quad (4.26)$$

respectivement pour la méthode d'ajustement avec MINUIT (4.25) et avec les multiplicateurs de Lagrange (4.26). Nous allons voir le comportement de ces méthodes d'étalonnage pour différents bruits de fond combinatoires. Cette étude est réalisée en jouant sur l'efficacité d'étiquetage des jets issus de b .

Les deux méthodes d'étalonnage sont donc peu sensibles au bruit de fond combinatoire.

4.2.5 Validation avec la reconstruction complète

Afin de valider l'algorithme d'étalonnage écrit dans le *framework* d'ATHENA, nous avons appliqué la méthode d'étalonnage sur des événements $t\bar{t}$ simulés à l'aide de Geant4 et reconstruits avec grâce à la reconstruction complète (*c.f.* 3.4.3). Nous avons utilisé l'échantillon de paires $t\bar{t}$ numéro 4100 générés et simulés pour l'atelier de physique de Rome. Cet échantillon a été généré avec Mc@NLO puis hadronisé avec HERWIG.

Il a tout d'abord été nécessaire de réaliser un prétraitement des collections de quadri-vecteurs contenues dans les fichiers AOD issus de la reconstruction complète. En effet, comme introduit précédemment (*c.f.* 3.4.4) nettoyer les collections pour corriger de la superposition des objets évoquée au chapitre précédent. Nous définissons une collection de référence construite à partir des électrons ayant été reconstruits *via* l'algorithme `egammaBuilder` et présentant les caractéristiques suivantes :

- une impulsion transverse supérieure à $20 \text{ GeV}/c$,
- une trace chargée reconstruite dans le détecteur interne,
- $e/(e + \pi) > 0.95$ où e et π sont des variables discriminantes construites à partir de l'énergie déposée dans les différentes couches du calorimètre, la forme des gerbes (η et ϕ) et le rapport E/p , et
- une énergie transverse à l'intérieur du cône d'isolation $\Delta R = 0.45$ autour de l'électron inférieure à 15 GeV .

La même sélection en impulsion transverse est appliquée à la collection de muons en demandant par ailleurs que le muon ait été reconstruit par l'algorithme de haut p_T . La collection de photons est ensuite allégée de tous les photons qui sont à l'intérieur d'un cône $\Delta R = 0.01$ par rapport à chacun des électrons précédemment sélectionnés ou alors qui n'ont pas une impulsion transverse suffisante ($p_T < 15 \text{ GeV}/c$). Nous demandons ensuite

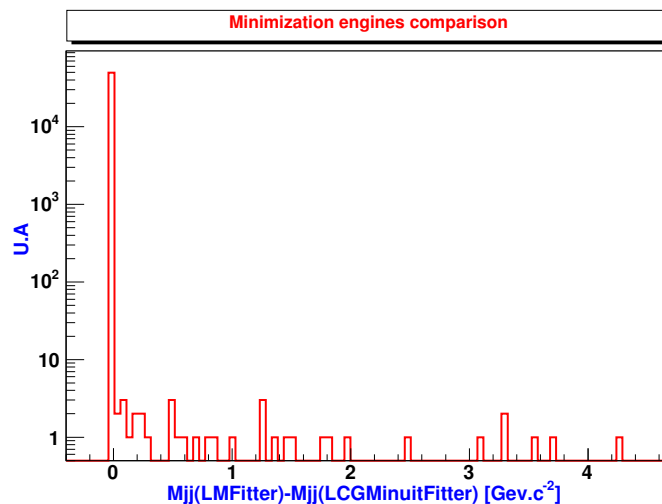


FIG. 4.12 – Différence entre la masse du W après application des poids obtenus événement par événement pour l'ajustement *via* MINUIT et pour l'ajustement en utilisant les multiplicateurs de Lagrange.

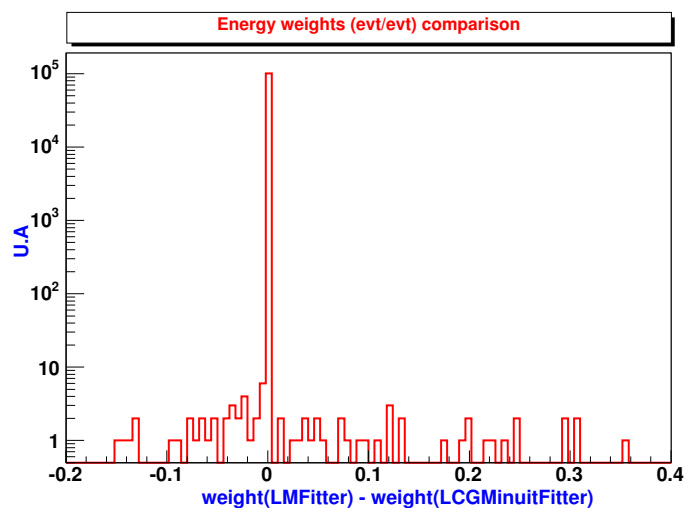


FIG. 4.13 – Différence entre les facteurs de correction en énergie obtenus événement par événement pour l'ajustement *via* MINUIT et pour l'ajustement en utilisant les multiplicateurs de Lagrange.

que les jets identifiés b aient une impulsion transverse minimale de $40 \text{ GeV}/c$ et une valeur supérieure à 3 pour la variable discriminante caractérisant la probabilité d'étiquetage. Nous appliquons enfin la même sélection en impulsion transverse à la collection de jets non étiquetés en demandant qu'il n'y ait aucune superposition avec aucune des collections d'électrons, de muons, de photons et de jets étiquetés b . Ici encore, le critère d'isolation est un cône d'ouverture $\Delta R = 0.01$.

Les histogrammes de contrôle synthétisant le résultat du nettoyage effectué par ce jeu de coupures sont reportés dans la figure 4.14. Nous voyons que lorsque nous comparons notre sélection aux objets reconstruits associés aux objets Monte-Carlo, le recouvrement est très largement satisfaisant.

Une fois le nettoyage des collections AOD effectué, nous pouvons appliquer les procédures de reconstruction des candidats W et top et d'étalonnage des jets légers qui ont été détaillées précédemment. Les rapports d'énergies avant et après étalonnage sont reportés sur la figure 4.15

La paramétrisation de la fonction d'étalonnage obtenue est alors :

$$F(E) = (1.002 \pm 0.001) + \frac{0.6 \pm 0.1}{E} \quad (4.27)$$

L'histogramme de contrôle de l'application de cette paramétrisation sur la masse invariante de la paire jj est reporté sur la figure 4.16. L'ajustement par une gaussienne sur l'intervalle $[70 \text{ GeV}/c^2, 90 \text{ GeV}/c^2]$ avant (eq. 4.28) et après (eq 4.29) étalonnage donne :

$$\begin{aligned} M_{jj}^{RAW} &= 78.5 \pm 0.1 \text{ GeV}/c^2 \\ \sigma_{jj}^{RAW} &= 8.8 \pm 0.3 \text{ GeV}/c^2 \end{aligned} \quad (4.28)$$

$$\begin{aligned} M_{jj}^{CALIB} &= 80.6 \pm 0.1 \text{ GeV}/c^2 \\ \sigma_{jj}^{CALIB} &= 8.1 \pm 0.2 \text{ GeV}/c^2 \end{aligned} \quad (4.29)$$

4.3 Etalonnage des jets de b *via* la résonance du boson Z^0

La reconstruction de résonances W en paires de jets légers a permis d'établir une stratégie pour l'établissement de l'échelle en énergie de ces jets. Les jets de b purement hadroniques devraient pouvoir être étalonnés par la fonction d'étalonnage obtenue dans le cadre des jets légers. Cependant cette fonction ne peut s'appliquer directement à tous les jets issus de quarks b . En effet, les désintégrations semi-leptoniques du quark b contribuent à l'énergie manquante et dégradent la reconstruction de l'énergie des jets de b .

On peut cependant adapter la technique utilisée pour les jets légers à la désintégration du boson Z^0 dans le canal $b\bar{b}$.

Contrairement à l'étalonnage des jets légers dont la sélection était aisée, la problématique qui est posée ici est différente dans la mesure où la reconstruction de Z^0 en paires $b\bar{b}$, est rendue plus ardue par la présence du bruit de fond QCD. L'approche choisie a donc été de produire le Z^0 en association avec un autre objet qui servira de déclenchement pour le système d'acquisition. Les différents canaux envisagés et leur section efficace obtenue avec Pythia 6.203 sont reportés dans le tableau 4.5. Nous pouvons remarquer que le bruit de fond QCD se situe à au moins 3 ordres de grandeur au-dessus du canal le plus

