



Caractérisation de grands détecteurs Micromegas pour le projet New Small Wheel et recherche de la production de quatre quarks top avec le détecteur ATLAS au LHC

Thibault Chevalérias

► To cite this version:

Thibault Chevalérias. Caractérisation de grands détecteurs Micromegas pour le projet New Small Wheel et recherche de la production de quatre quarks top avec le détecteur ATLAS au LHC. Physique des Hautes Energies - Expérience [hep-ex]. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLS172 . tel-02268557

HAL Id: tel-02268557

<https://theses.hal.science/tel-02268557>

Submitted on 21 Aug 2019

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.

Caractérisation de grands détecteurs Micromegas pour le projet New Small Wheel et recherche de la production de quatre quarks top avec le détecteur ATLAS au LHC

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud
au Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA),
au sein du Département de Physique des Particules (DPhP)
de l'Institut de Recherche sur les lois Fondamentales de l'Univers (IRFU)

École doctorale n°576 Particules Hadrons Énergie et Noyau : Instrumentation,
Image, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat : Physique des particules

Thèse présentée et soutenue à Gif-sur-Yvette, le 3 juillet 2019, par

THIBAUT CHEVALÉRIAS

Composition du Jury :

Laurent Serin Directeur de recherche, LAL Orsay	Président
Daniel Bloch Directeur de recherche, IPHC Strasbourg	Rapporteur
Martine Bosman Professeure de Recherche, IFAE Barcelone	Rapporteuse
Didier Lacour Directeur de recherche, LPNHE Paris	Examineur
Jean-Pierre Meyer Ingénieur chercheur, CEA Saclay	Examineur
Frédéric Déliot Ingénieur chercheur, CEA Saclay	Directeur de thèse

Introduction

Le modèle standard de la physique des particules décrit les composants élémentaires de la matière ainsi que leurs interactions. Il rassemble en une même théorie la description de tous les types de phénomènes physiques excepté la gravitation. L'observation du boson de Higgs par les collaborations ATLAS et CMS en 2012 complète la liste des particules élémentaires prédite par le modèle qui présente toutefois plusieurs limitations, par exemple les problèmes de hiérarchie et de naturalité. Il ne décrit pas tous les phénomènes observés, notamment dans le secteur des neutrinos et des observations réalisées en cosmologie. C'est pourquoi la recherche de phénomènes dits de nouvelle physique est un aspect majeur de la recherche contemporaine en physique des particules. Ce travail de thèse s'inscrit dans la recherche de nouvelle physique en s'intéressant à la production de quatre quarks top au LHC, un phénomène très rare et prometteur.

Le LHC est le collisionneur le plus énergétique du monde avec 13 TeV d'énergie dans son centre de masse suite à son redémarrage en 2015. Des collisions proton-proton ont été analysées par les détecteurs situés le long du faisceau. Le détecteur ATLAS est l'un d'entre eux, il est constitué de plusieurs couches permettant de détecter les particules produites lors des collisions et de mesurer leur propriétés. Les données récoltées par le détecteur ATLAS entre 2015 et 2018 représentent un potentiel d'étude impressionnant, avec une luminosité intégrée d'environ 139 fb^{-1} . Les recherches effectuées dans cette thèse concernent d'abord une recherche du processus de production de quatre quarks top qui a été réalisée avec les données prises en 2015 et en 2016, ainsi qu'une analyse dédiée au processus à quatre tops du modèle standard sur les données prises entre 2015 et 2018 qui est en cours.

L'analyse du processus de production de quatre quarks top représente un défi, d'abord parce que le taux de tels événements est très faible. La stratégie poursuivie dans ce travail est de s'intéresser aux événements comportant au moins deux leptons ayant la même charge électrique. Ce canal ne représente qu'environ 12 % des événements à quatre tops, mais présente un niveau de bruit de fond très faible. Le point clé de l'analyse est la modélisation des bruits de fond, d'un côté la production d'une paire de quarks top accompagnée d'un boson qui doit être modélisée à haut nombre de jets et de jets de quarks b , et de l'autre côté les bruits de fond provenant d'effets instrumentaux appelés

bruits de fond réductibles, notamment les leptons dits non-prompts ou faux. Plusieurs aspects ont été particulièrement étudiés dans cette thèse : l'optimisation de la sélection de l'analyse, la phase d'analyse statistique et de production des résultats, ainsi que le développement d'une nouvelle méthode d'estimation du bruit de fond dû aux faux leptons.

Un arrêt de deux ans du LHC a débuté fin 2018, et plusieurs détecteurs disposés le long du faisceau vont être mis à jour. Les petites roues du spectromètre à muons du détecteur ATLAS doivent notamment être remplacées en vue de la phase haute-luminosité du LHC prévue pour 2026. En effet la haute luminosité qui permettra une augmentation significative du volume de données récoltées va également provoquer une irradiation nettement plus forte qu'à l'heure actuelle, que certaines parties du détecteur ne pourront pas supporter. Les nouvelles petites roues appelées NSW sont constituées de détecteurs gazeux basés sur les technologies sTGC et Micromegas. Le quart des détecteurs Micromegas pour le projet NSW est construit à l'Institut de Recherche sur les lois Fondamentales de l'Univers (Irfu). Il est nécessaire de caractériser ces détecteurs notamment en termes d'efficacité de détection et de gain, c'est pourquoi un banc de test utilisant les muons cosmiques provenant de l'atmosphère a été mis en place. Dans le cadre de cette thèse, un code d'analyse des données du banc de test a été développé afin de réaliser la caractérisation des premiers détecteurs Micromegas produits à l'Irfu. Ce code est destiné à être utilisé tout au long de la production de la centaine de détecteurs Micromegas pour NSW.

Ce manuscrit se découpe en six chapitres. Le chapitre 1 présente de manière concise le modèle standard de la physique des particules et certaines de ses limites. Le LHC ainsi que le fonctionnement du détecteur ATLAS sont présentés dans le chapitre 2. Le chapitre 3 s'intéresse à la caractérisation des détecteurs Micromegas pour le projet NSW, il décrit notamment le banc de test mis en place ainsi que l'analyse des données de ce banc, puis les résultats obtenus sur les premiers détecteurs produits à l'Irfu. Le chapitre 4 présente ensuite le processus de production de quatre quarks top au LHC ainsi que quelques modèles de nouvelle physique associés à ce processus. Ces modèles sont étudiés dans l'analyse de données détaillée dans le chapitre 5. Il s'agit de l'analyse effectuée sur les données prises en 2015 et en 2016. Enfin, le chapitre 6 donne une vue d'ensemble de l'analyse en cours de réalisation, sur les données prises entre 2015 et 2018. Quelques résultats préliminaires sont présentés, notamment ceux d'une nouvelle méthode d'estimation du bruit de fond dû aux faux leptons développé lors de ce travail de thèse.

Table des matières

Introduction	iii
1 Le modèle standard de la physique des particules	1
1.1 Vue d'ensemble	1
1.2 L'électrodynamique quantique (QED)	3
1.3 La théorie électrofaible	3
1.4 Le mécanisme de brisure de symétrie électrofaible	5
1.5 La chromodynamique quantique (QCD)	7
1.6 Bilan et limites du modèle standard	8
1.7 Tests expérimentaux du modèle standard	9
Références	11
2 Le détecteur ATLAS au LHC	13
2.1 Le grand collisionneur de hadrons (LHC)	13
2.2 Le détecteur ATLAS	17
2.2.1 Le système de coordonnées d'ATLAS	19
2.2.2 Le trajectographe interne	20
2.2.3 Les calorimètres	21
2.2.4 Le spectromètre à muons	27
2.2.5 Le détecteur LUCID	29
2.2.6 Le système de déclenchement et d'acquisition	29
2.3 Reconstruction des objets physiques	30
2.3.1 Les électrons	30
2.3.2 Les muons	32
2.3.3 Les jets	35
2.3.4 Les jets de quarks b	36
2.3.5 L'énergie transverse manquante	38
2.3.6 Effet de l'empilement	39
2.3.7 Suppression des chevauchements	39

2.4	Les simulations	39
	Références	43
3	Caractérisation de grands détecteurs Micromegas pour le projet <i>New Small Wheel</i>	47
	<i>Small Wheel</i>	47
3.1	Introduction	47
3.2	Les détecteurs Micromegas pour <i>New Small Wheel</i>	50
3.2.1	La technologie Micromegas	51
3.2.2	Spécificités des détecteurs Micromegas pour le projet NSW	52
3.2.3	Les modules LM1	53
3.3	Le banc cosmique de test des modules	55
3.3.1	Mise en place	55
3.3.2	L'électronique de lecture et le système de déclenchement	57
3.3.3	Le trajectographe externe « Mcube »	58
3.4	Analyse des données du banc cosmique	59
3.4.1	Pré-traitement des données	59
3.4.2	Sélection des événements	61
3.4.3	Caractérisation de la qualité des modules	66
3.4.4	Critères requis pour le contrôle qualité des modules	74
3.5	Caractérisation du Module 1	75
3.5.1	Caractérisation du Module 1 avant nettoyage	75
3.5.2	Caractérisation du doublet « ETA » après nettoyage	77
3.6	Caractérisation du Module 2	82
3.6.1	Influence de la tension d'amplification	83
3.6.2	Influence de la tension de dérive	85
3.6.3	Étude de l'influence de la référence externe Mcube	86
3.6.4	Comparaison avec une méthode de reconstruction de traces interne	89
3.7	Conclusion et perspectives	92
	Références	94
4	Le processus de production de quatre quarks top	97
4.1	Production de quatre quarks top au LHC	98
4.1.1	Section efficace théorique	98
4.1.2	Signatures expérimentales	101
4.2	Production de quatre quarks top dans les modèles de nouvelle physique	102
4.2.1	Modèle effectif d'interaction de contact entre quatre quarks top	103
4.2.2	Modèle de dimensions supplémentaires	105
4.2.3	Modèle avec deux doublets de Higgs	106
	Références	108

5	Recherche de production de quatre quarks tops avec le détecteur ATLAS au LHC avec $36,1 \text{ fb}^{-1}$	113
5.1	Introduction	113
5.2	Sélection des objets utilisés dans l'analyse	114
5.3	Sélection des événements	116
5.3.1	Pré-sélection	117
5.3.2	Méthode d'optimisation des régions de signal	118
5.3.3	Définition des régions de signal	120
5.3.4	Définition des régions de validation	121
5.4	Estimation des bruits de fond et des signaux	123
5.4.1	Bruits de fond irréductibles	123
5.4.2	Bruits de fond réductibles	124
5.4.3	Accord entre simulation et données dans les régions de validation	131
5.5	Incertitudes systématiques	132
5.5.1	Reconstruction des objets	135
5.5.2	Luminosité	135
5.5.3	Bruits de fond irréductibles	135
5.5.4	Bruits de fond réductibles	142
5.5.5	Bilan	143
5.6	Analyse statistique	144
5.7	Résultats	147
5.7.1	Nombre d'événements dans les régions de signal	148
5.7.2	Ajustement du modèle	150
5.7.3	Résultats sur la production de quatre quarks top selon le modèle standard	153
5.7.4	Résultats sur les modèles de nouvelle physique	155
5.7.5	Résultats combinés avec l'analyse à un seul lepton	157
5.8	Conclusion	159
	Références	161
6	Vers la mise en évidence de la production modèle standard de quatre quarks tops avec le détecteur ATLAS	165
6.1	Vue d'ensemble de l'analyse	166
6.1.1	Sélection des objets physiques	166
6.1.2	Pré-sélection des événements	167
6.1.3	Définition des régions	167
6.1.4	Stratégie d'extraction du signal	169
6.1.5	Méthodes d'estimation des bruits de fond	170
6.2	Méthode des templates	172
6.2.1	Composition du bruit de fond dû aux faux leptons	172

6.2.2	Définition des régions de contrôle	174
6.2.3	Définition des régions de validation	178
6.3	Résultats préliminaires	178
6.3.1	Incertitudes systématiques et premiers résultats	178
6.3.2	Résultats avec la méthode des templates	180
6.3.3	Perspectives	186
6.4	Conclusion	190
	Références	192
	Conclusion	195