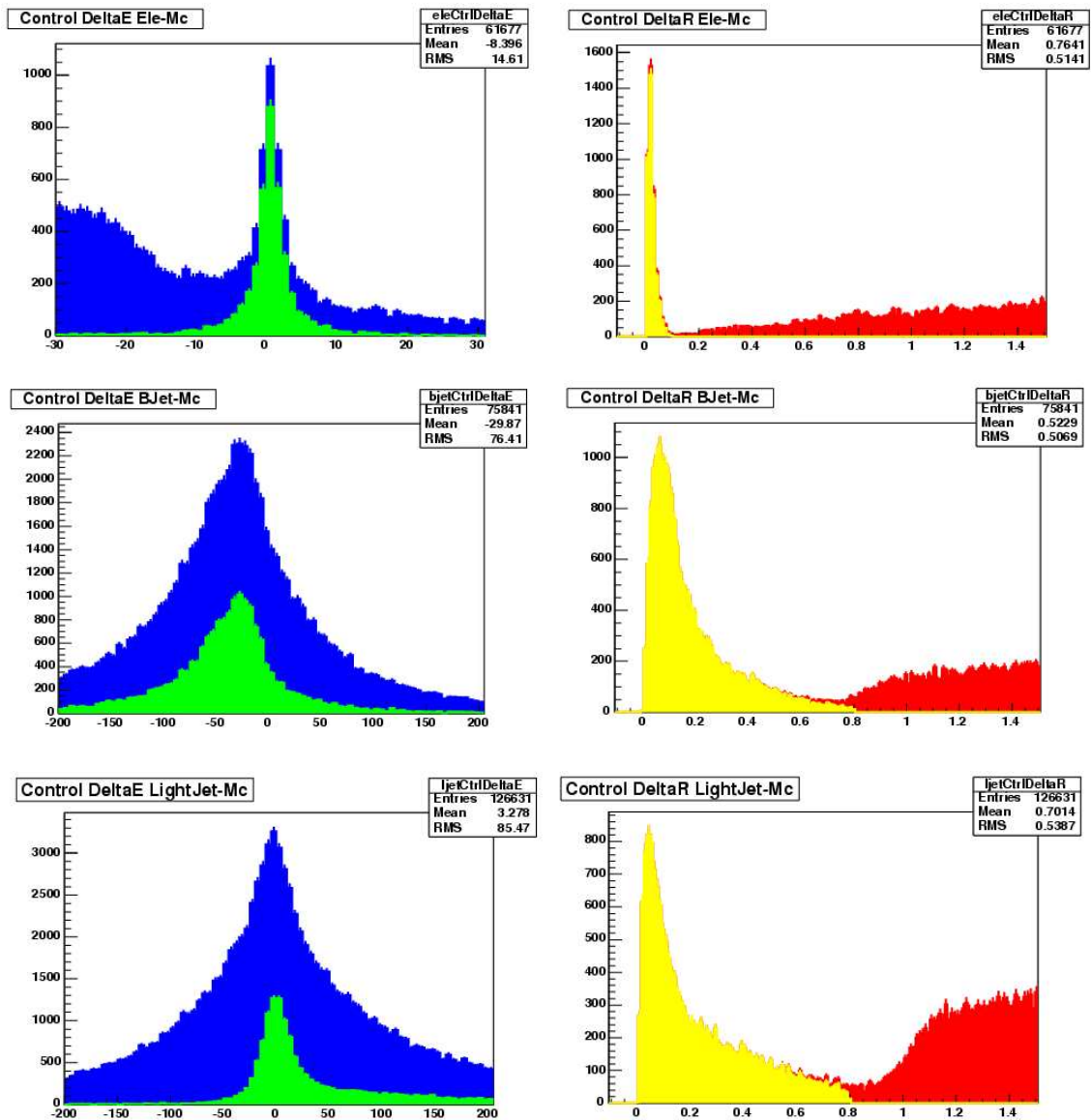


FIG. 4.14 – Histogrammes de contrôle du traitement de la superposition des collections. La première ligne comporte les informations relatives à la collection d'électrons et les deux dernières relèvent des jets étiquetés b et non étiquetés, respectivement. La colonne de gauche représente la distribution de la différence en énergie entre un objet reconstruit et son association Monte-Carlo, avant (en bleu) et après (en vert) nettoyage. Dans la colonne de droite est reportée la distance ΔR entre un objet reconstruit et son association Monte-Carlo avant (en rouge) et après (en jaune) nettoyage.

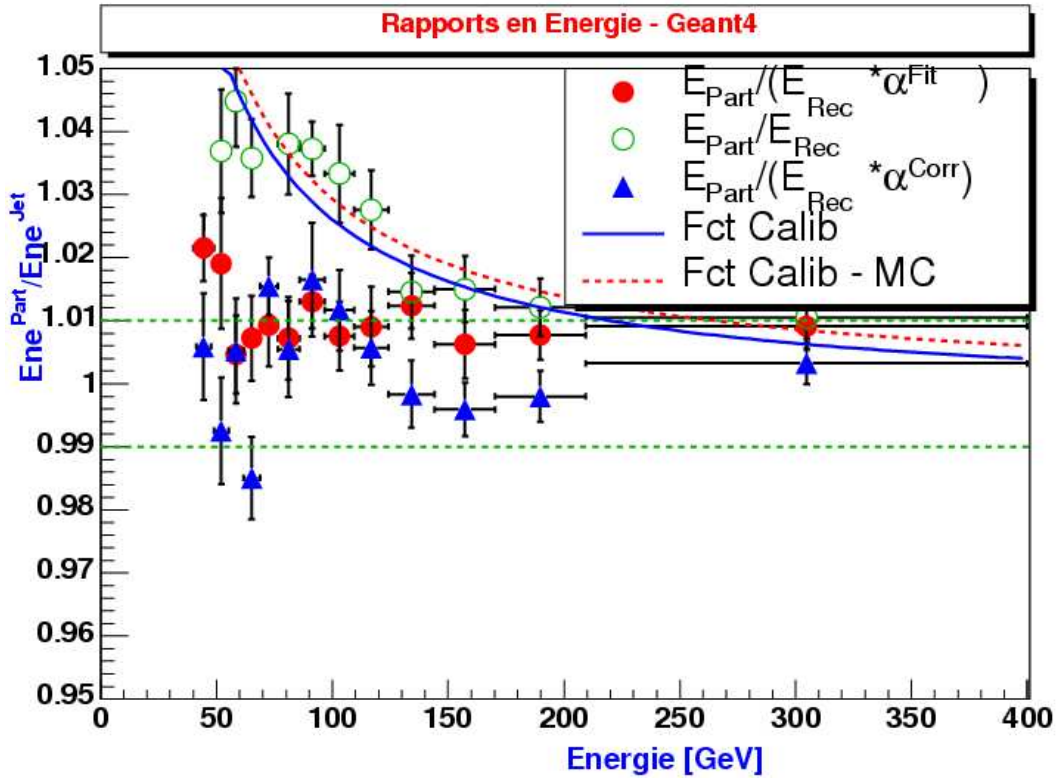


FIG. 4.15 – Rapports entre l'énergie des partons initiaux sur l'énergie des jets légers, pour différentes étapes de l'étalonnage avec l'ajustement par MINUIT. Les deux lignes horizontales en pointillés verts représentent les limites à $\pm 1\%$ de la connaissance de l'échelle absolue en énergie. Les cercles verts vides représentent les rapports en énergie entre le parton initial et l'énergie du jet avant étalonnage. Les cercles rouges pleins représentent les rapports en énergie entre le parton et l'énergie du jet corrigée par le facteur déterminé événement par événement par l'ajustement du χ^2 . Les triangles bleus représentent le même rapport, l'énergie du jet étant corrigée par la fonction ajustant l'évolution des poids (voir équation 4.4). Cette fonction est matérialisée par la ligne continue bleue. La ligne rouge en pointillés représente l'ajustement des poids Monte-Carlo (les cercles verts vides) par l'équation 4.4. Ces rapports ont été obtenus à partir de l'échantillon $t\bar{t}$ 4100 simulé *via* Geant4 et reconstruit avec la reconstruction complète à l'occasion de l'Atelier de Physique de Rome [30].

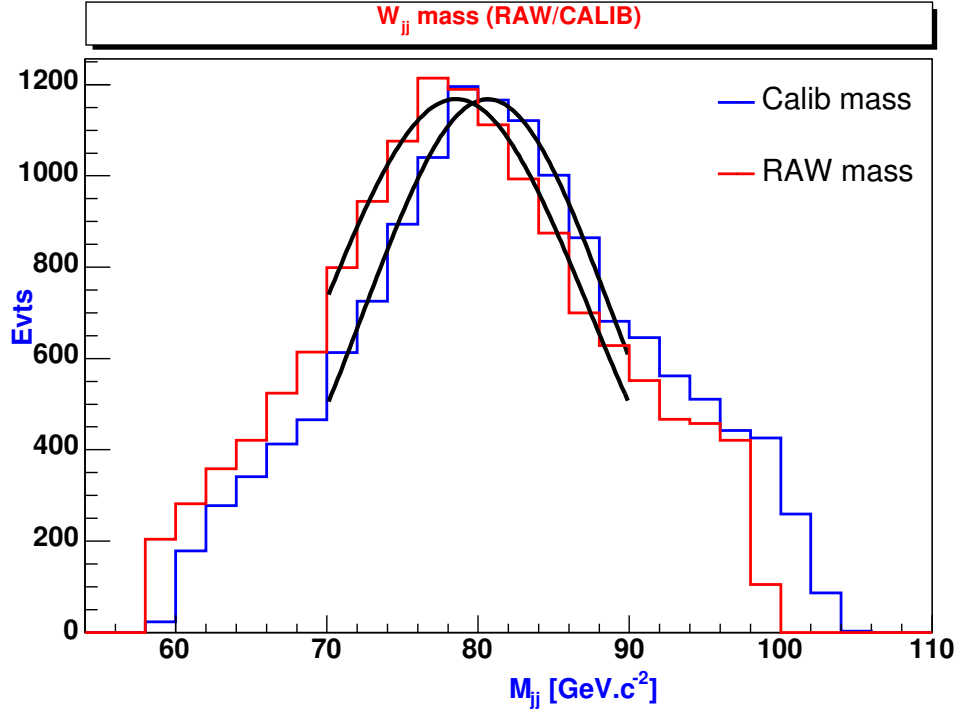


FIG. 4.16 – Masse invariante de la paire jj avant (en rouge à gauche) et après (en bleu à droite) application de la paramétrisation de la fonction d'étalonnage. Les valeurs obtenues lors de l'ajustement par les gaussiennes sont données dans les equations 4.28 et 4.29.

Processus	section efficace (fb)	Nbre évts attendus ($\mathcal{L} = 10fb^{-1}$)
$Z(b\bar{b}) + j$	2.3×10^6	23×10^6
$Z(b\bar{b})\gamma$	7.1×10^3	71×10^3
$Z(b\bar{b})W$	3.4×10^3	34×10^3
$Z(b\bar{b})Z$	1.4×10^3	14×10^3

TAB. 4.5 – Sections efficaces obtenues pour la production de bosons $Z^0 \rightarrow b\bar{b}$. Chacun des processus a été généré avec Pythia 6.203 en demandant un $\hat{p}_T > 20 GeV$ et au moins un Z^0 se désintégrant en une paire de quarks b .

abondant ce qui laisse présager de la difficulté d'extraction du signal. Malgré une section efficace très importante, le canal $Z + jet$ a été écarté du fait du bruit de fond multi-jet qui lui est associé. Nous nous sommes alors dirigés vers le canal $Z\gamma$ qui allie taux de production important et facilités de déclenchement et de reconstruction.

4.3.1 Étude du canal $Z^0\gamma \rightarrow b\bar{b}\gamma$

Dans ce canal, l'objet associé sur lequel le système d'acquisition va déclencher est un photon. Le menu de déclenchement utilisé est donc le menu HLT $\gamma 60i$ qui demande un photon isolé d'énergie transverse d'au moins $60 GeV$. Les bruits de fond considérés pour ce canal sont le bruit de fond QCD $b\bar{b}$ ainsi que le γjj . Les sections efficaces calculées avec Pythia 6.203 et AlpGen 1.3.3 (respectivement pour le $b\bar{b}$ et le signal d'une part et

Processus	Section efficace (fb)	Evts pour 10 fb^{-1}
bb	3.1×10^9	31×10^9
γjj	33.5×10^6	335×10^6
$Z\gamma$	7.1×10^3	71×10^3

TAB. 4.6 – Tableau des sections efficaces des bruits de fond considérés et du signal ($Z(b\bar{b}) + \gamma$). Le nombre d'événements attendus pour un an de prise de données à basse luminosité est également reporté.

le γjj d'autre part) sont reportées dans le tableau 4.6.