Table des figures

1.1	Propriétés des particules élémentaires dans le modèle standard	2
1.2	Représentation du potentiel de Higgs en fonction des parties réelle et imaginaire d'un champ complexe	6
1.3	Comparaison de l'ajustement global des paramètres de la théorie électrofaible à partir des résultats obtenus auprès de différents collisionneurs, et les mesures individuelles.	10
2.1	Schéma du LHC montrant les points de croisement des faisceaux le long de l'anneau ainsi que les expériences associées.	14
2.2	Complexe d'accélérateurs du Cern. La chaîne d'accélération permettant l'injection des protons dans le LHC y est représentée.	15
2.3	Luminosité intégrée enregistrée par le détecteur ATLAS durant les différentes prises de données effectuées entre 2011 et 2018 (gauche), et durant le Run 2 (droite).	16
2.4	Organisation dans le temps des différentes prises de données au LHC, avec les prévisions pour le projet HL-LHC.	17
2.5	Valeurs d'empilement moyen μ dans le détecteur ATLAS pour les prises de données du Run 2.	17
2.6	Vue d'ensemble du détecteur ATLAS et de ses principaux composants.	18
2.7	Vue d'ensemble du trajectographe interne et de ses différents composants (gauche). Organisation et dimensions des sous-parties du détecteur interne dans le tonneau (droite).	20
2.8	Résolution du trajectographe interne sur les paramètres d'impact transverse d_0 et longitudinal z_0 des traces en fonction de leur impulsion transverse.	22
2.9	Vue d'ensemble du système de calorimétrie du détecteur ATLAS.	23
2.10	Schéma en coupe du calorimètre électromagnétique d'ATLAS.	24
2.11	Résolution sur l'énergie des électrons en fonction de leur énergie transverse et de $ \eta $	25
2.12	Schéma d'un module du calorimètre hadronique à tuiles scintillantes.	26

2.13	Résolution sur l'impulsion transverse des jets dans les données prises en 2017 pour des jets dans la région $0,2 \leq \eta < 0,7$	26
2.14	Disposition des composants du spectromètre à muons en coupe transverse (gauche) et longitudinale (droite).	27
2.15	Résolution relative sur la masse reconstruite de la paire de muons mesurée à partir des résonances $J/\psi \rightarrow \mu^+\mu^-$ et $Z \rightarrow \mu^+\mu^-$ dans les données prises en 2015.	28
2.16	Efficacité de sélection des électrons en fonction de leur énergie transverse et des différents critères d'identification.	32
2.17	Efficacité de sélection des muons <i>medium</i> en fonction de leur impulsion transverse.	34
2.18	Comparaison de l'échelle d'énergie des jets mesurées sur les données prises entre 2015 et 2017 et sur des simulations Monte-Carlo, en fonction de l'impulsion transverse des jets.	36
2.19	Schéma représentant une collision proton-proton simulée par un générateur Monte-Carlo.	41
3.1	Vue en coupe des détecteurs à muons au sein du détecteur ATLAS.	48
3.2	(a) Efficacité de reconstruction de segments de traces en fonction de leur position radiale. (b) Taux de reconstruction de « faux » segments, c'est-à-dire n'étant pas associés à un muon réel, en fonction de leur position radiale.	49
3.3	Disposition des détecteurs NSW sur une roue.	50
3.4	(a) Vue d'un secteur NSW. (b) Vue des deux quadruplets MM d'un secteur NSW.	51
3.5	Schéma du concept d'un détecteur Micromegas.	53
3.6	(a) Représentation des quatre couches d'un quadruplet MM LM1. (b) Représentation d'une d'un quadruplet MM LM1 couche vue de dessus.	54
3.7	Photo du banc cosmique mis en place à l'Irfu.	56
3.8	Photo d'une carte d'adaptation avec ses huit connecteurs.	58
3.9	(a) Amplitudes en l'absence de signal avant correction de piédestal et suppression du bruit cohérent. (b) Amplitudes en l'absence de signal après correction de piédestal et suppression du bruit cohérent.	62
3.10	(a) Amplitudes en présence d'un signal avant correction de piédestal et suppression du bruit cohérent. (b) Amplitudes en présence d'un signal après correction de piédestal et suppression du bruit cohérent.	62
3.11	Distribution du temps calculé des amas reconstruits dans un circuit imprimé.	65
3.12	Distribution des amas reconstruits.	66
3.13	Distribution des positions des traces provenant de Mcube dans le plan d'un détecteur.	67
3.14	Distribution des amplitudes des amas reconstruits sur un circuit imprimé, avec ajustement d'une fonction de Landau.	69

3.15 Évolution du gain d'une chambre Micromegas en fonction du champ d'amplification, pour différents mélanges de gaz.	70
3.16 Distribution des résidus entre les traces de référence et les pistes des amas sélectionnés.	72
3.17 Graphe d'efficacité en fonction du numéro de piste.	73
3.18 Représentation des critères de détection des pistes bruyantes (gauche) et déconnectées (droite).	74
3.19 Courbes d'efficacité des différentes couches du Module 1 avant nettoyage, en fonction de la tension d'amplification.	76
3.20 Courbes d'efficacité des différentes couches du Module 1 avant nettoyage, en fonction de la tension de dérive.	77
3.21 Comparaison du gain obtenu avec le doublet ETA du Module 1 en utilisant différents mélanges de gaz, en fonction de la tension d'amplification.	79
3.22 Comparaison de l'efficacité obtenue avec le doublet ETA du Module 1 après nettoyage en utilisant différents mélanges de gaz, en fonction de la tension d'amplification.	80
3.23 Comparaison de l'efficacité obtenue avec le doublet ETA du Module 1 après nettoyage en utilisant différents mélanges de gaz, en fonction de la tension d'amplification (zones des haute tension trop basse ignorées).	81
3.24 Efficacité en fonction du gain pour les couches du doublet « ETA » du Module 1 après nettoyage (zones des hautes tensions trop basses ignorées).	82
3.25 Courbes de gain des différentes couches du Module 2, en fonction de la tension d'amplification.	83
3.26 Courbes d'efficacité des différentes couches du Module 2, en fonction de la tension d'amplification.	84
3.27 Courbes de gain des différentes couches du Module 2, en fonction de la tension de dérive.	85
3.28 Courbes d'efficacité des différentes couches du Module 2, en fonction de la tension de dérive.	86
3.29 Efficacité des couches du Module 2 en fonction de la valeur maximale de la variable χ_X^2	87
3.30 Efficacité des couches du Module 2 en fonction de la valeur maximale de la variable χ_Y^2	88
3.31 Efficacité des couches du Module 2 en fonction de la coupure sur l'angle des traces par rapport à la verticale α_z	89
3.32 Courbes d'efficacité externe (trait plein) et interne (trait pointillé) de différentes couches du Module 2, en fonction de la tension d'amplification.	90
3.33 Courbes d'efficacité externe (trait plein) et interne (trait pointillé) de différentes couches du Module 2, en fonction de la tension de dérive.	91

3.34	Efficacité des quatre couches du Module 4 en deux dimensions.	93
4.1	Diagrammes de Feynman représentatifs pour l'amplitude de Born $pp \rightarrow t\bar{t}\bar{t}\bar{t}$. .	98
4.2	Diagrammes de Feynman représentatifs pour l'amplitude $pp \rightarrow t\bar{t}\bar{t}\bar{t}$ à une boucle.	98
4.3	Diagrammes de Feynman représentatifs pour l'amplitude $pp \rightarrow t\bar{t}\bar{t}\bar{t}$ avec une émission réelle de parton.	99
4.4	Diagrammes de Feynman représentatifs pour l'amplitude $pp \rightarrow t\bar{t}\bar{t}\bar{t}$ avec des vertex électrofaibles.	99
4.5	États finaux et leurs rapports d'embranchement pour la désintégration de quatre bosons W	102
4.6	Diagramme de Feynman de production de quatre quarks top via interaction de contact.	104
4.7	Diagramme de Feynman de production de quatre quarks top dans le modèle à deux dimensions supplémentaires compactifiées.	106
4.8	Diagramme de Feynman de production de quatre quarks top dans le modèle à deux doublets de Higgs.	107
5.1	Distributions de grandeurs cinématiques pour l'estimation du bruit de fond et de plusieurs signaux représentatifs.	119
5.2	Taux d'électrons lâches passant les coupures strictes, en fonction des valeurs de (p_T, η) , et $\Delta R(e, \text{jet})$	128
5.3	Taux de muons lâches passant les coupures strictes, en fonction des valeurs de $p_T, \eta $, et $\Delta R(\mu, \text{jet})$	129
5.4	Taux ε de mauvaise identification de la charge d'un électron par intervalle (η , p_T)	132
5.5	Distributions de la grandeur E_T^{miss} dans les régions de validation.	133
5.6	Distributions de la grandeur H_T dans les régions de validation.	134
5.7	Nombres d'événements $t\bar{t}W$ dans les régions de signal estimés avec différentes valeurs des échelles de factorisation et de renormalisation.	137
5.8	Nombres d'événements $t\bar{t}Z$ dans les régions de signal estimés avec différentes valeurs des échelles de factorisation et de renormalisation.	138
5.9	Nombres d'événements $t\bar{t}W$ dans les régions de signal estimés via deux géné- rateurs Monte-Carlo alternatifs.	139
5.10	Nombres d'événements $t\bar{t}Z$ dans les régions de signal estimés via deux généra- teurs Monte-Carlo alternatifs.	141
5.11	Nombre d'événements de bruit de fond et de données dans les huit régions de signal.	148
5.12	Distributions de variables cinématiques des événements entrant dans la région de signal SR3b2 ℓ _L ou dans la région de signal SR3b3 ℓ _L.	151

5.13	Distributions de variables cinématiques des événements entrant dans n'importe quelle région de signal.	152
5.14	Effet des paramètres de nuisance les plus influents sur la force du signal de production de quatre quarks tops dans le cadre du modèle standard.	154
5.15	Contraintes sur le modèle à interaction de contact dans le plan (Λ, C_{4t}) . . .	156
5.16	Limite attendue et observée sur la section efficace (gauche) et dans le plan (m_{KK}, ξ) (droite), dans le cadre du modèle à deux dimensions supplémentaires compactifiées.	156
5.17	Limites sur le modèle à deux doublets de Higgs.	157
5.18	Combinaison des résultats obtenus avec ceux de l'analyse dans les canaux à un seul lepton et à deux leptons de même charge électrique. Combinaison des limites (gauche) et des meilleurs ajustements (droite).	159
6.1	Taux de mauvaise identification de la charge des électrons en fonction de $ \eta $ pour quatre intervalles de p_T	171
6.2	Nombre d'événements du bruit de fond simulé dans les régions de signal du canal à deux leptons de même charge.	175
6.3	Nombre d'événements du bruit de fond simulé dans la région de signal SR4b. . .	176
6.4	Composition du bruit de fond des différentes régions de contrôle, de validation, et de signal.	179
6.5	Valeurs ajustées des facteurs de normalisation de la méthode des templates. .	181
6.6	Matrice de corrélation des paramètres de nuisance et des facteurs de normalisation.	182
6.7	Distributions des variables d'ajustement dans les régions de contrôle CR $t\bar{t}W$ (haut) et CR $t\bar{t}CO$ (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.	183
6.8	Distributions des variables d'ajustement dans les régions de contrôle CR1b3 ℓ (haut) et CR $t\bar{t}Z$ (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.	184
6.9	Distributions de la grandeur H_T dans les régions de validation VR1b2 ℓ (haut) et VR1b3 ℓ (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.	185
6.10	Distribution de la grandeur H_T dans la région de validation VR2b3 ℓ , avant (gauche) et après (droite) l'ajustement du modèle statistique.	186
6.11	Nombres d'événements dans les régions de contrôle (haut) et dans les régions de validation (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.	187
6.12	Effet des paramètres de nuisance les plus influents sur la force du signal de production de quatre quarks tops dans le cadre du modèle standard.	188

6.13 Comparaison de la modélisation au NLO du signal $t\bar{t}t\bar{t}$ avec les générateurs MG5_AMC@NLO+PYTHIA et MG5_AMC@NLO+HERWIG.	189
---	-----

Liste des tableaux

2.1	Efficacité d'identification des jets de quarks b et facteurs de rejet des jets de quarks c et des jets légers pour les quatre points de fonctionnement de l'algorithme MV2c10.	38
3.1	Résumé des coupures et autres actions appliquées aux différents objets lors de l'analyse des données.	60
5.1	Résumé de la sélection des objets utilisés dans l'analyse.	115
5.2	Définition des régions de signal de l'analyse.	121
5.3	Efficacités de sélection dans les régions de signal pour différents modèles de production de quatre quarks top.	122
5.4	Définition des régions de validation de l'analyse.	122
5.5	Définition des régions de contrôle utilisées pour mesurer les taux des vrais et faux leptons.	127
5.6	Incertitudes systématiques globales appliquées aux processus $t\bar{t}W$ et $t\bar{t}Z$ dans les différentes régions de signal.	140
5.7	Variations des nombres d'événements VV par rapport à la valeur nominale, pour les différentes variations d'échelles étudiées.	140
5.8	Incertitudes systématiques exprimées en pourcentage du nombre total d'événements de bruit de fond dans les régions de signal, par type d'incertitude.	144
5.9	Nombre d'événements de bruit de fond et de données dans les huit régions de signal.	149
5.10	Valeurs ajustées de force du signal pour plusieurs modèles représentatifs.	153
5.11	Limites supérieures attendues et observées sur la production de quatre quarks top dans le modèle standard et le modèle à interaction de contact.	155
6.1	Résumé des catégories de l'analyse et de la pré-sélection des événements.	168
6.2	Définition des régions de signal du canal à deux leptons de même charge.	169
6.3	Définition des régions de contrôle pour la méthode des templates.	177
6.4	Définition des régions de validation pour la méthode des templates.	178

Acronymes

- 2HDM** *Two-Higgs Doublet Model*. 97, 103, 106, 107, 113, 122, 153, 155–157
- 2UED** *Two Universal Extra Dimensions*. 97, 103, 105, 106, 113, 118, 122, 153, 155–157
- ALICE** *A Large Ion Collider Experiment*. 13
- ATLAS** *A Toroidal LHC ApparatuS*. iii, iv, 6, 10, 13, 14, 16–20, 23, 28–30, 34, 35, 37, 38, 40–42, 47, 97, 114, 159, 160, 165, 170, 191, 195
- BSM** *Beyond Standard Model*. 9, 97, 102, 113, 114, 117, 118, 120, 121, 123, 136, 142, 144, 148–150, 153, 159
- Cern** Conseil Européen pour la Recherche Nucléaire. 13, 15, 50
- CI** *Contact Interaction*. 97, 103, 104, 113, 122, 153, 155, 156, 159
- CMS** *Compact Muon Solenoid*. iii, 6, 13, 97, 165, 191, 196
- CSC** *Cathode-Strip Chamber*. 27, 33, 47, 52
- DPhP** Département de Physique des Particules. 92
- EFT** *Effective Field Theory*. 103, 191
- HL-LHC** *High Luminosity LHC*. 16–18, 28, 47, 81, 191
- Irfu** Institut de Recherche sur les lois Fondamentales de l’Univers. iv, 29, 50, 51, 53, 55–57, 74, 75, 92, 195
- KK** Kaluza-Klein. xiii, 105, 106, 155, 156
- LHC** *Large Hadron Collider*. iii, iv, 6, 10, 13–17, 19, 40, 97, 99, 102, 103, 117, 159, 165, 167, 190, 191
- LHCb** *LHC beauty*. 13
- LO** *leading-order*. 40, 99, 100, 124

- MC** Monte-Carlo. xii, 22, 25, 31, 32, 34, 36, 39–42, 118, 123, 124, 127, 131, 135–142, 160, 170–174, 180, 186, 188, 191
- MDT** *Monitored Drift Tube*. 27, 33, 52
- MS** modèle standard de la physique des particules. iii, iv, 1, 8, 9, 104, 114, 122, 148–150, 153, 155
- NLO** *next-to-leading-order*. 40, 100, 116, 124, 165, 186, 189
- NSW** *New Small Wheel*. iv, 47–52, 57, 59, 68, 74, 92, 195
- PDF** fonctions de distribution de partons. 40, 100, 165, 186
- PS** *Proton Synchrotron*. 13
- PSB** *Proton Synchrotron Booster*. 13
- QCD** *Quantum Chromo-Dynamics*. 7–9, 40, 99, 100, 116, 124, 165
- QED** *Quantum Electro-Dynamics*. 3–5, 40
- RPC** *Resistive Plate Chamber*. 27
- SPS** *Super Proton Synchrotron*. 13
- TGC** *Thin Gap Chamber*. 27
- VLQ** *Vector-Like Quarks*. 113, 118

1 Le modèle standard de la physique des particules

1.1 Vue d'ensemble

Le modèle standard de la physique des particules (MS) est une théorie quantique des champs relativistes décrivant les constituants élémentaires de la matière, appelés particules élémentaires, ainsi que leurs interactions. Il repose sur des invariances sous des transformations locales des champs, associées à des groupes de symétrie dont découlent les interactions fondamentales entre les particules. Le modèle standard est ainsi formulé en termes de groupes de symétrie et utilise le théorème de Noether [1] reliant les symétries aux lois de conservation. Toute transformation laissant l'action invariante est ainsi associée à une grandeur qui se conserve. Par exemple, la physique reste inchangée si l'on modifie la phase complexe d'un champ. Cette invariance correspond au groupe de jauge $U(1)$ et est associée à la conservation de la charge électrique Q . On note alors ce groupe $U(1)_Q$, qui représente l'interaction électromagnétique. De manière similaire, on associe l'interaction faible à la conservation de l'isospin faible et au groupe de jauge $SU(2)_L$, ainsi que l'interaction forte à la conservation de la charge de couleur et au groupe de jauge $SU(3)_c$. Dans le modèle standard, les interactions électromagnétique et faible sont rassemblées dans une même interaction électrofaible décrite dans la section 1.3, dont la symétrie est brisée selon le mécanisme détaillé dans la section 1.4.

Le MS décrit la matière comme étant constituée de fermions de spin $1/2$: six leptons et six quarks. Il existe trois leptons chargés qui sont l'électron e^- , le muon μ^- , et le tau τ^- , ainsi que les trois leptons neutres associés appelés neutrinos ν_e , ν_μ , ν_τ . Les quarks sont également répartis en trois familles contenant chacune un couple de charge électrique positive et négative. Les quarks u (up) et d (down) sont les plus légers, suivis par les quarks s (strange) et c (charm), puis les quarks b (bottom) et t (top). Chacun de ces fermions possède une anti-particule lui correspondant, doublant le nombre total de particules. Les caractéristiques de ces particules sont résumées sur la figure 1.1. Il existe une très grande différence de masse entre les particules du modèle standard, par exemple dans le secteur des leptons on trouve l'électron ayant une masse d'environ 511 keV ainsi que le τ qui a une masse d'environ 1,78 GeV.

Les particules élémentaires interagissent via quatre forces fondamentales. Les particules

	masse → charge → spin →	$\approx 2,3 \text{ MeV}/c^2$ $2/3$ $1/2$	$\approx 1,275 \text{ GeV}/c^2$ $2/3$ $1/2$	$\approx 173,07 \text{ GeV}/c^2$ $2/3$ $1/2$	0 0 1	$\approx 126 \text{ GeV}/c^2$ 0 0
		u up	c charm	t top	g gluon	H boson de Higgs
QUARKS		$\approx 4,8 \text{ MeV}/c^2$ $-1/3$ $1/2$	$\approx 95 \text{ MeV}/c^2$ $-1/3$ $1/2$	$\approx 4,18 \text{ GeV}/c^2$ $-1/3$ $1/2$	0 0 1	
		d down	s strange	b bottom	γ photon	
		$0,511 \text{ MeV}/c^2$ -1 $1/2$	$105,7 \text{ MeV}/c^2$ -1 $1/2$	$1,777 \text{ GeV}/c^2$ -1 $1/2$	$91,2 \text{ GeV}/c^2$ 0 1	
		e électron	μ muon	τ tau	Z^0 boson Z^0	
LEPTONS		$< 2,2 \text{ eV}/c^2$ 0 $1/2$	$< 0,17 \text{ MeV}/c^2$ 0 $1/2$	$< 15,5 \text{ MeV}/c^2$ 0 $1/2$	$80,4 \text{ GeV}/c^2$ ± 1 1	
		ν_e neutrino électronique	ν_μ neutrino muonique	ν_τ neutrino tauique	$W^\pm$ boson $W^\pm$	
					BOSONS DE JAUGE	

Figure 1.1 – Propriétés des particules élémentaires dans le modèle standard [2].

médiatrices des différentes forces fondamentales sont des bosons, appelées bosons de jauge. Il s'agit de particules de spin 1. Le photon γ est le médiateur de l'interaction électromagnétique, les huit gluons ceux de l'interaction nucléaire forte, et les bosons W^+ , W^- et Z^0 sont les médiateurs de l'interaction nucléaire faible. À cette liste s'ajoute le boson de Higgs H de spin 0 qui n'est pas un boson médiateur d'une force mais qui est responsable du mécanisme qui donne leur masse aux bosons de jauge, détaillé dans la section 1.4. Les propriétés de ces bosons sont résumées sur la figure 1.1. La quatrième force fondamentale est la gravité qui est absente du modèle standard, et est d'intensité très inférieure aux autres forces.

Ce chapitre décrit d'abord les interactions fondamentales : l'interaction électromagnétique dans la section 1.2, puis l'interaction électrofaible qui résulte de l'unification des interactions électromagnétique et faible dans la section 1.3. Cette dernière est brisée via le mécanisme de Higgs, à l'origine de la masse des fermions et des bosons $W^\pm$ et Z^0 , qui est présenté dans la section 1.4. L'interaction forte est ensuite décrite dans la section 1.5. Finalement, les sections 1.6 et 1.7 présentent respectivement quelques limites du modèle standard puis une introduction aux tests expérimentaux réalisés auprès des collisionneurs.

1.2 L'électrodynamique quantique (QED)

L'interaction électromagnétique agit sur toutes les particules possédant une charge électrique, à savoir tous les fermions chargés ainsi que les bosons $W^\pm$. Son médiateur est le photon qui est sans masse et neutre électriquement. La formulation en termes de théorie quantique des champs de l'électromagnétisme est appelé électrodynamique quantique, ou *Quantum Electro-Dynamics* (QED). Les interactions entre les bosons $W^\pm$ et les photons ne sont pas décrits par la QED, mais par la théorie électrofaible présentée dans la section 1.3. Au sein du modèle standard, la QED est associée au groupe de symétrie $U(1)_Q$ et à la conservation de la charge électrique Q . Elle est décrite par le lagrangien suivant :

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.1)$$

où ψ et $\bar{\psi}$ représentent les champs des fermions chargés, les γ^μ sont les matrices de Dirac, $D_\mu = \partial_\mu + ieA_\mu$ est la dérivée covariante avec e la charge élémentaire, et $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ est le tenseur électromagnétique. Le champ A_μ représente le potentiel électromagnétique et donc le champ de photons. La dérivée covariante permet d'assurer l'invariance de jauge de la théorie QED, qui est l'invariance vis-à-vis de la modification de la phase complexe d'un champ, associée à la conservation de la charge électrique. Comme la théorie classique de l'électromagnétisme, la QED doit être invariante par transformation de jauge du champ électromagnétique, c'est-à-dire à un gradient de champ scalaire près : $A_\mu \rightarrow A_\mu - \frac{1}{e}\partial_\mu\alpha$. Ceci implique une redéfinition de la phase des champs de fermions $\psi \rightarrow e^{i\alpha}\psi$, et de la dérivée ∂_μ qui devient la dérivée covariante $D_\mu = \partial_\mu + ieA_\mu$. Ainsi la transformation de jauge laisse le lagrangien $\mathcal{L}_{\text{QED}}$ invariant.

Le premier terme du lagrangien représente la propagation des fermions chargés via le terme $\bar{\psi}\partial_\mu\psi$ et l'interaction électromagnétique avec le champ de photons A_μ via le terme ieA_μ . Le terme $\bar{\psi}m\psi$ correspond à la masse des fermions, et finalement le terme purement de jauge $F_{\mu\nu}F^{\mu\nu}$ représente la propagation des photons. Il n'y a pas de terme de masse pour les photons qui sont donc des particules de masse nulle, l'interaction électromagnétique est ainsi de portée infinie.

1.3 La théorie électrofaible

L'interaction nucléaire faible agit sur toutes les particules du modèle standard, via les bosons $W^\pm$ et Z^0 . Le groupe de symétrie de cette interaction est $SU(2)_L$ et la charge associée est la troisième composante de l'isospin faible T , notée T_3 . Cette interaction est la seule qui permet de modifier la nature des particules. On peut par exemple décrire une désintégration β^- ($n \rightarrow p^+ + e^- + \bar{\nu}_e$) comme la transformation d'un quark d en quark u via l'émission d'un boson W^- , puis la désintégration de ce boson en e^- et $\bar{\nu}_e$.