Le nombre d'événements de bruit de fond à générer étant très important, la simulation rapide du détecteur a été choisie afin d'obtenir le plus rapidement possible une statistique suffisante et représentative. La procédure de présélection a été effectuée comme suit :

- sélection des événements contenant au moins un photon d'énergie transverse de $60 GeV$ (la courbe d'efficacité de déclenchement (*turn-on curve*) est donc modélisée par une Heaviside), ceci représente la coupure $HLT\gamma 60i$,
- sélection des événements ayant au moins deux jets dont au moins un étant labellé b (Sel1),
- sélection des événements ayant au moins deux jets d'impulsion transverse $p_T > 40 GeV/c$ (Sel2).

Les efficacités de ces coupures sont reportées dans le tableau 4.7 pour différentes efficacités d'étiquetage de jets issus de b . L'efficacité de sélection HLT est évidemment stable en fonction de l'efficacité d'étiquetage des b , alors que l'efficacité des coupures 1 et 2 augmente lorsque ϵ_b augmente. La coupure 1 étant sensible au nombre de jets étiquetés b , l'amélioration de l'étiquetage augmente évidemment la multiplicité des jets de b reconstruits. Les figures 4.17, 4.18 et 4.19 illustrent les coupures de présélection. Nous pouvons remarquer dans le tableau 4.7 que la coupure Sel2, appliquée en prévision de la procédure d'étalonnage en aval de cette présélection, dégrade la signification. Nous nous sommes alors orientés vers une autre coupure, Sel2', qui demande que le jet de plus haute impulsion transverse satisfasse : $p_T > 40 GeV/c$. Ceci permet alors d'obtenir les efficacités de sélection reportées dans les trois dernières lignes du tableau 4.7 (et ce pour chaque efficacité d'étiquetage des jets de b). La signification maximale de 4.381 est obtenue pour l'efficacité d'étiquetage des jets issus de b la plus basse. Cet effet peut paraître contre-intuitif mais sera expliqué par la suite.

La reconstruction de ces événements présélectionnés se déroule comme suit. Les événements sont divisés en deux catégories : ceux qui présentent au moins deux jets identifiés de b (labellés $2b$) et ceux n'ayant qu'un seul jet de b reconstruit (labellés bl). La stratégie de reconstruction pour les événements $2b$ consiste à apparier les deux jets labellés b de plus haute impulsion transverse pour former le candidat Z^0 . Ceci se traduit par les masses invariantes reportées dans les graphiques reportés en haut à droite sur les figures 4.20, 4.21 et 4.22 pour différentes efficacités d'étiquetage de jets de b .

Pour les événements bl , plusieurs stratégies ont été investiguées :

- $HiPt$: combinaison des deux jets de plus haute impulsion transverse,
- $MinDeltaR_{jj}$: appariement des deux jets les plus proche en ΔR ,
- $MinDeltaPtZG$: combinaison des deux jets réalisant le meilleur balancement en impulsion transverse avec le photon de plus haute impulsion transverse,

$\epsilon_b = 50\% R_u = 200$					
Coupure	$Z\gamma$		γjj		S/\sqrt{B}
	eff (%)	Nbr evts	eff (%)	Nbr evts	
$\gamma 60i$	9.33 ± 0.03	6640 ± 20	9.38 ± 0.01	$(3.14 \pm 0.01) \times 10^7$	1.19
Sel1	4.15 ± 0.02	2960 ± 20	0.27 ± 0.01	$(8.88 \pm 0.05) \times 10^5$	3.14
Sel2	1.54 ± 0.01	1100 ± 10	0.12 ± 0.01	$(3.95 \pm 0.04) \times 10^5$	1.75
Sel2'	3.92 ± 0.07	2790 ± 50	0.24 ± 0.01	$(7.94 \pm 0.09) \times 10^5$	3.13
m_{bb}	3.61 ± 0.07	2570 ± 50	0.17 ± 0.01	$(5.59 \pm 0.08) \times 10^5$	3.43
m_{bb} in range	2.62 ± 0.06	1860 ± 40	0.05 ± 0.01	$(1.80 \pm 0.01) \times 10^5$	4.38

$\epsilon_b = 60\% R_u = 100$					
Coupure	$Z\gamma$		γjj		S/\sqrt{B}
	eff (%)	Nbr evts	eff (%)	Nbr evts	
$\gamma 60i$	9.33 ± 0.03	6640 ± 20	9.38 ± 0.01	$(3.14 \pm 0.01) \times 10^7$	1.19
Sel1	4.82 ± 0.02	3430 ± 20	0.39 ± 0.01	$(1.32 \pm 0.01) \times 10^6$	3.00
Sel2	1.80 ± 0.01	1280 ± 10	0.19 ± 0.01	$(6.35 \pm 0.04) \times 10^5$	1.61
Sel2'	4.56 ± 0.08	3240 ± 60	0.36 ± 0.01	$(1.20 \pm 0.01) \times 10^6$	2.95
m_{bb}	4.20 ± 0.08	2980 ± 50	0.25 ± 0.01	$(8.42 \pm 0.01) \times 10^5$	3.25
m_{bb} in range	3.04 ± 0.06	2160 ± 50	0.08 ± 0.01	$(2.67 \pm 0.01) \times 10^5$	4.17

$\epsilon_b = 70\% R_u = 30$					
Coupure	$Z\gamma$		γjj		S/\sqrt{B}
	eff (%)	Nbr evts	eff (%)	Nbr evts	
$\gamma 60i$	9.33 ± 0.03	6640 ± 20	9.38 ± 0.01	$(3.14 \pm 0.01) \times 10^7$	1.19
Sel1	5.48 ± 0.02	3900 ± 20	0.78 ± 0.01	$(2.62 \pm 0.01) \times 10^6$	2.41
Sel2	2.04 ± 0.01	1450 ± 10	0.33 ± 0.01	$(5.88 \pm 0.06) \times 10^5$	1.37
Sel2'	5.17 ± 0.08	3670 ± 60	0.70 ± 0.01	$(2.34 \pm 0.01) \times 10^6$	2.40
m_{bb}	4.77 ± 0.08	3390 ± 60	0.50 ± 0.01	$(1.68 \pm 0.01) \times 10^6$	2.62
m_{bb} in range	3.43 ± 0.07	2430 ± 50	0.16 ± 0.01	$(5.42 \pm 0.01) \times 10^5$	3.31

TAB. 4.7 – Efficacités de sélection des événements pour le bruit de fond et le signal. Les coupures sont appliquées de manière séquentielle : chaque ligne correspond au pourcentage d'événements survivant après la coupure par rapport au nombre total d'événements analysés. La première ligne correspond à l'application du déclenchement HLT, les deuxième et troisième lignes représentent les coupures “2 jets dont au moins un étiqueté b ” et “2 jets dont l'impulsion transverse est d'au moins 40 GeV/c ” respectivement. La ligne Sel 2' correspond à la coupure “au moins 1 jet d'impulsion transverse d'au moins 40 GeV/c ”. Les deux dernières lignes correspondent aux événements pour lesquels la masse du boson candidat reconstruit est inférieure à 200 GeV/c^2 et comprise dans l'intervalle $[60 GeV/c^2; 110 GeV/c^2]$, respectivement. Le nombre d'événements retenus après chaque étape de la sélection a été normalisé au nombre d'événements attendus pour une luminosité $\mathcal{L} = 10 fb^{-1}$.

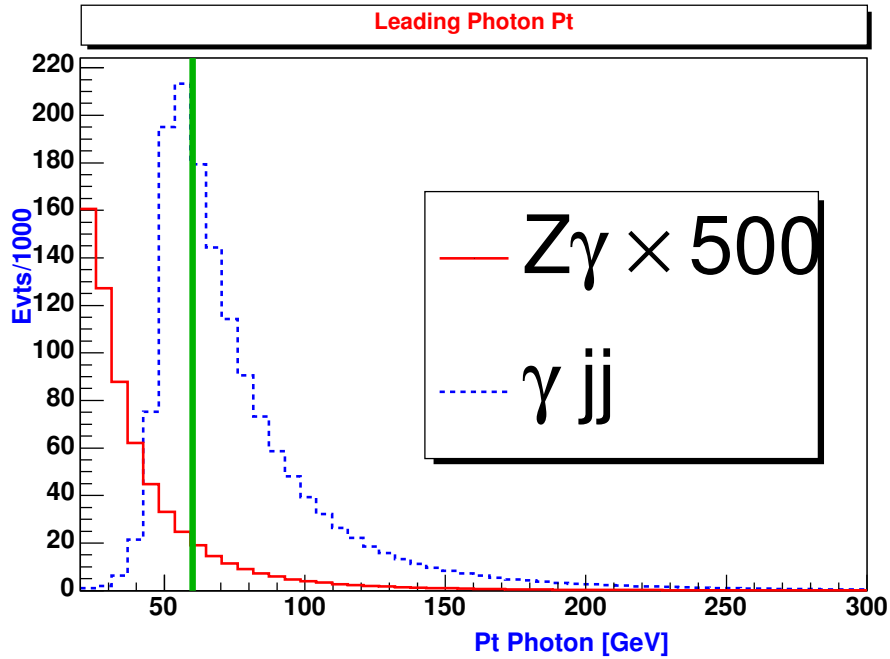


FIG. 4.17 – Distribution de l'impulsion transverse du photon de plus haute énergie transverse avant coupure HLT. Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus. La coupure appliquée par le menu $\gamma 60i$ est représentée par le trait plein vert vertical. Nous pouvons noter l'effet de la génération du bruit de fond avec $\hat{p}_T > 50$ GeV.

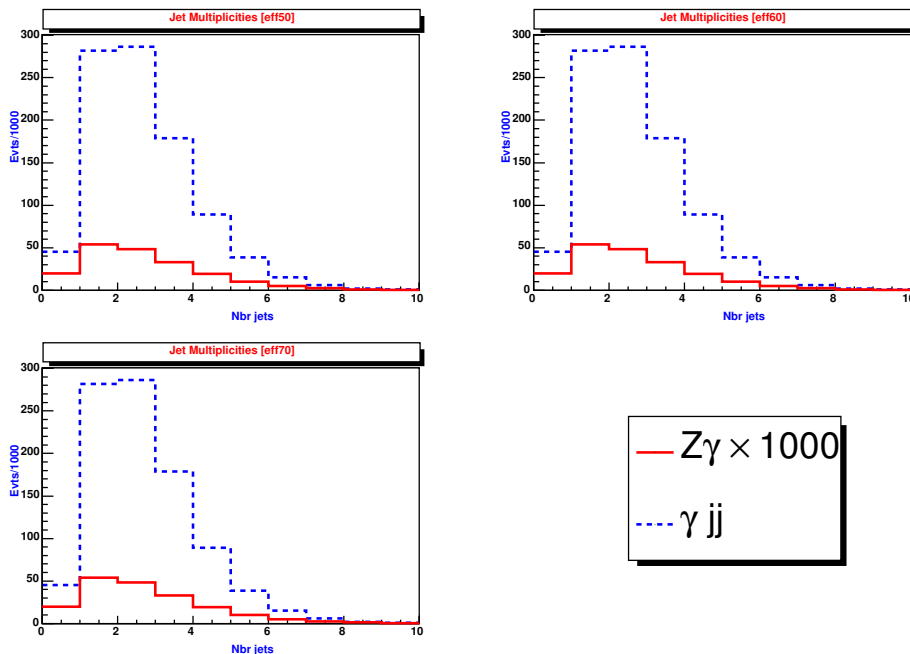


FIG. 4.18 – Multiplicités des jets (toutes saveurs confondues) après coupure HLT. Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus.