L'interaction faible agit différemment selon la chiralité des particules, gauche ou droite. On décompose donc les champs fermioniques ainsi :

$$\psi = \frac{1 - \gamma_5}{2} \psi + \frac{1 + \gamma_5}{2} \psi = \psi_L + \psi_R, \quad (1.2)$$

où γ_5 est une matrice de Dirac, et ψ_L et ψ_R sont respectivement les parties gauche et droite du fermion ψ . Les particules élémentaires sont alors décrites sous forme de doublets d'isospin faible pour leur composante gauche, par exemple pour les fermions de la première génération :

$$\psi_{eL} = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \psi_{qL} = \begin{pmatrix} u_L \\ d_L \end{pmatrix}. \quad (1.3)$$

Ces doublets ont un isospin $T = \frac{1}{2}$, avec une troisième composante $T_3 = +\frac{1}{2}$ pour ν_e et u_L , et $T_3 = -\frac{1}{2}$ pour e_L et d_L . Les composantes droites des particules sont des singulets (e_R, u_R, d_R) et ont un isospin nul $T = 0$. Le modèle standard n'inclut pas de partie droite pour les neutrinos, en effet seuls des neutrinos gauches ont été observés à ce jour.

À partir d'une énergie de l'ordre de 100 GeV, les interactions électromagnétique et faible sont unifiées au sein d'une même interaction dite électrofaible, qui inclut la théorie QED [3, 4, 5]. Cette nouvelle interaction est fondée sur la conservation d'une grandeur appelée hypercharge Y . Elle est exprimée en fonction de la charge électrique et la troisième composante de l'isospin via la formule de Gell-Mann et Nishijima :

$$Y = 2(Q - T_3). \quad (1.4)$$

Ainsi l'interaction électrofaible est décrite par le groupe $SU(2)_L \otimes U(1)_Y$. Le groupe de symétrie de l'interaction faible $SU(2)_L$ fait apparaître trois générateurs W^1, W^2 et W^3 , et le nouveau groupe $U(1)_Y$ en introduit un quatrième B . Le lagrangien de l'interaction électrofaible est le suivant :

$$\mathcal{L}_{EW} = i\bar{\psi}_L \gamma^\mu D_{\mu,L} \psi_L + i\bar{\psi}_R \gamma^\mu D_{\mu,R} \psi_R - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \quad (1.5)$$

où ψ_L et $\bar{\psi}_L$ représentent les champs des fermions gauches existant sous forme de doublets comme expliqué précédemment, et de même ψ_R et $\bar{\psi}_R$ sont champs des fermions droits se présentant sous forme de singulets. $D_{\mu,L} = \partial_\mu + ig'YB_\mu/2 + ig\sigma_a W_\mu^a/2$ et $D_{\mu,R} = \partial_\mu + ig'YB_\mu/2$ sont les dérivées covariantes associées aux fermions gauches et droits respectivement, avec g' la constante de couplage électromagnétique, g la constante de couplage faible, et les σ_a sont les trois matrices de Pauli. Comme dans le cas de la théorie QED, on a $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ pour le champ électromagnétique. Pour les champs des bosons faibles, on a $W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g\epsilon^{abc}W_\mu^b W_\nu^c$, avec a, b et c les indices des trois champs W et ϵ^{abc} le tenseur antisymétrique. Le dernier terme provient du caractère non-commutatif des matrices de Pauli associées au groupe de symétrie $SU(2)_L$.

Les deux premiers termes de l'équation 1.5 représentent la propagation libre des fermions ainsi que leur interaction avec les bosons électrofaibles. Dans le cas des fermions droits, l'interaction est limitée au champ électromagnétique. Les deux derniers termes quant à eux décrivent la propagation et les interactions des quatre champs électrofaibles B et W^a . L'existence du terme $g\epsilon^{abc}W_\mu^b W_\nu^c$ permet des interactions entre les bosons W^a , une spécificité de l'interaction faible qui n'existe pas dans la théorie QED. Les champs physiques correspondant aux bosons qui interagissent avec les fermions du modèle standard sont obtenus à partir des champs B et W^a . Les bosons chargés W^+ et W^- sont une combinaison des champs W^1 et W^2 , et sont anti-particules l'un de l'autre : $W^\pm = (W^1 \mp iW^2)/\sqrt{2}$. Les bosons Z^0 et γ (photon) sont une combinaison des champs W^3 et B correspondant à une rotation dans l'espace (W^3, B) . L'angle de cette rotation est appelé angle de Weinberg θ_W :

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix} \quad (1.6)$$

avec $\cos \theta_W = g/\sqrt{g^2 + g'^2} = m_W/m_Z$. On retrouve la charge élémentaire $e = g \sin \theta_W$.

Aucun terme de masse n'est présent dans l'équation 1.5 puisqu'un terme du type $\bar{\psi}m\psi$ fait apparaître les termes $\bar{\psi}_L\psi_R$ et $\bar{\psi}_R\psi_L$ qui ne respectent pas la symétrie $SU(2)$. Il n'est donc pas possible à ce stade de faire apparaître un terme de masse pour les fermions, ce qui est un problème. De plus, les bosons $W^\pm$ et Z^0 sont observés expérimentalement avec des masses non-nulles : $m_W \simeq 80 \text{ GeV}$ et $m_Z \simeq 91 \text{ GeV}$ [6]. Ces problèmes sont résolus via une brisure spontanée de la symétrie électrofaible $SU(2)_L \otimes U(1)_Y$ décrite dans la section 1.4.

1.4 Le mécanisme de brisure de symétrie électrofaible

Le problème des masses dans la théorie électrofaible est résolu par un mécanisme de brisure spontanée de symétrie [7, 8, 9], qui permet de décrire à la fois l'origine des masses des quarks et des leptons ainsi que celle des bosons de jauge. Un boson scalaire Φ appelé boson de Higgs est introduit en tant que doublet complexe de $SU(2)_L$, associé à un potentiel

$$V(\Phi) = \mu^2|\Phi|^2 + \lambda|\Phi|^4 \quad (1.7)$$

avec $\lambda > 0$ appelé constante d'auto-couplage du boson de Higgs. Ce potentiel a pour minimum $\Phi = 0$ si $\mu^2 \geq 0$, toutefois si $\mu^2 < 0$ ce n'est plus le cas. La figure 1.2 représente le potentiel $V(\phi)$ pour un champ complexe ϕ dans le cas $\mu^2 < 0$, en fonction de sa partie réelle et de sa partie imaginaire. Ce potentiel est symétrique mais le point central est instable et il existe une infinité de configurations $|\phi| = \sqrt{-\mu^2/(2\lambda)}$ minimisant le potentiel. La symétrie est brisée quand le champ ϕ prend une valeur correspondant à l'un

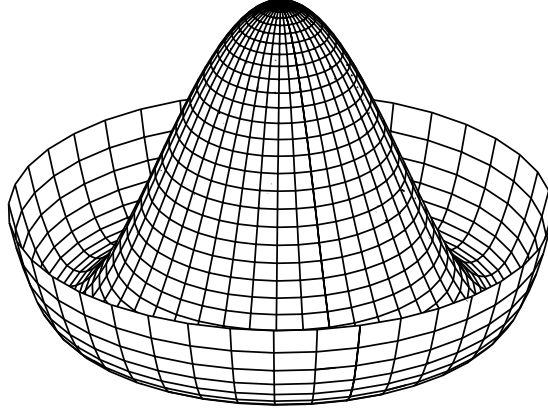


Figure 1.2 – Représentation du potentiel de Higgs en fonction des parties réelle et imaginaire d'un champ complexe.

de ces minima. Dans le cas du champ de Higgs, on dit que le boson de Higgs acquiert une valeur moyenne dans le vide v non nulle, et le doublet correspondant devient

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \quad (1.8)$$

où h représente le champ de Higgs fluctuant autour de sa valeur moyenne $v = \sqrt{-\mu^2/\lambda} = 246 \text{ GeV}$. Le boson de Higgs a été récemment mis en évidence par les expériences ATLAS et CMS au LHC [10, 11].

Après brisure de symétrie le champ de Higgs s'exprime selon l'équation 1.8, et l'équation 1.7 devient

$$V(\Phi) = -\frac{\lambda v^4}{4} + \lambda v^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4. \quad (1.9)$$

Finalement, le lagrangien associé au boson de Higgs est

$$\mathcal{L}_h = |D_\mu \Phi|^2 - V(\Phi) \quad (1.10)$$

où D_μ est la dérivée covariante électrofaible. Le premier terme représente la propagation libre du champ de Higgs, ainsi que son interaction avec les bosons $W^\pm$ et Z^0 . C'est de ce terme que proviennent les masses de ces bosons. Le boson de Higgs n'interagit cependant pas avec le photon et les gluons qui restent non massifs, car il ne possède ni charge électrique, ni charge de couleur. Le deuxième terme est le potentiel de l'équation 1.9 et il inclut des termes d'interaction en h^3 et h^4 , c'est-à-dire à trois et quatre bosons H .

L'interaction du champ de Higgs avec les fermions leur donne leur masse. Il est représenté par un terme de Yukawa de la forme

$$\mathcal{L}_y = -y^{ij} \bar{\psi}_{Li} \Phi \psi_{Rj} \quad (1.11)$$

avec y^{ij} une matrice représentant les couplages de Yukawa des fermions considérés. Les champs $\bar{\psi}_{Li}$ et ψ_{Rj} représentent soit les trois familles de leptons avec i et j les indices des familles : e, μ, τ , soit les trois familles de quarks de type *up* : u, c, t , soit les trois familles de quarks de type *down* : d, s, b . Les matrices y^{ij} peuvent alors être diagonalisées, et on obtient des états propres de masse. La rotation à effectuer dans l'espace des familles pour la diagonalisation n'est pas nécessairement la même pour les leptons, les quarks de type *up*, et les quarks de type *down*. La masse d'un fermion f état propre de masse s'exprime finalement en fonction du couplage de Yukawa correspondant après diagonalisation λ_f :

$$m_f = \frac{\lambda_f}{\sqrt{2}} v. \quad (1.12)$$

Il est possible de transformer un quark en un autre via l'émission ou l'absorption d'un boson W , par exemple $t \rightarrow bW$. Les types de quarks qui se couplent avec le boson W sont appelés états propres de « saveur ». Il y a donc un couplage entre saveurs de type *up* : u, c, t , et saveurs de type *down* : d, s, b via le boson W . Toutefois la rotation pour diagonaliser la matrice y^{ij} est différente pour les deux types de quarks, ce qui implique que les états propres de masse ne sont pas alignés avec les états propres de saveur. C'est ce phénomène qui explique par exemple que la désintégration $t \rightarrow bW$ a une probabilité proche de 1, mais que les désintégrations $t \rightarrow sW$ et $t \rightarrow dW$ sont également possibles mais peu probables. Les probabilités de ces différents changements de saveur sont résumées dans la matrice CKM [12, 13] qui possède des valeurs diagonales proches de 1, mais des valeurs hors diagonales non nulles. Ceci explique que le changement de saveur via l'émission ou l'absorption d'un boson W est le plus probable entre quarks de la même génération, mais n'est pas impossible entre quarks appartenant à des générations différentes.

1.5 La chromodynamique quantique (QCD)

Finalement, l'interaction nucléaire forte est décrite par la chromodynamique quantique, ou *Quantum Chromo-Dynamics* (QCD). Elle agit sur tous les quarks, mais pas sur les leptons. Les quarks possèdent une charge de couleur, et les anti-quarks une anti-charge de couleur. Les charges de couleur sont au nombre de trois : rouge (R), vert (G), et bleu (B). On peut donc dire qu'il existe trois versions de chaque quark dans le modèle standard, une pour chaque charge de couleur. Le gluon est le médiateur de la QCD, et possède une charge ainsi qu'une anti-charge de couleur. Les quarks peuvent ainsi changer de couleur via l'émission ou l'absorption d'un gluon. Le gluon est un boson sans masse, qui peut lui-même interagir via l'interaction forte.

La théorie QCD est associée au groupe de symétrie $SU(3)_c$ possédant huit générateurs, correspondant aux huit gluons. Elle s'exerce sur les quarks et les gluons, appelés partons.