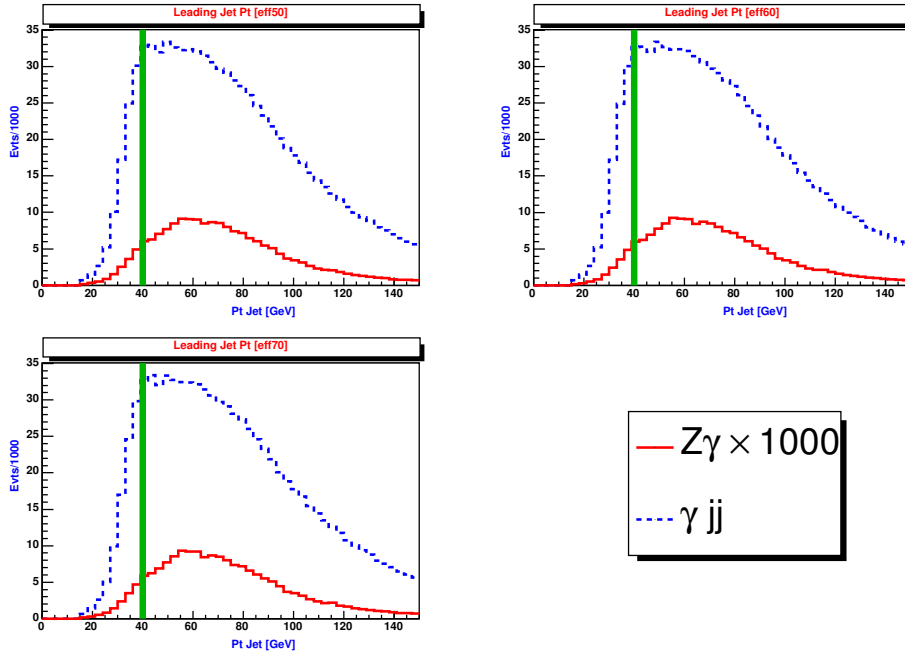


FIG. 4.19 – Distribution de l’impulsion transverse du jet de plus haute énergie transverse après coupure HLT et pour différentes efficacités d’étiquetage de jets de b ($\epsilon_b = 50\%$ en haut à gauche, 60% en haut à droite et 70% en bas à gauche). Le signal est en traits pleins rouges et le bruit de fond en traits pointillés bleus.

- *HiMom* : combinaison des deux jets de plus haute impulsion,
- *HiPtMinDeltaEta_{jj}* : combinaison des deux jets de plus haute impulsion transverse et les plus proche en η .

La masse invariante des bosons candidats correspondant à la superposition des événements $2b$ et bl (pour la stratégie *HiPt*) est reportée sur les histogrammes en bas à gauche sur les figures 4.20, 4.21 et 4.22.

Nous avons ensuite ajusté par une gaussienne chacun des pics en masse obtenus pour les événements de classe $2b$, bl et $2b + bl$. Comme nous pouvions nous y attendre, la meilleure résolution ($\sigma_{\epsilon_b=50\%}^{2b} = 13 \pm 1 \text{ GeV}/c^2$) est obtenue pour la classe $2b$ et la plus mauvaise ($\sigma_{\epsilon_b=50\%}^{bl} = 14 \pm 1 \text{ GeV}/c^2$) pour la classe bl . Ceci illustre le fait que la calibration appliquée par *At1Fast* sur les jets de b et les jets légers est différente. Ainsi, un jet de b n’aura pas la même énergie suivant qu’il sera étiqueté comme issu de b ou léger. D’autre part, et pour les mêmes raisons, nous reconstruisons un meilleur pic en masse (plus proche de la valeur générée) pour la classe bb .

Les valeurs des significances pour chaque classe d’événement ($2b$ et bl) et pour différentes valeurs de l’efficacité d’étiquetage des jets issus de b sont reportées dans le tableau 4.8. Nous remarquons que l’évolution de la signification diffère entre les deux classes d’événements. La signification dans la fenêtre en masse pour la classe bl diminue lorsque l’efficacité d’étiquetage des jets issus de b augmente. Ceci explique les valeurs de significances que nous relevons dans le tableau 4.7 et qui étaient alors la manifestation de cette compétition entre les deux classes d’événements.

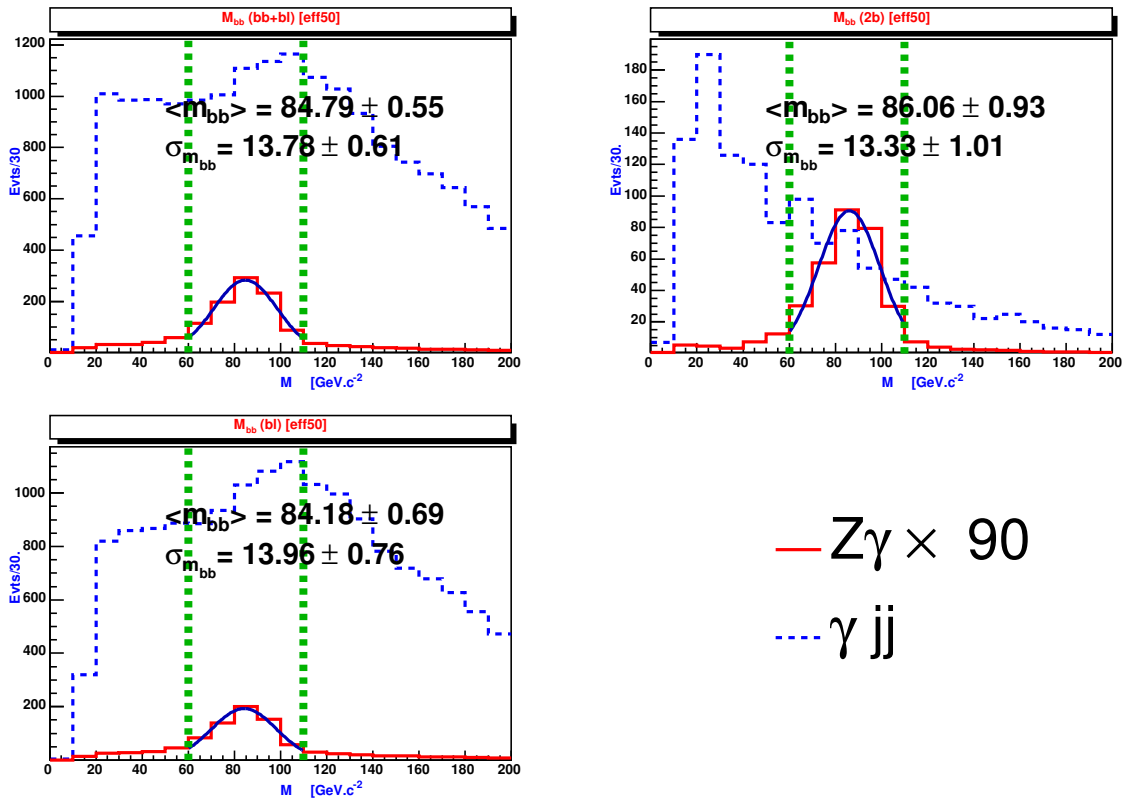


FIG. 4.20 – Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 50%.

Coupe	$\epsilon_b = 50\%$	$\epsilon_b = 60\%$	$\epsilon_b = 70\%$
classe $2b$			
m_{bb}	3.44	3.91	3.15
m_{bb} in range	5.41	6.14	4.50
classe bl			
m_{bb}	2.60	2.26	1.61
m_{bb} in range	3.11	2.66	1.84

TAB. 4.8 – Valeurs des significances S/\sqrt{B} pour les efficacités d'étiquetage des jets issus de b pour les coupures $m_{bb} < 200 \text{ GeV}/c^2$ et $60 \text{ GeV}/c^2 \leq m_{bb} < 110 \text{ GeV}/c^2$.

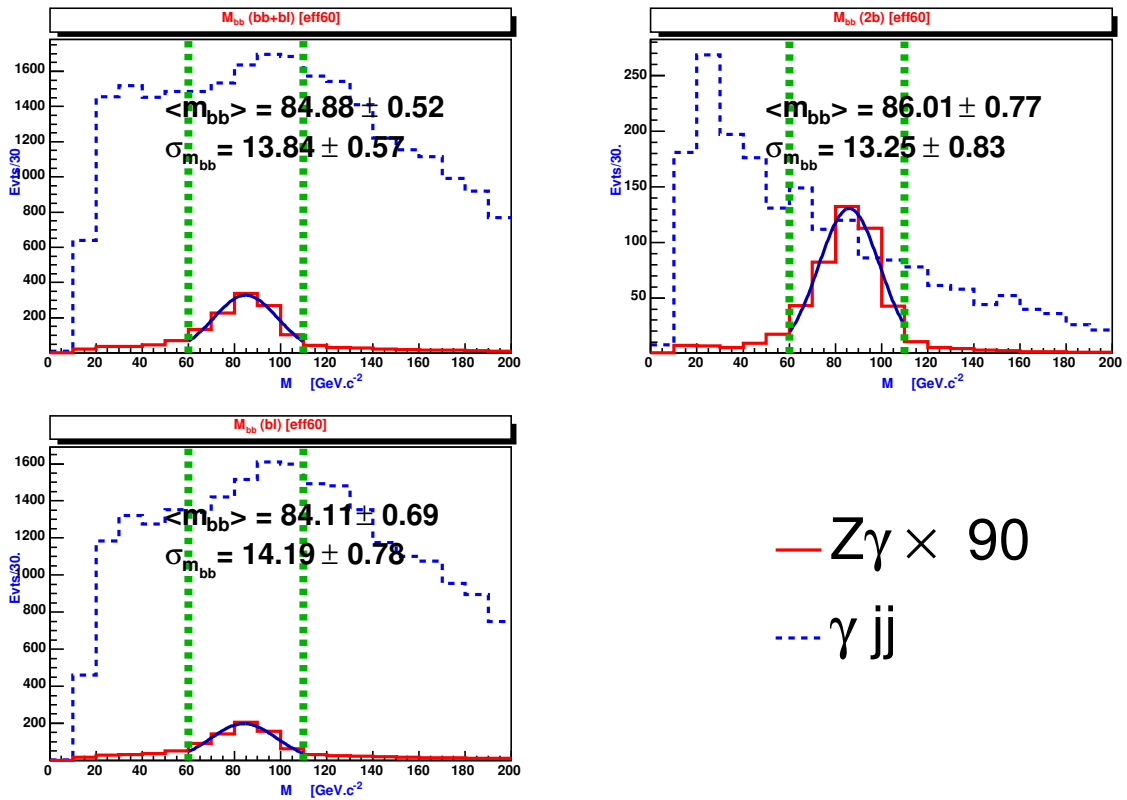


FIG. 4.21 – Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 60%.