Le caractère non-abélien du groupe de symétrie entraîne l'existence de couplages à trois ou quatre gluons. Le lagrangien de la QCD est le suivant :

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu}, \quad (1.13)$$

où ψ et $\bar{\psi}$ représentent les champs des quarks, un pour chaque saveur de quark et un pour chaque couleur, $D_\mu = \partial_\mu + ig_s \mathcal{A}_\mu^a \lambda_a/2$ est la dérivée covariante avec a les indices des huit champs de gluons, et les termes $G_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + g_s f^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c$ sont les tenseurs chromodynamiques. Les termes $\mathcal{A}_\mu^a$ représentent les huit champs de gluons, les termes f^{abc} sont les constantes de structure de $SU(3)$, g_s est la constante de couplage, et les matrices 3×3 λ^a sont les huit matrices de Gell-Mann. On appelle souvent couplage fort la grandeur $\alpha_s = g_s^2/4\pi$. Le premier terme de l'équation 1.13 représente la propagation libre des quarks, et l'interaction avec les champs de gluons. Le second terme est un terme de masse pour les quarks, et enfin le dernier terme représente la propagation des champs de gluons, ainsi que les interactions entre gluons.

Les particules hadroniques observées dans la nature se présentent sous la forme d'un état lié d'un quark et d'un anti-quark ou d'un état lié de trois quarks. Les particules du premier type sont appelées mésons, et celles du second baryons. Aucun parton, quark ou gluon, n'est observable directement, seuls les mésons et les baryons le sont. Ceci est modélisé via la charge de couleur que portent les partons. La QCD stipule en effet que seules les particules ayant une charge de couleur neutre peuvent être observées, ce qui est possible en associant une charge de couleur avec son anti-charge comme pour les mésons, ou bien en associant les trois charges ou les trois anti-charges au sein d'un baryon. Cette propriété de la théorie QCD est appelée confinement.

Cette modélisation des particules hadroniques est cependant incomplète. Les deux (respectivement trois) quarks constitutifs d'un méson (respectivement baryon) ne sont pas les seuls partons à l'intérieur d'un hadron. Ils sont appelés quarks de valence et cohabitent avec d'autres quarks et gluons désignés par le terme « mer de quarks et de gluons ». Ces derniers sont des particules virtuelles produites via des émissions de gluons, des créations de paires de quarks et anti-quarks, ou encore par annihilation de ces paires.

La constante de couplage α_s de l'interaction forte varie avec l'énergie du processus considéré. À haute énergie les partons sont libres, on parle de liberté asymptotique. La théorie est renormalisable avec une dépendance de α_s à une échelle de renormalisation μ_R qui introduit donc une incertitude théorique. Pour des faibles énergies, la théorie est non-perturbative et les partons sont confinés dans les hadrons de couleur neutre.

1.6 Bilan et limites du modèle standard

Finalement le groupe de jauge du MS est $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$, avec la symétrie électrofaible brisée par le mécanisme de Higgs. Le lagrangien du MS est égal à la somme du

lagrangien électrofaible $\mathcal{L}_{EW}$ (équation 1.5) avec les termes de Higgs $\mathcal{L}_h$ (équation 1.10) et des couplages de Yukawa $\mathcal{L}_y$ (équation 1.11), et du lagrangien QCD $\mathcal{L}_{QCD}$ (équation 1.13) privé de son terme de masse, puisque la masse des quarks est déjà décrite par le terme de Yukawa.

Le modèle standard présente cependant plusieurs problèmes qui motivent la recherche de nouveaux modèles. La différence très importante entre l'échelle d'énergie électrofaible de l'ordre de 100 GeV et celle de la gravitation représentée par la masse de Planck située à environ 10^{19} GeV constitue le problème de hiérarchie. Le calcul de la masse du boson de Higgs fait apparaître un réglage très fin des paramètres afin qu'elle soit stable à haute énergie, c'est le problème de naturalité ou d'ajustement fin. Le MS contient beaucoup de paramètres libres, et n'explique pas la masse des neutrinos ni le phénomène d'oscillations qui ont été observés expérimentalement. La gravitation décrite par la relativité générale est absente du MS, tout comme la matière noire et l'énergie sombre qui sont les composants majoritaires de l'univers d'après le modèle standard de la cosmologie Λ CDM. L'asymétrie entre la matière et l'antimatière reste également une question ouverte.

Cette liste non-exhaustive d'arguments amène à penser le MS comme un cas limite d'une théorie plus large. Un effort majeur de la recherche contemporaine en physique des particules est donc d'élaborer et de tester de nouveaux modèles corrigeant tout ou partie des problèmes du MS. On qualifie ces théories d'« au-delà du modèle standard », ou *Beyond Standard Model* (BSM), ou encore de « Nouvelle Physique ». Quelques modèles BSM sont présentés dans le chapitre 4 et sont testés dans l'analyse de données présentée dans le chapitre 5.

1.7 Tests expérimentaux du modèle standard

Les particules élémentaires naturellement accessibles sont peu nombreuses et se rencontrent généralement à basse énergie. Une partie importante du travail de recherche est ainsi effectuée au niveau des collisionneurs qui accélèrent les particules à haute énergie et transforment cette énergie cinétique en énergie de masse afin de créer de nouvelles particules. Il est possible de distinguer deux stratégies complémentaires dans la recherche de physique au-delà du modèle standard avec les accélérateurs :

- l'étude de régions encore inexplorées de l'espace des phases, typiquement à très haute énergie, afin d'observer directement des phénomènes de nouvelle physique ;
- et la recherche indirecte de phénomènes BSM via la mesure de précision de grandeurs prédites par le MS. Toute déviation significative de la mesure par rapport à la prédiction peut être l'indice d'une nouvelle physique à une énergie inatteignable à ce jour mais ayant des effets à plus basse énergie. Il est également possible de tester la cohérence du MS de cette façon. Par exemple, la figure 1.3 montre la différence entre l'ajustement global des paramètres de la théorie électrofaible à

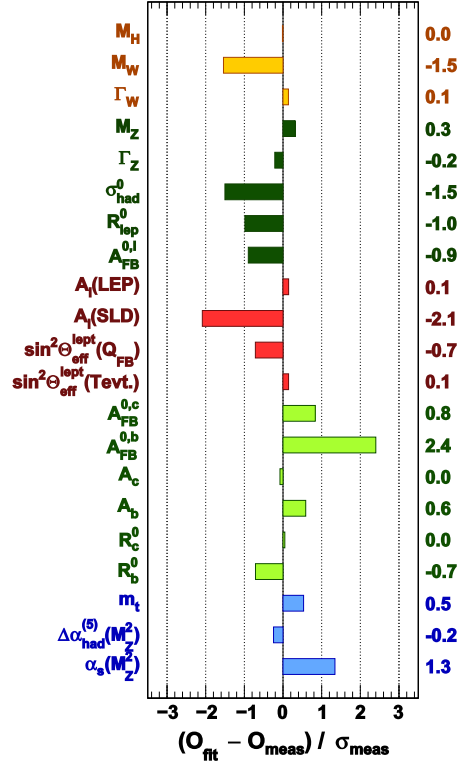


Figure 1.3 – Comparaison de l’ajustement global des paramètres de la théorie électrofaible à partir des résultats obtenus auprès de différents collisionneurs, avec les mesures individuelles. Les résultats sont exprimés en unités d’incertitudes expérimentales des mesures individuelles [14].

partir des résultats obtenus auprès de différents collisionneurs, avec les mesures individuelles. La figure montre une bonne compatibilité entre tous les paramètres en général.

L’accélérateur de particules le plus énergétique construit à ce jour est le *Large Hadron Collider* (LHC). Il concentre une grande partie de la recherche contemporaine en physique des particules. Le détecteur ATLAS est l’un des détecteurs étudiant les collisions au sein du LHC dans le but de tester le modèle standard, ainsi que de nombreux modèles de nouvelle physique. La description de ce détecteur est l’objet du chapitre 2.

Références

- [1] Emmy NOETHER. “Invariant variation problems”. In : *Transport Theory and Statistical Physics* 1.3 (jan. 1971), p. 186-207. DOI : 10.1080/00411457108231446.
- [2] Wikimedia COMMONS. *File :Standard Model of Elementary Particles-fr.svg* — *Wikimedia Commons*. 2018. URL : https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles-fr.svg&oldid=321732592.
- [3] Sheldon L. GLASHOW. “The renormalizability of vector meson interactions”. In : *Nuclear Physics* 10 (fév. 1959), p. 107-117. DOI : 10.1016/0029-5582(59)90196-8.
- [4] Abdus SALAM et J. C. WARD. “Weak and electromagnetic interactions”. In : *Il Nuovo Cimento* 11.4 (fév. 1959), p. 568-577. DOI : 10.1007/bf02726525.
- [5] Steven WEINBERG. “A Model of Leptons”. In : *Physical Review Letters* 19.21 (nov. 1967), p. 1264-1266. DOI : 10.1103/physrevlett.19.1264.
- [6] Particle Data GROUP. “Review of Particle Physics”. In : *Physical Review D* 98.3 (août 2018). DOI : 10.1103/physrevd.98.030001.
- [7] F. ENGLERT et R. BROUT. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In : *Physical Review Letters* 13.9 (août 1964), p. 321-323. DOI : 10.1103/physrevlett.13.321.
- [8] Peter W. HIGGS. “Broken Symmetries and the Masses of Gauge Bosons”. In : *Physical Review Letters* 13.16 (oct. 1964), p. 508-509. DOI : 10.1103/physrevlett.13.508.
- [9] G. S. GURALNIK, C. R. HAGEN et T. W. B. KIBBLE. “Global Conservation Laws and Massless Particles”. In : *Physical Review Letters* 13.20 (nov. 1964), p. 585-587. DOI : 10.1103/physrevlett.13.585.
- [10] ATLAS COLLABORATION. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In : *Physics Letters B* 716.1 (sept. 2012), p. 1-29. DOI : 10.1016/j.physletb.2012.08.020.
- [11] CMS COLLABORATION. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In : *Physics Letters B* 716.1 (sept. 2012), p. 30-61. DOI : 10.1016/j.physletb.2012.08.021.
- [12] Nicola CABIBBO. “Unitary Symmetry and Leptonic Decays”. In : *Physical Review Letters* 10.12 (juin 1963), p. 531-533. DOI : 10.1103/physrevlett.10.531.
- [13] Makoto KOBAYASHI et Toshihide MASKAWA. “CP-Violation in the Renormalizable Theory of Weak Interaction”. In : *Progress of Theoretical Physics* 49.2 (fév. 1973), p. 652-657. DOI : 10.1143/ptp.49.652.

- [14] J. HALLER et al. “Update of the global electroweak fit and constraints on two-Higgs-doublet models”. In : *The European Physical Journal C* 78.8 (août 2018). DOI : 10.1140/epjc/s10052-018-6131-3.

Le détecteur ATLAS au LHC

2.1 Le grand collisionneur de hadrons (LHC)

Le grand collisionneur de hadrons, ou *Large Hadron Collider* (LHC) [1], est le plus grand accélérateur de particules construit à ce jour. Il s'agit d'un anneau souterrain d'environ 27 km de circonférence situé au Cern (Conseil Européen pour la Recherche Nucléaire) à Genève en Suisse. Le LHC est également le collisionneur de particules le plus énergétique à ce jour, avec 13 TeV d'énergie dans le centre de masse pour les collisions proton-proton.

Deux faisceaux de particules sont accélérés en sens opposés et se croisent à quatre emplacements le long de l'anneau. Les quatre principales expériences de physique du LHC sont situées à ces quatre points de croisements des faisceaux. Les expériences ATLAS (*A Toroidal LHC ApparatuS*) [2] et CMS (*Compact Muon Solenoid*) [3] étudient une large gamme de processus physiques et recherchent des phénomènes de nouvelle physique. L'expérience LHCb (*LHC beauty*) [4] est dédiée à la physique du quark b , et l'expérience ALICE (*A Large Ion Collider Experiment*) [5] s'intéresse aux collisions d'ions lourds. L'organisation des quatre détecteurs le long de l'anneau du LHC est montré sur la figure 2.1.

Les particules accélérées dans le LHC sont le plus souvent des protons dans chaque direction. Cependant d'autres modes de collisions ont été utilisés avec des ions lourds : proton-plomb et plomb-plomb. Dans le cadre de cette thèse, seules les collisions proton-proton sont étudiées.

Les protons dans le LHC proviennent d'atomes d'hydrogène ionisés accélérés en plusieurs étapes : l'accélérateur linéaire Linac 2, le PSB (*Proton Synchrotron Booster*), le PS (*Proton Synchrotron*), et enfin le SPS (*Super Proton Synchrotron*). Chacun de ces accélérateurs permet l'augmentation de l'énergie des protons : 50 MeV, 1,4 GeV, 25 GeV, 450 GeV. Ils sont finalement accélérés jusqu'à une énergie de 6,5 TeV dans le LHC. La chaîne d'accélérateurs menant au LHC, ainsi que d'autres accélérateurs au Cern sont représentés sur la figure 2.2.

Au sein du LHC les protons sont accélérés le long de l'anneau par huit cavités radiofréquences, la courbure de leur trajectoire est assurée par des aimants dipôles

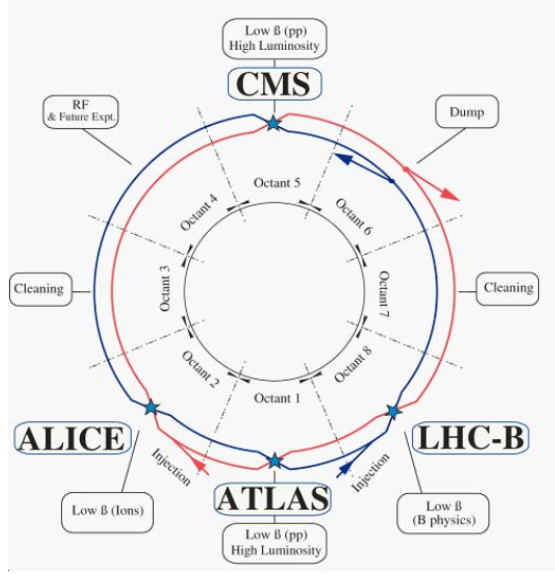


Figure 2.1 – Schéma du LHC montrant les points de croisement des faisceaux le long de l'anneau ainsi que les expériences associées.

supraconducteurs refroidis à environ 2 K et produisant un champ magnétique d'environ 8 T, et d'autres aimants comme des quadrupôles focalisent le faisceau de particules. Les deux faisceaux de protons ne sont pas continus mais sont composés de paquets de particules appelés *bunchs*. Il y a au maximum 2800 *bunchs* par faisceau, espacés dans le temps de 25 ns. Autrement dit le taux de croisement de faisceaux est de 40 MHz en un point de collision comme le centre du détecteur ATLAS. Lors d'un croisement de faisceaux plusieurs interactions proton-proton peuvent avoir lieu, il s'agit du phénomène d'empilement, ou *pile-up*. L'empilement moyen μ est un paramètre important des prises de données au LHC. Cet effet peut provenir du même croisement de faisceau que la collision d'intérêt (empilement synchrone), ou peut être rémanent d'un croisement de faisceau précédent (empilement asynchrone).

Le nombre de collisions proton-proton au cours du temps est représenté par la luminosité L , dite luminosité instantanée. Elle peut être exprimée en fonction des paramètres de fonctionnement du collisionneur :

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (2.1)$$

où N_b est le nombre de particules dans un paquet, n_b le nombre de paquets par faisceau, f_{rev} la fréquence de révolution, γ_r le facteur gamma relativiste, ϵ_n l'émittance transverse normalisée (par γ_r) du faisceau, β^* la fonction d'amplitude au point de collision qui dépend de l'émittance et de la taille transverse du faisceau, et enfin F est le facteur

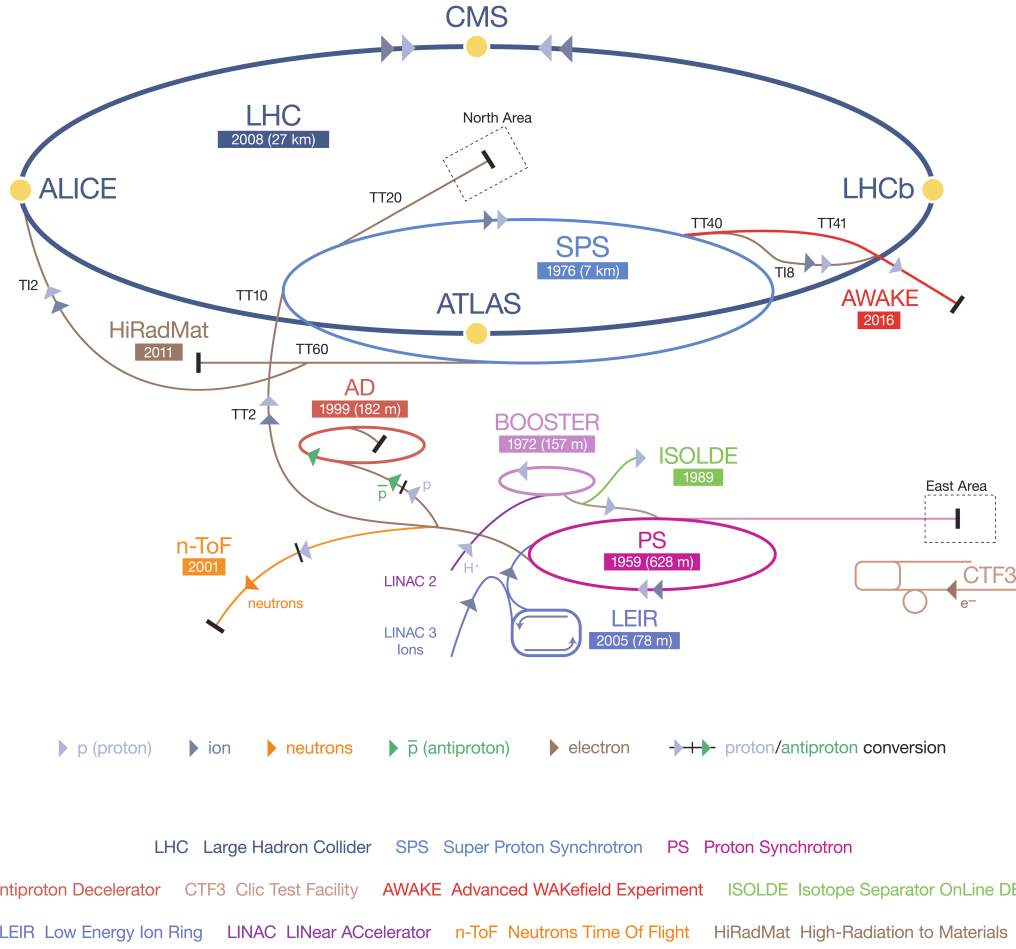


Figure 2.2 – Complexe d’accélérateurs du Cern. La chaîne d’accélération permettant l’injection des protons dans le LHC y est représentée. Le trajet des protons est indiqué par les flèches de couleur gris clair. [6]

de réduction géométrique dû à l’angle de croisement des faisceaux au niveau du point d’interaction.

Lors de l’étude d’un processus de section efficace σ , on peut estimer le nombre d’événements par seconde

$$\frac{dN}{dt} = L \times \sigma. \quad (2.2)$$

Afin de caractériser une quantité de données récoltée pendant une certaine période, la luminosité intégrée $\mathcal{L}$ est utilisée $\mathcal{L} = \int L dt$. Le nombre d’événements attendu pour un phénomène de section efficace σ est alors

$$N = \mathcal{L} \times \sigma. \quad (2.3)$$

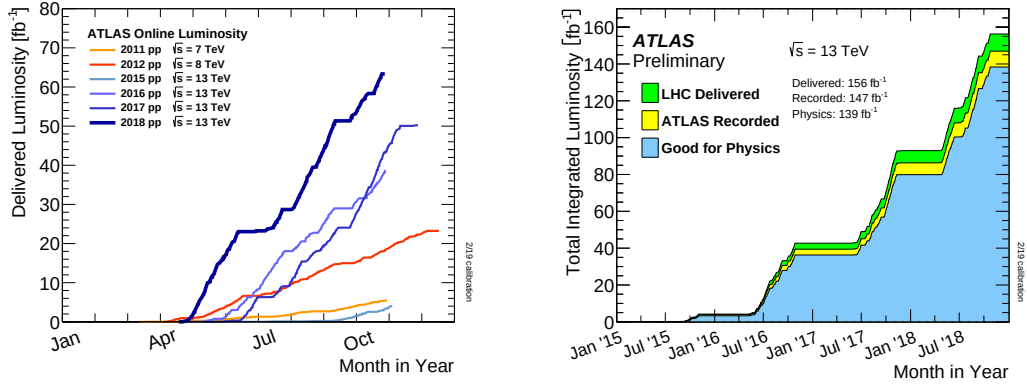


Figure 2.3 – Luminosité intégrée fournie par le LHC durant les différentes prises de données effectuées entre 2011 et 2018 (gauche), et enregistrée et exploitée par le détecteur ATLAS durant le Run 2 (droite).

Un détecteur comme ATLAS ne peut pas enregistrer la totalité des collisions ayant lieu. Il existe en effet des périodes de temps mort d'un ou plusieurs sous-détecteurs ou de leur électronique. C'est pourquoi la luminosité intégrée enregistrée par un détecteur est inférieure à la luminosité intégrée délivrée par l'accélérateur. La figure 2.3 montre la luminosité intégrée fournie par le LHC dans le détecteur ATLAS pour les différentes prises de données effectuées entre 2011 et 2018, ainsi que la luminosité intégrée enregistrée et exploitée par ATLAS durant tout le Run 2.

La première campagne de prise de données au LHC, appelée Run 1 s'est déroulée en 2011 et 2012, avec des énergies dans le centre de masse de 7 TeV et 8 TeV respectivement. Elle a été suivie par un arrêt d'environ deux ans avant le démarrage d'une nouvelle campagne de prise de données appelée Run 2 à une énergie de 13 TeV dans le centre de masse. Cette campagne débutée en 2015 s'est terminée fin 2018 et est suivie d'un second arrêt de deux ans avant un redémarrage pour le Run 3, puis pour le projet HL-LHC [7] prévu pour 2026 environ. L'organisation dans le temps des différentes prises de données est montré sur la figure 2.4.

Les valeurs de luminosité intégrée exploitée par le détecteur ATLAS au cours des différentes prises de données sont

- pour le Run 1 : $4,6 \text{ fb}^{-1}$ en 2011 et $20,3 \text{ fb}^{-1}$ en 2012 ;
- pour le Run 2 : $3,2 \text{ fb}^{-1}$ en 2015, $33,0 \text{ fb}^{-1}$ en 2016, $44,3 \text{ fb}^{-1}$ en 2017, et $58,5 \text{ fb}^{-1}$ en 2018.

La luminosité intégrée du Run 2 est donc d'environ 139 fb^{-1} . Les valeurs d'empilement moyen μ ont grandement varié tout au long du Run 2, et leurs valeurs sont présentées dans la figure 2.5. L'empilement moyen est proche de 40 pour les prises de données de 2017 et 2018. Il valait environ 20 durant le Run 1, et va continuer d'augmenter, pour

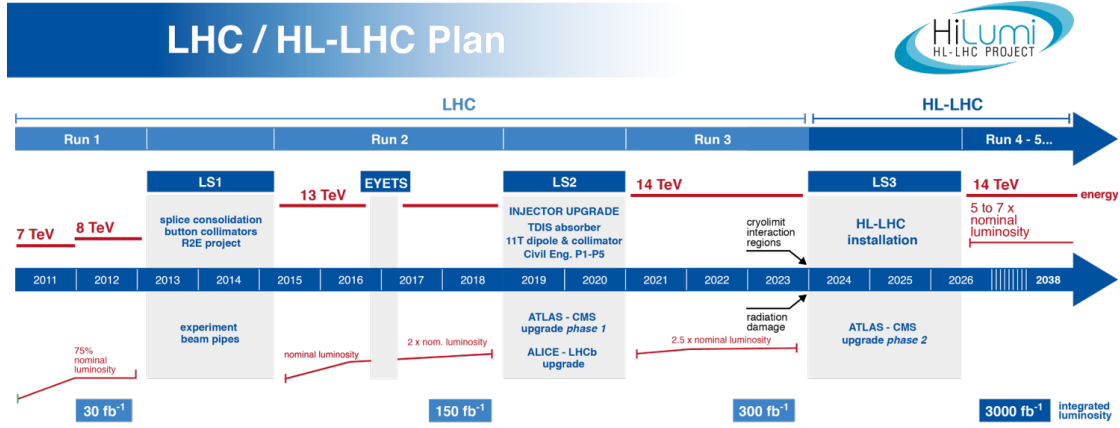


Figure 2.4 – Organisation dans le temps des différentes prises de données au LHC, avec les prévisions pour le projet HL-LHC.