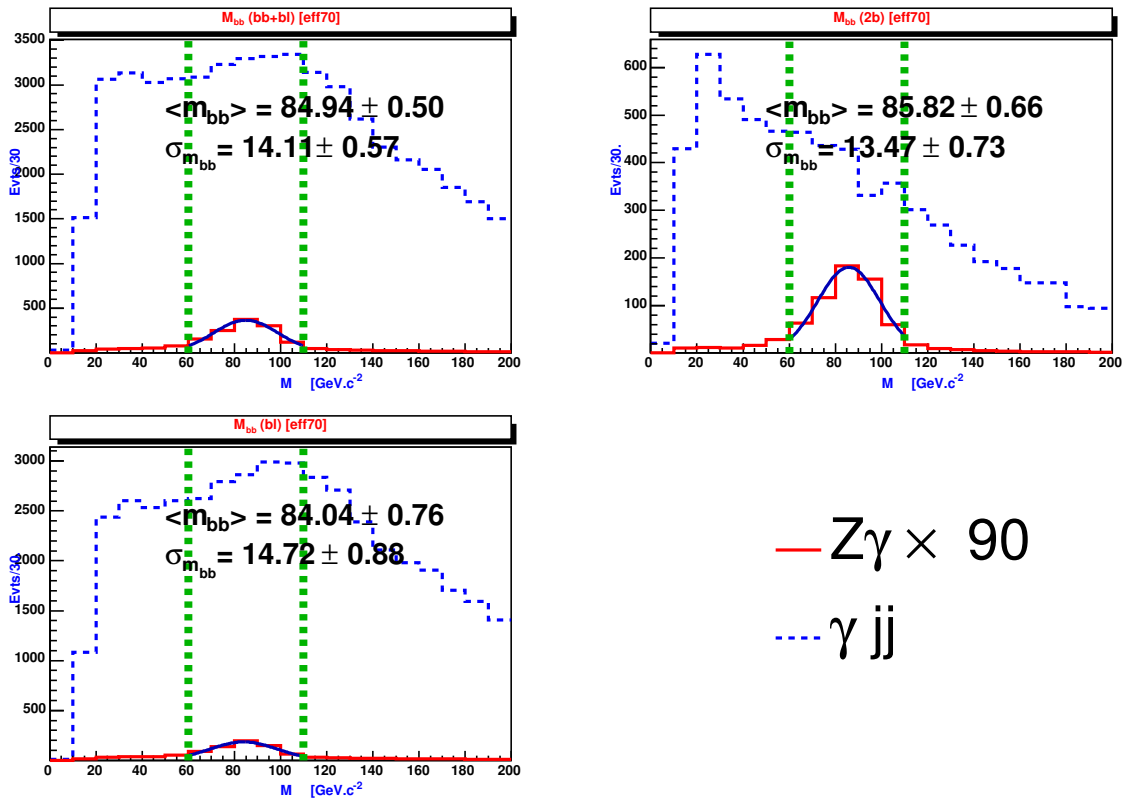


FIG. 4.22 – Masses invariantes pour les classes d'événements $2b$ et bl (en haut à gauche), $2b$ (en haut à droite) et bl (en bas à droite). Le signal est en traits pleins rouges et le bruit de fond est en traits pointillés bleus. L'efficacité d'étiquetage des b est de 70%.

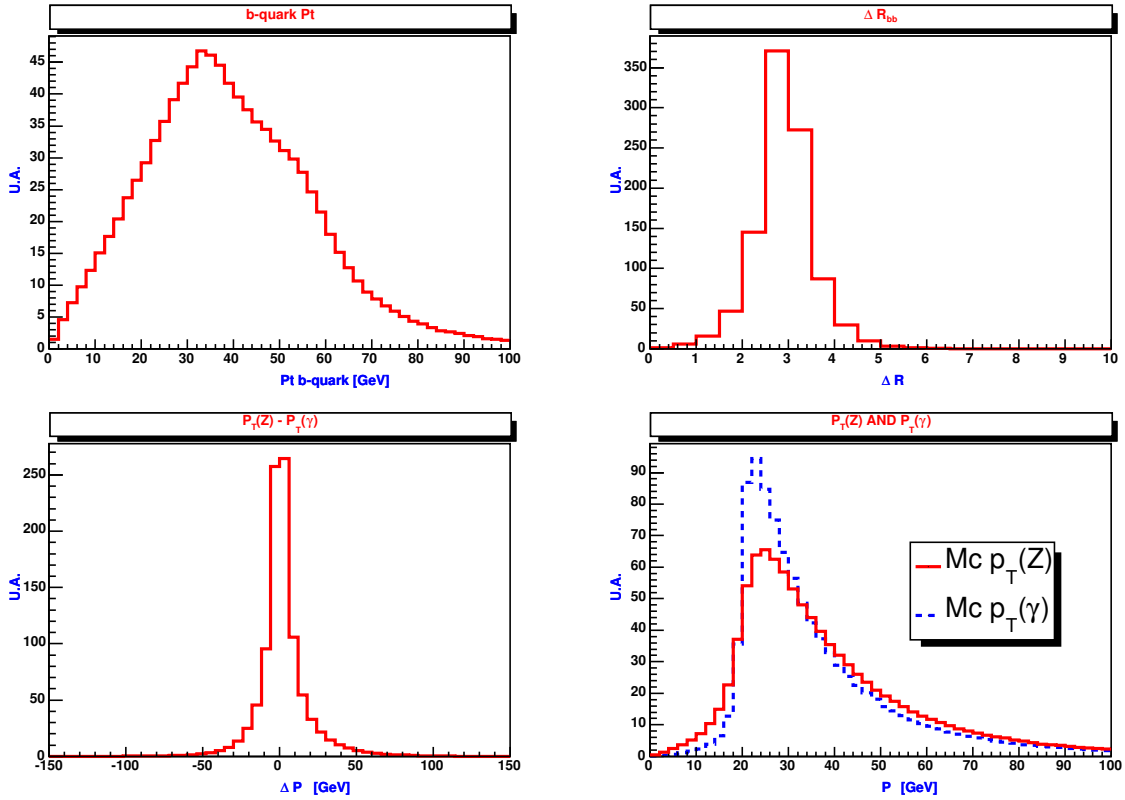


FIG. 4.23 – Distributions de l’impulsion transverse des quarks provenant du Z^0 en haut à droite et du rayon du cône reliant les quarks b issus du boson Z^0 en haut à droite. Différence entre les impulsions transverses du Z^0 et du photon et leur superposition, respectivement en bas à droite et en bas à gauche.

4.3.1.1 Etude de stratégie de reconstruction

La première étape de validation de l’algorithme de reconstruction a été de construire les différentes distributions caractérisant les quarks issus du Z^0 ainsi que les grandeurs reliant le Z^0 et le γ afin d’étudier les performances d’une stratégie de reconstruction. Les distributions de la figure 4.23 ont permis de confirmer l’orientation vers une reconstruction privilégiant les jets de haut p_T sans toutefois pouvoir utiliser le cône d’ouverture entre les quarks ou bien la balance en impulsion transverse entre le photon et le boson Z^0 .

La deuxième étape de la validation a été d’apparier les jets aux quarks contenus dans la collection Monte-Carlo. Le critère d’appariement retenu est une coupure en $\Delta R < 0.8$. Une attention toute particulière a été portée à l’algorithme d’association afin de ne pas introduire de double comptage du à un double appariement : chaque quark est associé une et une seule fois à un seul jet. Cette association est d’abord utilisée afin d’obtenir les grandeurs reconstruites si l’efficacité d’étiquetage était de 100% et ainsi connaître l’efficacité et la pureté des différents algorithmes de reconstruction étudiés. Les différentes efficacités d’association sont reportées dans le tableau 4.9. Afin de valider l’algorithme d’association, le nombre d’événements pour lesquels au moins une association quark-jet puis deux associations quark-jet a été calculé et reporté dans les deux premières lignes du tableau 4.9. Nous avons vérifié que les quarks pour lesquels aucune association n’avait

Associations	$\epsilon_b = 50\%$	$\epsilon_b = 60\%$	$\epsilon_b = 70\%$
1 association $q \rightarrow j$	$(99.97 \pm 0.01)\%$	$(99.97 \pm 0.01)\%$	$(99.96 \pm 0.01)\%$
2 associations $q \rightarrow j$	$(87.1 \pm 0.3)\%$	$(86.7 \pm 0.3)\%$	$(86.1 \pm 0.3)\%$
jets ass. : bb	$(32.3 \pm 0.4)\%$	$(40.1 \pm 0.4)\%$	$(49.0 \pm 0.4)\%$
jets ass. : bl	$(67.7 \pm 0.4)\%$	$(59.9 \pm 0.4)\%$	$(51.0 \pm 0.4)\%$
ass. bl HiPt	$(73.4 \pm 0.5)\%$	$(73.0 \pm 0.5)\%$	$(72.1 \pm 0.5)\%$
ass. bl NoHiPt	$(26.6 \pm 0.5)\%$	$(27.1 \pm 0.5)\%$	$(27.9 \pm 0.5)\%$

TAB. 4.9 – Efficacités d’association pour les différentes efficacités d’étiquetage des b pour le signal. La première et la deuxième ligne représentent le nombre d’associations quark-jet dans les événements signaux. Les deux lignes suivantes donnent la répartition de la saveur des jets associés aux quarks parmi les événements pour lesquels deux associations ont pu être trouvées. Les deux dernières lignes donnent enfin le pourcentage d’événements parmi les bl pour lesquels le jet léger associé est le jet de plus haute impulsion transverse ou pas.