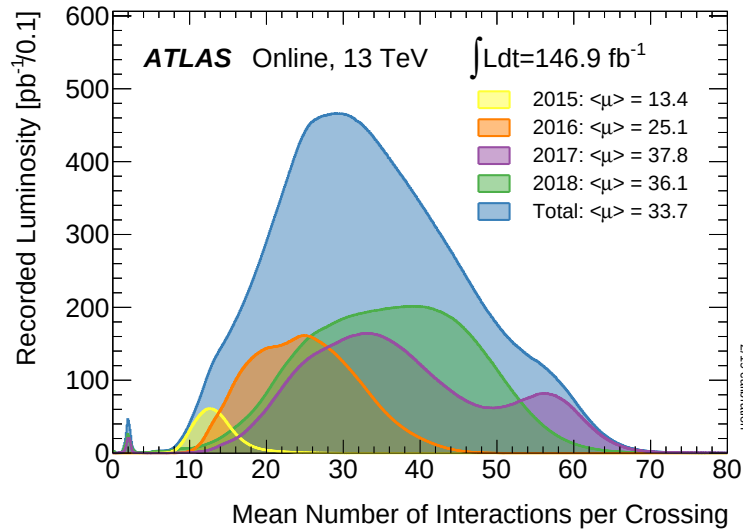


Figure 2.5 – Valeurs d'empilement moyen μ dans le détecteur ATLAS pour les prises de données du Run 2.

atteindre jusqu'à 200 au HL-LHC.

2.2 Le détecteur ATLAS

Le détecteur ATLAS [2] est un détecteur de très grande taille, environ 25 m de diamètre et 44 m de long. Sa géométrie d'ensemble est celle d'un cylindre couché traversé par l'axe du faisceau de protons du LHC. Les deux extrémités plates du cylindre sont appelées

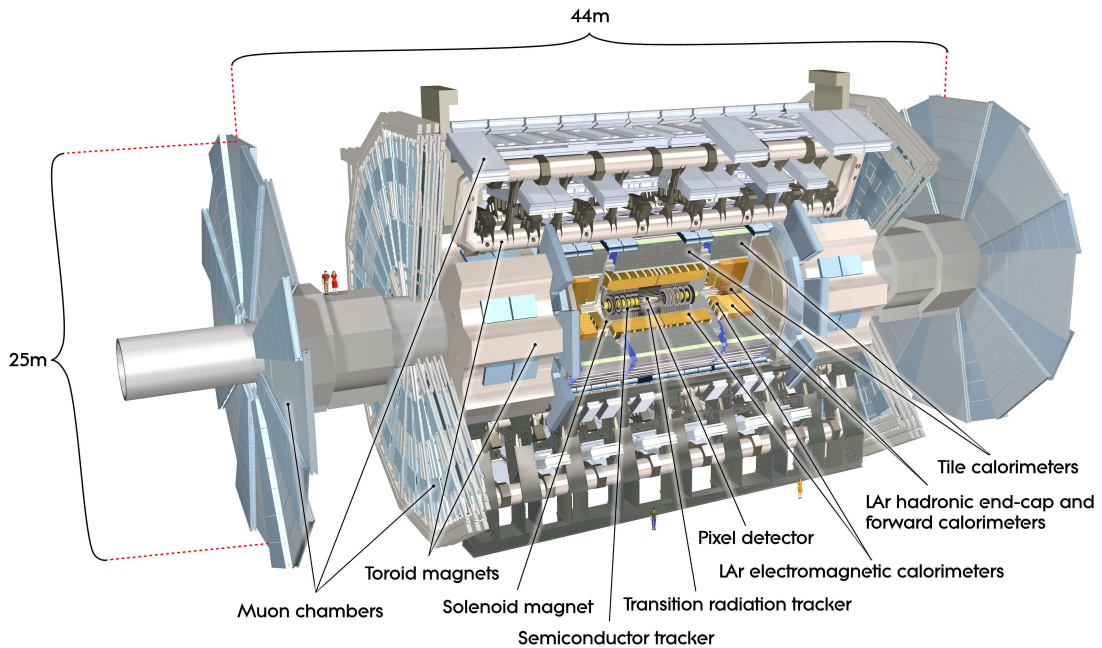


Figure 2.6 – Vue d’ensemble du détecteur ATLAS et de ses principaux composants.

détecteurs avant-arrière ou « bouchons » tandis que le corps du cylindre symétrique autour de l’axe du faisceau est appelé détecteur central ou « tonneau ». La figure 2.6 montre une vue d’ensemble du détecteur ATLAS et de ses principaux composants. Certains de ces composants vont être remplacés ou mis à jour en vue du projet HL-LHC, notamment les « petites roues » du spectromètre à muons. Les nouveaux détecteurs destinés à remplacer des petites roues sont l’objet du chapitre 3.

Les différents sous-détecteurs constitutifs d’ATLAS sont empilés en couches autour de l’axe du faisceau. Le détecteur doit être en mesure d’identifier le plus grand nombre de particules créées lors des collisions, et de mesurer leurs caractéristiques. C’est pourquoi ATLAS couvre presque tout l’espace autour du faisceau, et que plusieurs couches de détection sont mises en place. Chaque sous-détecteur a un rôle particulier, et leur combinaison permet une identification des particules générées lors des collisions. En partant du faisceau de proton et en allant vers la couche extérieure du détecteur, on trouve

- le trajectographe interne dont le rôle est de mesurer la trajectoire des particules chargées ;
- le calorimètre électromagnétique qui mesure l’énergie des électrons et photons tout en les stoppant ;
- le calorimètre hadronique dont la fonction est similaire à celle du calorimètre électromagnétique, pour les particules hadroniques ;

- le spectromètre à muons qui mesure l'impulsion des muons, particules capables de sortir complètement du détecteur.

Le détecteur ATLAS comporte également deux systèmes d'aimants supraconducteurs afin de courber les traces des particules chargées et ainsi de mesurer leur impulsion.

- L'aimant solénoïdal est situé entre le détecteur interne et le calorimètre électromagnétique. Il produit un champ de 2 T le long de l'axe du faisceau afin de courber les traces des particules dans le plan transverse.
- Huit aimants toroïdaux sont situés autour du calorimètre hadronique, et huit autres plus petits au niveau de chaque bouchon. Ils produisent un champ d'environ 0,5 à 1 T au niveau du spectromètre à muons.

L'identification des particules est effectuée en combinant les informations des différents sous-détecteurs. Par exemple un électron engendre une trace courbée dans le trajectographe, puis s'arrête complètement en laissant un dépôt d'énergie dans le calorimètre électromagnétique. Un muon produit une trace dans tous les sous-détecteurs et peut en sortir complètement. Les particules neutres ne laissent pas de trace mais peuvent être identifiées grâce à leur dépôt d'énergie dans les calorimètres, sauf les neutrinos qui quittent le détecteur sans interagir. Afin d'estimer l'énergie des neutrinos, l'énergie transverse manquante E_T^{miss} est utilisée. Elle est construite à partir de la conservation de l'impulsion dans le plan transverse, comme précisé dans la section 2.3.5.

2.2.1 Le système de coordonnées d'ATLAS

Le point d'interaction est fixé comme étant le centre du système de coordonnées utilisé dans le détecteur ATLAS. L'axe x pointe en direction du centre de l'anneau du LHC, l'axe y pointe vers le haut, et l'axe z est dirigé suivant le faisceau de manière à former un repère direct. Dans le plan transverse les coordonnées cartésiennes (x, y) ainsi que les coordonnées cylindriques (R, ϕ) peuvent être utilisées. ϕ est appelé angle azimutal. L'angle polaire θ est défini comme l'angle par rapport à l'axe z , ainsi $\theta \in [0, \pi]$. On définit la rapidité

$$\varphi = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (2.4)$$

avec E l'énergie de la particule et p_z la projection de son impulsion suivant l'axe z . L'intérêt d'utiliser la rapidité φ plutôt que l'angle polaire θ est qu'elle est invariante par transformation de Lorentz. Dans le cas d'un objet massif ultra-relativiste $E \gg m$, la pseudo-rapacité

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right), \quad (2.5)$$

est une bonne approximation de la rapidité. Dans le cas d'une particule sans masse, on a $\varphi = \eta$. La valeur $\eta = 0$ ($\theta = \pi/2$) correspond à une position perpendiculaire à l'axe z , tandis qu'un objet sur l'axe du faisceau est repéré par $\eta = \pm\infty$ ($\theta = 0$ ou $\theta = \pi$). Le

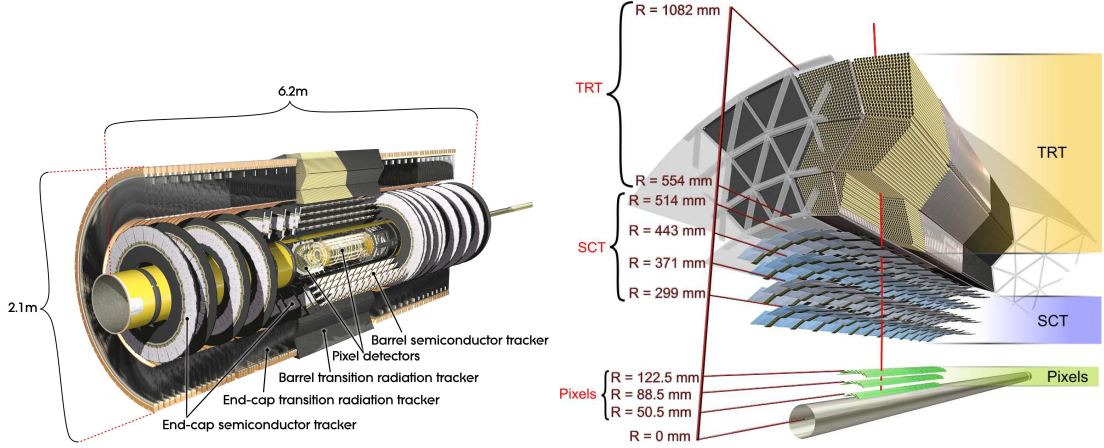


Figure 2.7 – Vue d’ensemble du trajectographe interne et de ses différents composants (gauche). Organisation et dimensions des sous-parties du détecteur interne dans le tonneau (droite). La couche-b n’est pas représentée.

signe de η correspond à la position avant ($\theta \in [0, \pi/2]$) ou arrière ($\theta \in [\pi/2, \pi]$) dans le détecteur, on s’intéresse donc généralement à $|\eta|$ par symétrie. Le détecteur ATLAS couvre l’espace défini par $|\eta| < 4,9$.

L’impulsion transverse p_T des traces reconstruites, ainsi que l’énergie transverse manquante E_T^{miss} dans un événement sont définies comme des projections dans le plan transverse (x, y) . On définit la distance angulaire ΔR entre deux particules à partir de leur différence de pseudo-rapacité $\Delta\eta$ et de leur différence d’angle azimutal $\Delta\phi$,

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (2.6)$$

2.2.2 Le trajectographe interne

Le sous-détecteur le plus proche du faisceau de protons est le trajectographe interne [8]. Il se situe dans la zone de plus forte densité de traces, et a pour rôle de détecter les vertex primaires et secondaires, et de mesurer la charge et l’impulsion des particules chargées le traversant. Ce sous-détecteur intervient également dans l’identification des jets de quark b , décrite dans la section 2.3.4. Il est constitué de trois sous-parties

- le détecteur à pixels et la couche-b ou *b-layer* ;
- le trajectographe à semi-conducteurs (SCT) ;
- le trajectographe à rayonnement de transition (TRT).

Le détecteur interne couvre la région de l’espace définie par $|\eta| < 2,5$. Il est représenté sur la figure 2.7.

Le détecteur à pixels est constitué de quatre couches de détection en silicium semi-conducteur, il fournit typiquement quatre positions tri-dimensionnelles par trace. La

couche la plus proche du faisceau est appelée couche-b ou *b-layer* [9]. Elle se situe à environ 33 mm du faisceau et n'est pas représentée sur la figure 2.7 puisqu'elle a été ajoutée entre le Run 1 et le Run 2. Les couches du détecteur à pixels sont arrangées de manière cylindrique dans le tonneau, et consistent en trois disques dans les bouchons. Ce sous-détecteur est le plus proche du faisceau, il doit donc avoir une grande granularité et une très bonne résolution spatiale. Il est constitué de pixels de taille $50\,\mu\text{m} \times 400\,\mu\text{m}$ ($250\,\mu\text{m}$ pour la couche-b), permettant une résolution de $115\,\mu\text{m}$ en z , et de $10\,\mu\text{m}$ en $R - \phi$ ($40\,\mu\text{m}$ et $8\,\mu\text{m}$ pour la couche-b).

Le trajectographe à semi-conducteurs (SCT) est constitué de micro-pistes en silicium semi-conducteur sur quatre couches dans le tonneau et neuf disques dans les bouchons. Il fournit typiquement quatre points de mesure pour la trajectoire et permet une mesure de l'impulsion grâce au bras de levier plus grand que pour le détecteur à pixels. Les quatre couches peuvent être vues comme deux paires comportant une couche avec des pistes parallèles au faisceau, et une autre avec des pistes légèrement déviées d'un angle de $40\,\text{mrad}$. La résolution suivant z est de $580\,\mu\text{m}$ et la résolution en $R - \phi$ est de $17\,\mu\text{m}$.