été trouvée correspondaient à des jets ne passant pas la coupure en impulsion transverse (cf figure 4.24). Les deux lignes suivantes représentent le nombre d’événements pour lesquels les jets associés aux quarks ont été labellés beaux ou légers. Le pourcentage d’événements bb augmente bien avec l’efficacité d’étiquetage pour atteindre environ 50% avec une efficacité d’étiquetage des jets issus de b de 70%. Nous retrouvons alors la classification des événements en deux grandes catégories : les événements pour lesquels le boson sera construit à partir de la combinaison de deux jets étiquetés b et ceux pour lesquels un seul b -jet est disponible. L’exercice a ensuite consisté, pour les événements dont les jets associés sont de saveurs différentes, à repérer la position du jet léger associé dans la collection de jets légers, classés par impulsion transverse décroissante. Nous remarquons ainsi que dans environ un tiers des cas le jet léger de plus haute impulsion transverse n’est pas celui que nous devons prendre pour construire le Z^0 . Ce changement dans le classement des jets est dû à l’insertion d’un jet provenant d’une radiation de gluon dans l’état final. Nous avons alors étudié plusieurs méthodes de reconstruction afin d’essayer de récupérer ces événements. En effet, la stratégie de reconstruction basée uniquement sur l’impulsion transverse n’est pas adéquate pour ces événements.

Les différentes stratégies n’ont néanmoins pas permis de récupérer ces 30% d’événements pour lesquels le jet léger de plus haut p_T ne provient pas, selon l’association Monte-Carlo, du boson Z^0 . La figure 4.25, dans laquelle la distribution de la différence en ϕ entre un quark b et son compagnon est reportée pour différentes stratégies d’appariement, illustre cet état de fait.

La première conclusion à tirer de cette reconstruction du $Z\gamma$ est qu’en l’état actuel du jeu de coupures de présélections, il n’est pas possible d’extraire le signal du bruit de fond irréductible γbb . Le différentiel appliqué par la première coupure, le dernier niveau de déclenchement, n’est jamais rattrapé dans la suite de l’analyse.

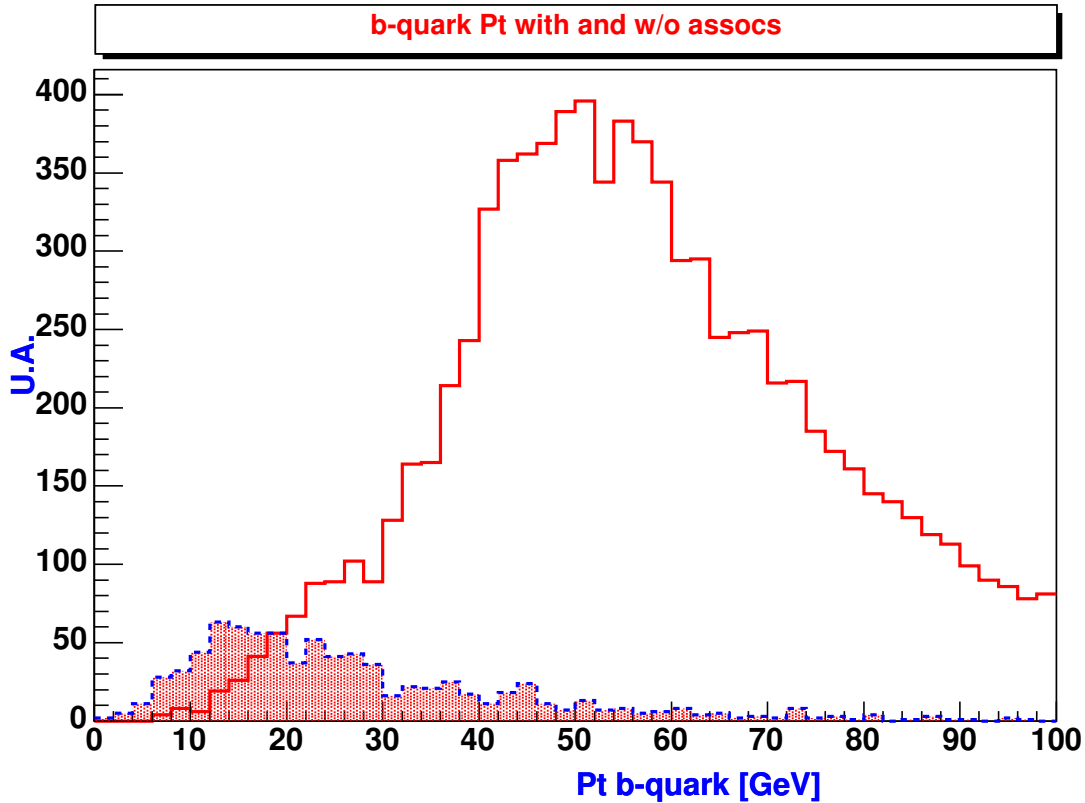


FIG. 4.24 – Distributions de l’impulsion transverse des quarks issus du boson Z^0 lorsqu’une association est trouvée (en traits pleins rouges) et lorsqu’aucune association n’a été trouvée. Nous pouvons noter l’effet de la coupure sur les jets ($p_T > 40$ GeV) qui dégrade la probabilité d’association.

Efficacité	$Z\gamma$	γbb	S/\sqrt{B}
$\epsilon_b = 50\%$	70.3%	22.5%	4.381
$\epsilon_b = 60\%$	70.3%	22.6%	4.172
$\epsilon_b = 70\%$	70.8%	21.7%	3.305

TAB. 4.10 – Tableau des efficacités de reconstruction du Z^0 pour le signal et le bruit de fond en fonction de l’efficacité d’étiquetage des jets de b . Ces efficacités correspondent au nombre d’événements compris dans la fenêtre en masse $[60; 110]$ et pour les classes d’événements $2b$ ou bl . La dernière colonne représente la signification correspondant aux nombres d’événements attendus dans la fenêtre en masse pour $10 fb^{-1}$.

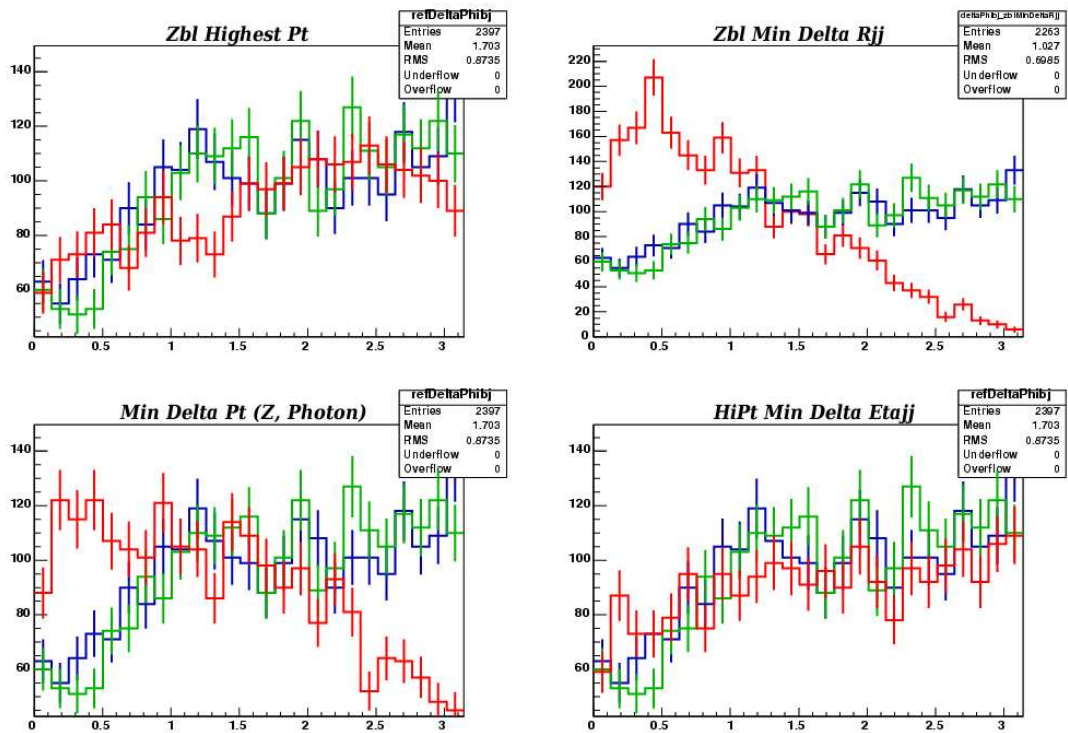


FIG. 4.25 – Différence entre les angles azimutaux du quark b et son compagnon pour le Monte-Carlo (vert), les jets associés aux quarks (bleu) et les jets appariés par les différentes stratégies de reconstruction bl (rouge).

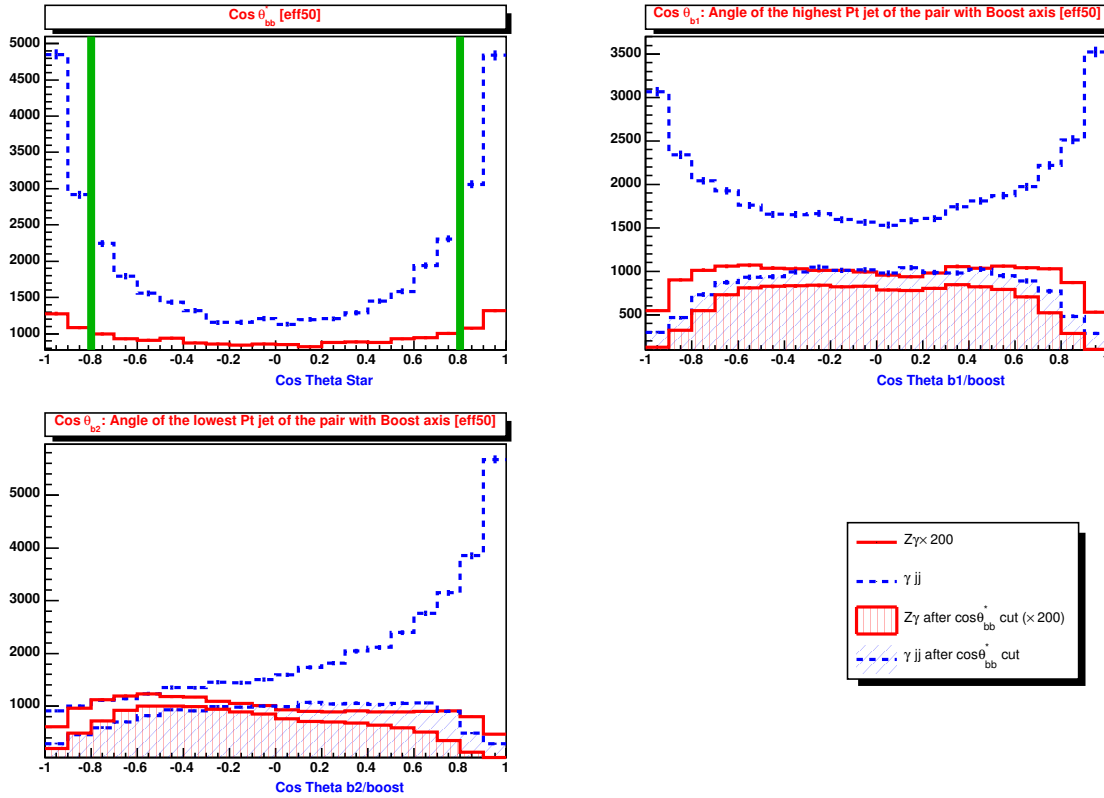


FIG. 4.26 – Distribution du cosinus de l’angle de la paire de jets avec l’axe faisceau dans le référentiel du Z^0 reconstruit (en haut à gauche). Le signal est en traits pleins rouges et le bruit de fond en pointillés bleus. La figure en haut à droite (en bas à gauche) représente la distributions du cosinus de l’angle entre la direction de la poussée du Z^0 reconstruit avec le jet de plus haute impulsion transverse (de plus petit p_T). Les aires hachurées correspondent à l’application de la coupure $|\cos \theta^*| < 0.8$.

4.4 Conclusion

Nous avons vu que l’extraction du signal $Z\gamma$ où le Z^0 se désintègre en une paire de quarks b n’est pas possible en utilisant le système de déclenchement HLT ainsi qu’un jeu de coupures séquentielles. En effet, la signficance maximale ($= 6.143$) obtenue avec une efficacité d’étiquetage des jets de b de 60% et en n’utilisant que les événements de la classe $2b$, ne permet pas de reconstruire le pic du Z . Une étude faisant usage de fonctions de distributions de probabilités (PDF) ou de réseaux de neurones est cependant envisageable. En effet, la coupure sur une variable telle que le cosinus de l’angle entre l’axe faisceau et la paire de jets dans le référentiel du Z^0 (figure 4.26) semble être intéressante, même si en l’état actuel de notre jeu de coupures, ne permet pas d’améliorer la signficance. Nous avons constaté que cette coupure permettait d’améliorer le rapport signal sur bruit, mais dégradait la signficance. Une étude utilisant le canal ZW est elle aussi envisageable. Cependant, la toxicité du bruit de fond $W + jets$ doit être évaluée.

Conclusions

Dans le cadre de ce travail, deux axes de recherche ont été développés. Nous avons tout d'abord mis en place un certain nombre d'outils logiciels qui ont ensuite été pour la plupart réutilisés dans l'étude de l'étalonnage des jets dans ATLAS.

Les outils pour les études de physique avec l'environnement logiciel d'ATLAS sont en passe d'arriver à maturité. Ils ont en effet été testés à grande échelle à l'occasion de l'atelier de physique de Rome en 2005. L'optimisation de l'arbre d'héritage des classes de données pour l'analyse a permis une meilleure factorisation du code modélisant les particules. Des outils pour l'analyse ont été développés et mis à disposition de la collaboration ATLAS. Les classes de filtrage des particules sur des critères cinématiques, de saveur et de charge permettent d'écrire facilement des séquences de sélection pour les particules issues du Monte-Carlo et de la reconstruction. Ces classes de filtrage sont ensuite réutilisées comme fondations pour les classes d'associations entre particules.

La maintenance du code de sélection des particules Monte-Carlo, ainsi que l'expertise acquise au cours de ce travail, ont permis de lancer les bases d'une refonte de l'organisation de ce paquetage. Cette nouvelle organisation, en cours de déploiement, tire d'avantage parti de l'héritage virtuel du C++ et de l'architecture en AlgTools d'ATHENA. Ce système de greffons (*plugins*) devrait permettre une maintenance plus aisée et une sélection plus robuste des particules au sein des événements Monte-Carlo. D'autre part, les classes développées dans le cadre du filtrage nécessaire à l'établissement d'une association entre deux objets d'analyse de physique (une particule, une trace, un amas de cellules, . . .), sont en cours d'intégration dans un cadre plus général. Ces classes seront en effet utilisées dans le contexte de l'étiquetage des superpositions d'objets et donc du double comptage de certains quadri-moments.

En parallèle, l'étude de l'étalonnage des jets légers dans l'environnement ATHENA a été effectuée. Ce travail a permis de mettre en place les outils logiciels nécessaires à l'extraction d'un des outils physiques fondamentaux pour les mesures de précision mettant en jeu des canaux comportant des jets dans l'état final. Ces outils seront à terme intégrés dans le code officiel d'ATLAS. L'étude de l'étalonnage *in situ* à l'aide de ces outils a permis de confirmer les résultats précédemment établis à partir de l'environnement logiciel ancienne génération et de données obtenues avec la simulation rapide. De plus, le développement d'une procédure d'ajustement contraint basé sur les multiplicateurs de Lagrange a permis d'accélérer la vitesse de convergence de cet ajustement. L'utilisation d'un lot de données obtenu avec la reconstruction complète du détecteur, équivalent à un peu plus d'un mois de prise de données à basse luminosité, indique que l'échelle en énergie peut être connue à 2%. L'étude d'un lot plus important serait intéressante afin de conforter le résultat de 1% obtenu avec la simulation rapide.

Enfin, la faisabilité de l'application aux jets de b de la méthode d'extraction de l'échelle

en énergie des jets légers, a été menée. L'application d'un jeu de coupures séquentielles n'a pas permis d'extraire du bruit de fond γjj le pic en masse du Z^0 dans le canal $Z^0(b\bar{b})\gamma$. Une étude utilisant des techniques plus avancées d'extraction du signal tel qu'un réseau de neurones pourrait optimiser la significativité de ce canal. L'introduction d'un menu de déclenchement pour l'étiquetage des b devrait également améliorer les possibilités d'extraction du signal.

Il reste que l'obtention de l'échelle en énergie des jets de b ne semble pas aisée en utilisant la méthode développée pour les jets légers. Il faudra également considérer les autres canaux (ZZ et ZW) avant de s'orienter le cas échéant vers d'autres méthodes. La modélisation de l'échelle en énergie des jets de b à partir de celle des jets légers serait en effet une étude intéressante à mener.

Bibliographie

- [1] The LEP Collaborations ALEPH, DELPHI, L3, OPAL and the LEP Electroweak Working Group, *A combination of preliminary Electroweak Measurements and constraints of the Standard Model*, CERN-PH-EP/2005-051
- [2] DONuT Collaboration, *Observation of tau neutrino interactions*, Phys. Lett. B504 218-224 (2001)
- [3] TheCDF Collaboration, the D0 Collaboration, and the TEVATRON Electroweak Working Group, *Combination of CDF and D0 Results on the Top-Quark Mass*, FERMILAB-TM-2323-E
- [4] ATLAS Collaboration, *ATLAS Technical Proposal*, CERN/LHCC 94-43 (1994)
- [5] CMS Collaboration, *CMS Technical Proposal*, CERN/LHCC 94-38 (1994)
- [6] LHCb Collaboration, *LHCb Technical Proposal*, CERN/LHCC 98-04 (1998)
- [7] ALICE Collaboration, *ALICE Technical Proposal*, CERN/LHCC 95-71 (1995)
- [8] Totem Collaboration, *Totem Technical Proposal*, CERN/LHCC 99-7 (1999)
- [9] ATLAS Collaboration, *ATLAS Letter of Intent for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*, CERN/LHCC/92-4 (1992)
- [10] S. Asai *et al.*, *Prospects for the search for a standard model Higgs boson in ATLAS using vector boson fusion*, Eur. Phys. J., C 32 (2004) p19-54, SN-ATLAS-2003-024
- [11] ATLAS Collaboration, *Inner Detector Technical Design Report*, Volume I, CERN/LHCC/97-16 (1997)
- [12] ATLAS Collaboration, *Pixel Detector Technical Design Report*, CERN/LHCC/98-13 (1998)
- [13] ATLAS Collaboration, *Inner Detector Technical Design Report*, Volume II, CERN/LHCC/97-17 (1997)
- [14] ATLAS Collaboration, *Liquid Argon Calorimeter Technical Design Report*, CERN/LHCC/96-41 (1996)
- [15] ATLAS Collaboration, *Tile Calorimeter Technical Design Report*, CERN/LHCC/96-42 (1996)
- [16] ATLAS Collaboration, *Muon Spectrometer Technical Design Report*, CERN/LHCC/97-22 (1997)
- [17] ATLAS Collaboration, *Trigger Performance Status Report*, CERN/LHCC/98-15 (1998)
- [18] ATLAS Collaboration, *ATLAS High-Level Trigger, Data Acquisition and Controls Technical Design Report*, CERN/LHCC 2003-069

- [19] ATLAS Collaboration, *First-Level Trigger Technical Design Report*, CERN/LHCC/98-14 (1998)
- [20] ATLAS Collaboration, *DAQ, EF, LVL2 and DCS Technical Progress Report*, CERN/LHCC/98-16 (1998)
- [21] *ATLAS C++ Coding Standard*, ATL-SOFT-2002-001
- [22] *Athena Developers Guide*,
<http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/architecture/General/Documentation/AthenaDeveloperGuide-8.0.0-draft.pdf>
- [23] G. Barrand *et al.*, *Gaudi - A Software Architecture and Framework for building HEP Data Processing Applications*, Internal Conference on Computing in High Energy Physics (CHEP)(200)
- [24] Application Software Group, *The Zebra System*, CERN Program Library Long Writeup Q100/Q101, Geneva : CERN, 1995.
- [25] Pool Of persistent Objects for LHC - A persistency framework, <http://lcgapp.cern.ch/project/persist>
- [26] Python : langage de script orienté objet, <http://www.python.org>
- [27] T. Sjöstrand, *A brief introduction to Pythia 8.040*, CERN-LCGAPP-2005-05
- [28] S. Gieseke *et al.*, *Herwig++ 1.0 : An Event Generator for e^+e^- Annihilation*, hep-ph/0311208
- [29] ATLAS Collaboration, *ATLAS detector and physics performance Technical Design Report*, CERN/LHCC/99-15 (1999)
- [30] ATLAS Collaboration, *Rome ATLAS Physics Workshop*, <http://agenda.cern.ch/fullAgenda.php?id=a044738> (2005)
- [31] B. Andersson, G. Gustafson, B. Söderberg, *A General Model For Jet Fragmentation*, Z. phys. C 20, page 317, 1983
- [32] R. Brun *et al.*, *GEANT3*, CERN/DD/EE/84-1 (1996)
- [33] The Geant4 Collaboration (S.Agostinelli *et al.*), *Geant4 - A Simulation Toolkit*, Nuclear Instruments and Methods in Physics Research, NIM A 506(2003), 250-303
- [34] *Geant4*, <http://wwwasd.web.cern.ch/wwwasd/geant4/geant4.html>
- [35] E. Richter-Was, D. Froidevaux and L. Poggioli, *ATLFAST 1.0 A package for particle-level analysis*, ATLAS Internal Notes ATL-PHYS-96-079 (1996) and ATL-PHYS-98-131 (1998)
- [36] *Atlfast++ web site*,
<http://www.hep.ucl.ac.uk/atlas/atlfast>
- [37] *LCG Application Project*, <http://lcgapp.cern.ch/>
- [38] Geant4 validation page, <http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/simulation/geant4/index.html>
- [39] C. Collins-Tooth, *Atlfast Comparator*,
<http://ppewww.ph.gla.ac.uk/~chrisc/logs/comparator-overview.html>
- [40] ATLAS Computing Group, *ATLAS Computing Technical Design Report*, ATLAS TDR-017, CERN-LHCC-2005-022

- [41] ATLAS Tracking software Wiki,
<https://uimon.cern.ch/twiki/bin/view/Atlas/TrackingSoftwareOverview>
- [42] ATLAS Navigation classes,
https://uimon.cern.ch/twiki/bin/view/Atlas/DataModelTools#Navigational_Classes
- [43] G. Blazey *et al.*, *Run II Jet Physics*,
hep-ex/0005012v2 (2000)
- [44] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber,
Nucl. Phys. B 406 (1993) 187.
- [45] J. Butterworth *et al.*, *KtJet : A C++ implementation of the Kt clustering algorithm*,
hep-ph/0210022 (2002)
- [46] ATLAS Physics Analysis Tools group,
<http://www.usatlas.bnl.gov/PAT>
- [47] Particle Data Group, <http://pdg.lbl.gov>
- [48] S. Binet, *Filter Tools & Examples*, ATLAS Software Workshop,
<http://agenda.cern.ch/fullAgenda.php?ida=a036309>
- [49] S. Binet, *MC Truth in ESD and AOD*, ATLAS Software & Computing Workshop,
<http://agenda.cern.ch/fullAgenda.php?ida=a045107>
- [50] S. Binet, *Status of AOD MC Truth*, PAT phone conference 09/06/05,
<http://agenda.cern.ch/fullAgenda.php?ida=a055041>
- [51] S.G.I. <http://www.sgi.com/tech/stl/> Scott Meyers,
Effective STL,
Ed. Addison-Wesley
- [52] F. Akesson *et al.*, *Outcome of UCL Workshop on the Architectural Aspects of Physics Analysis in Athena*,
http://www.usatlas.bnl.gov/PAT/ucl_workshop_summary.pdf
- [53] S. Binet, *EventView*, Physics Analysis Tools workshop,
<http://agenda.cern.ch/fullAgenda.php?ida=a051946>
- [54] F. Akesson *et al.*, *Physics Analysis Tools Workshop Summary Report*,
CERN-ATL-COM-SOFT-2005-006, ATL-SOFT-INT-2005-005
- [55] K. Cranmer, *Status of EventView & Recent Developments* <http://agenda.cern.ch/fullAgenda.php?ida=a054691>
- [56] *ESRAT Review*, ATLAS Software & Computing Workshop,
<http://agenda.cern.ch/fullAgenda.php?ida=a045109#s7>
- [57] S. Binet, Page CVS du paquet Pyragorn,
<http://atlas-sw.cern.ch/cgi-bin/viewcvs-atlas.cgi/users/binet/Grid/Pyragorn>
- [58] E. Cogneras, *Rome Top event Reprocessing (Meeting top 08/09/2005)*,
<http://agenda.cern.ch/fullAgenda.php?ida=a054530>
- [59] Collaboration ATLAS, *Results from a new combined test of an Electromagnetic Liquid Argon Calorimeter with a Hadronic Scintillating-Tile Calorimeter*, Nuclear Instruments and Methods in Physics Research A449, page 461 (2000)
- [60] V. Giangiobbe, P. Johansson, K. Jon-And, C. Santoni, *Hadronic calorimeter performance in the ATLAS combined testbeam 2004*, ATL-TILECAL-PUB-2005-008

- [61] F. Gianotti, *Detector performance and physics commissioning with physics data*, <http://gianotti.web.cern.ch/gianotti/commissioning.html>
- [62] F. Derue, A. Kaczmarska, Ph. Schwemling, *Reconstruction of DC1 $J/\psi \rightarrow e^+e^-$ decays and use for the low energy calibration of the ATLAS electromagnetic calorimeter*, ATL-COM-PHYS-2005-022; CERN-ATL-COM-PHYS-2005-022
- [63] F. Djama, *Using $Z^0 \rightarrow e^+e^-$ for Electromagnetic Calorimeter Calibration*, ATL-LARG-2004-008
- [64] ATLAS Test Beam webpage, <http://atlas.web.cern.ch/Atlas/GROUPS/GENERAL/TESTBEAM>
- [65] *First ATLAS Combined Test Beam Workshop*, July 2003, <http://agenda.cern.ch/fullAgenda.php?ida=a031931>
- [66] *Second ATLAS H8 Combined Test Beam Workshop*, November 2003, <http://agenda.cern.ch/fullAgenda.php?ida=a036132>
- [67] Collaboration ATLAS, *ATLAS Detector and Physics Performance Technical Design Report*, CERN/LHCC/99-14, page 143 (1999)
- [68] M. Cobal *et al.*, *Analysis results of the April 1996 combined test of the LArgon and TILECAL barrel calorimeter prototypes*, ATL-TILECAL-98-168 (1998)
- [69] M.P. Casado and M. Cavalli-Sforza, *H1-inspired analysis of the 1994 combined test of the Liquid Argon and Tilecal calorimeter prototypes*, ATL-TILECAL-96-075 (1996)
- [70] C. Biscarat, *Calibration of the hadronic calorimeter using E/p from single hadrons*, ATL-CAL-99-003
- [71] R. Lefèvre, C. Santoni, *In situ determination of jet energy scales and resolutions in ATLAS using Z^0 +jet events* ATL-COM-PHYS-2002-014 (2002)
- [72] R. Lefèvre, *Caractérisation et implantation des photomultiplicateurs du calorimètre à tuiles scintillantes d'ATLAS. Mesure des énergies des jets dans ATLAS*, Thesis / PhD (01 January 2001), Laboratoire de Physique Corpusculaire de Clermont-Ferrand (IN2P3/CNRS), Université Blaise Pascal - Clermont-Ferrand II
- [73] P. Roy, *Perspectives de mesure de la masse du quark Top avec le détecteur ATLAS*, Thesis / PhD (01 January 2002), Laboratoire de Physique Corpusculaire de Clermont-Ferrand (IN2P3/CNRS), Université Blaise Pascal - Clermont-Ferrand II
- [74] P. Roy *et al.*, *Measurement of the top quark mass in the inclusive lepton plus jets channel*, ATL-PHYS-99-026
- [75] I. Borjanović *et al.*, *Investigation of top mass measurements with the ATLAS detector at LHC*, hep-ex/0403021
- [76] T. Sjöstrand, P. Edén, C. Friberg, L. Lnnblad, G. Miu, S. Mrenna and E. Norrbin, *Computer Phys. Commun.* 135 (2001) 238 (LU TP 00-30, hep-ph/0010017)
- [77] Pythia Web Site, <http://www.thep.lu.se/~torbjorn/Pythia.html>
- [78] R. Bonciani *et al.*, *NLL Resummation of the Heavy-Quark Hadroproduction Cross-Section*, *Nucl. Phys B*529 (1998) 424
- [79] C++ Minuit Home page, <http://seal.web.cern.ch/seal/work-packages/mathlibs/minuit/index.html>

-
- [80] P. Avery, *Applied fitting theory VI, Formulas for kinematic fitting* (1998),
<http://www.phys.ufl.edu/~avery/fitting.html>

Résumé

Ce document présente le travail réalisé pour doter l'environnement logiciel de la collaboration ATLAS, ATHENA, d'une bibliothèque d'outils pour l'analyse de physique ainsi que pour l'extraction d'une fonction d'étalonnage de l'énergie des jets à partir d'évènements de physique (étalonnage *in-situ*).

La partie logicielle expose les différents composants de l'architecture logicielle qui a la gestion du flot de données simulées et reconstruites, ainsi que les différentes étapes de ce flot, avant et pendant la prise de données. La construction d'une bibliothèque d'outils facilitant la reconstruction d'objets physiques, leur association avec les entités Monte-Carlo et les interfaces de programmation de ces objets est ensuite détaillée, l'accent étant mis sur l'importance d'avoir un langage et des outils communs à l'ensemble de la collaboration afin de partager l'effort de validation de ces outils et ainsi obtenir des résultats de physique reproductibles.

Dans la partie analyse, l'implémentation dans le *framework* ATHENA d'un algorithme d'étalonnage de l'énergie des jets légers en utilisant la désintégration de bosons W en une paire de jets est traitée. A partir de l'application de cet algorithme sur des données simulées *via* la simulation rapide et la simulation complète, il semble envisageable de connaître l'échelle en énergie des jets légers à hauteur du pour-cent. Enfin, l'étude de faisabilité de l'extraction de l'échelle en énergie des jets de b en utilisant le processus $\gamma Z^0 \rightarrow \gamma b \bar{b}$ est exposée. Il est montré que l'application de coupures séquentielles ne permet pas d'extraire le signal par rapport au bruit de fond. Cependant, une approche multivariable pourrait améliorer la sélection, permettant de collecter un nombre suffisant de paires Z^0 pour réaliser l'étalonnage en énergie des jets de b .

Mots-clés : ATLAS, ATHENA, $Z \rightarrow b \bar{b}$, calibration

Abstract

This thesis presents the work achieved to instrument the ATLAS software framework, ATHENA, with a library of tools and utils for the physics analysis as well as the extraction of the jet energy scale using physics events (*in-situ* calibration).

The software part presents the various components of the ATHENA framework which handles the simulated and reconstructed data flow as well as the different stages of this process, before and during the data taking. The building of a library of tools easing the reconstruction of physics objects, their association with Monte-Carlo particles and their API is then explained. The need for common language and collaboration-wide utils is emphasized as it allows to share the workload of validating these tools and to get reproducible physics results.

The analysis part deals with the implementation of a light jet energy scale calibration algorithm within the C++ framework. This calibration algorithm makes use of W bosons decaying into light jets within semileptonic $t\bar{t}$ events. From the processing of fast and full simulation data with this algorithm, it seems possible to reach a percent level knowledge of the light jet energy scale. Finally, the feasibility study of the b -jet energy scale calibration using $\gamma Z^0 \rightarrow \gamma b \bar{b}$ events is presented. It is shown that a purely sequential approach is not sufficient to extract the signal nor to collect a sufficient amount of Z^0 to calibrate the b -jet energy scale.

Keywords : ATLAS, ATHENA, $Z \rightarrow b \bar{b}$, calibration