Le trajectographe à rayonnement de transition (TRT) constitue la partie externe du trajectographe interne. Il est constitué de tubes polymères à dérive de $4\,\text{mm}$ contenant un mélange de gaz Xe-CO₂-O₂ et un fil de tungstène. Le principe de détection est l'ionisation du gaz par les photons émis par les particules chargées lors d'un changement de milieu. L'amplitude du rayonnement de transition permet de distinguer les électrons des hadrons légers. Les tubes sont disposées parallèlement au faisceau sur trois couches dans le tonneau, et arrangées radialement sur deux disques dans les bouchons. Le TRT fournit en moyenne 36 points de mesure par trace jusqu'à $|\eta| < 2,1$, pour une résolution en $R - \phi$ de $130\,\mu\text{m}$ dans le tonneau et une résolution identique suivant z dans les bouchons.

La résolution globale du trajectographe interne sur le paramètre d'impact transverse σ_{d_0} est d'environ $22\,\mu\text{m}$ pour des traces de haute impulsion. La résolution sur les paramètres d'impact transverse et longitudinal des traces en fonction de leur impulsion transverse est montrée sur la figure 2.8. La résolution globale relative sur l'impulsion des traces est d'environ $\sigma_{p_T}/p_T = 0,05\,\% p_T[\text{GeV}] \oplus 1\,\%$, où le symbole $\oplus$ désigne une somme quadratique.

2.2.3 Les calorimètres

Le trajectographe interne ainsi que l'aimant solénoïde sont situés à l'intérieur du système de calorimétrie, composé d'un calorimètre électromagnétique entouré par un calorimètre hadronique. Les deux calorimètres à échantillonnage couvrent une région de l'espace définie par $|\eta| < 4,9$, ils sont représentés sur la figure 2.9. Le rôle des calorimètres est de mesurer l'orientation ainsi que l'énergie de toutes les particules sauf les muons et les neutrinos. Les muons déposent de l'énergie dans les calorimètres mais n'y sont pas stoppés, tandis que les neutrinos n'interagissent pas du tout. Les calorimètres doivent

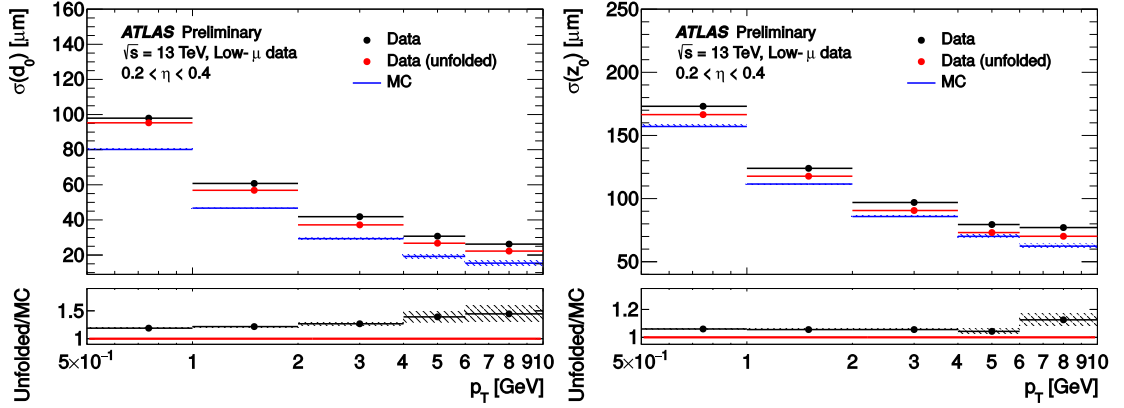


Figure 2.8 – Résolution du trajectographe interne sur les paramètres d’impact transverse d_0 et longitudinal z_0 des traces en fonction de leur impulsion transverse. Ces résultats sont issus de données prises en 2015 avec un faible empilement, pour des traces dans la région $0,2 < |\eta| < 0,4$. La résolution obtenue à partir des données est comparée à celle obtenue à partir d’une simulation Monte-Carlo (MC) [10].

donc être en mesure de stopper toutes les autres particules. C’est pourquoi le calorimètre électromagnétique a une épaisseur supérieure à 22 (respectivement 24) longueurs de radiation X_0 dans le tonneau (respectivement les bouchons). La grandeur X_0 est définie pour un électron de haute énergie comme la distance moyenne pour qu’il ne conserve qu’une proportion $1/e$ de son énergie initiale, le reste étant perdu par rayonnement de freinage ou *bremsstrahlung*. Pour un photon, la grandeur X_0 correspond à $7/9$ du libre parcours moyen avant de produire une paire e^+e^- dans le matériau. Similairement à la grandeur X_0 pour les électrons, on définit une longueur caractéristique d’interaction λ pour les particules hadroniques. L’épaisseur totale du calorimètre est d’environ $9,7\lambda$ (respectivement 10λ) dans le tonneau (respectivement les bouchons). Cette largeur atteint 11λ en incluant le support externe, ce qui permet de réduire grandement la probabilité que des particules traversent les calorimètres et déposent de l’énergie dans le spectromètre à muons. Le système de calorimétrie présente une bonne couverture spatiale qui permet une bonne estimation de l’énergie transverse manquante, décrite dans la section 2.3.5.

Le calorimètre électromagnétique

Le calorimètre électromagnétique s’étend sur la région $|\eta| < 3,2$. La granularité est très fine dans le tonneau et plus grossière dans les bouchons. La région de transition entre le tonneau et les bouchons $1,37 < |\eta| < 1,52$ est généralement ignorée par les analyses de physique puisqu’elle contient des câbles et d’autres types de matériel. Le calorimètre est constitué de trois couches en profondeur dans le tonneau :

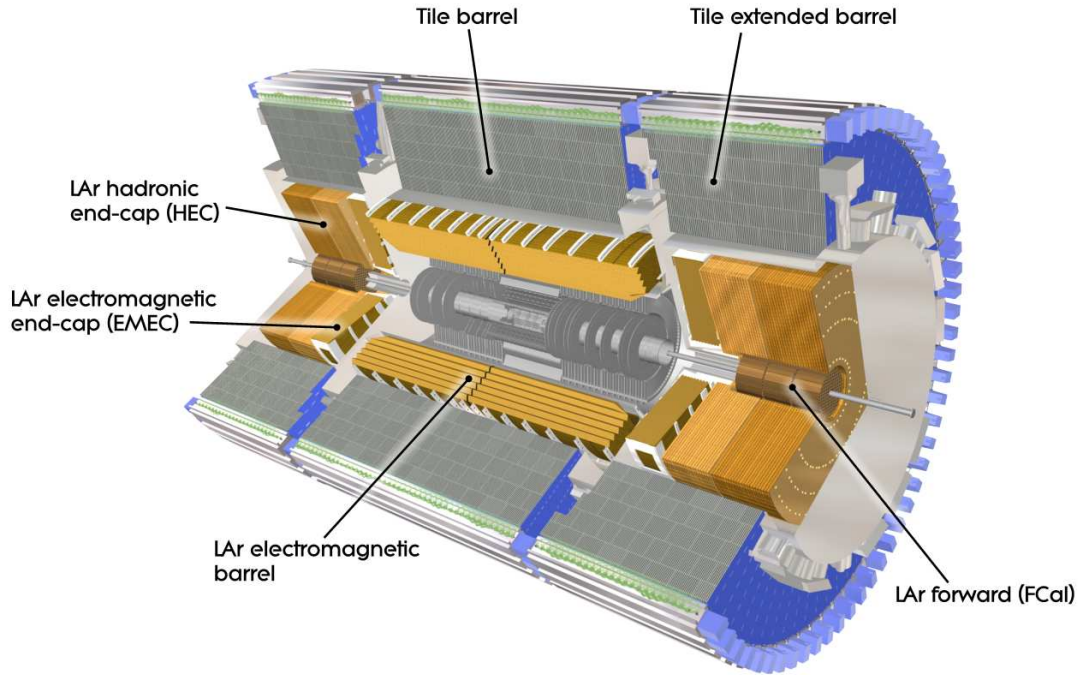


Figure 2.9 – Vue d’ensemble du système de calorimétrie du détecteur ATLAS. Il est constitué de calorimètres à argon liquide désignés par « LAr » et d’un calorimètre à tuiles désigné par « Tile ».

- la première est très granulaire et permet d’identifier des gerbes proches pour distinguer les photons des désintégrations de pions neutres ;
- la seconde est la plus grande et contient la plus grande partie de la gerbe ;
- enfin la troisième couche est fine et permet d’estimer les pertes dues aux gerbes sortant du calorimètre.

Un pré-échantillonneur est situé devant le calorimètre dans la région $|\eta| < 1,8$. Il permet d’estimer les pertes d’énergie des électrons et photons avant leur entrée dans le calorimètre.

Le rôle du calorimètre électromagnétique est de mesurer la direction ainsi que l’énergie des gerbes électromagnétiques, et ainsi d’identifier les électrons et photons. Il s’agit d’un calorimètre à échantillonnage, c’est-à-dire que des couches de milieu actif sont alternées avec des couches d’absorbeur dense. Le milieu actif est de l’argon liquide, remplissant l’espace entre les électrodes de lecture en cuivre et les absorbeurs en plomb, séparés par une structure en nid d’abeille. Les électrodes et les absorbeurs sont disposés en accordéon afin d’assurer une bonne couverture angulaire, comme montré sur la figure 2.10. La gerbe électromagnétique est détectée via une différence de potentiel appliquée entre les électrodes et les plaques d’absorbeur.

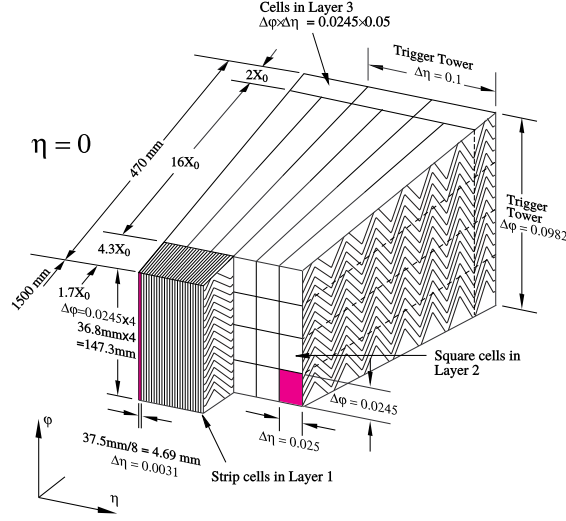


Figure 2.10 – Schéma en coupe du calorimètre électromagnétique d'ATLAS. La taille $(\Delta\eta, \Delta\phi)$ des cellules des différentes couches du calorimètre est donnée, ainsi que leur profondeur en termes de longueur de radiation X_0 . Les trois couches du calorimètre dans le tonneau sont visibles mais le pré-échantillonneur est absent.

La résolution en énergie du calorimètre électromagnétique s'exprime sous la forme d'une somme quadratique de trois termes :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.7)$$

avec $a/\sqrt{E}$ le terme stochastique associé à la gerbe électromagnétique, b/E le terme de bruit provenant de l'électronique et de l'empilement, et c est le terme constant modélisant la non-uniformité de la réponse du calorimètre. Le terme a vaut typiquement $10\% \sqrt{\text{GeV}}$, le terme b a une valeur d'environ 300 MeV, et le terme constant c est de l'ordre de $1 - 2\%$. À haute énergie, c'est le terme constant qui domine la résolution puisqu'il est indépendant de l'énergie. La résolution sur l'énergie des électrons est montrée sur la figure 2.11 en fonction de leur énergie transverse et de $|\eta|$. On observe des valeurs de résolution variant entre environ 1 et 10 %.

Le calorimètre hadronique

Les quarks et les gluons produits dans une interaction proton-proton subissent une hadronisation et génèrent plusieurs particules hadroniques collimatées. Cette gerbe hadronique appelée *jet* dépose une partie de son énergie dans le calorimètre électromagnétique, mais n'y est pas complètement stoppée. Le rôle du calorimètre hadronique est de stopper tous les hadrons et de mesurer les paramètres des jets, en particulier leur énergie. Le calorimètre hadronique est constitué de différentes sous-parties visibles sur la figure 2.9.

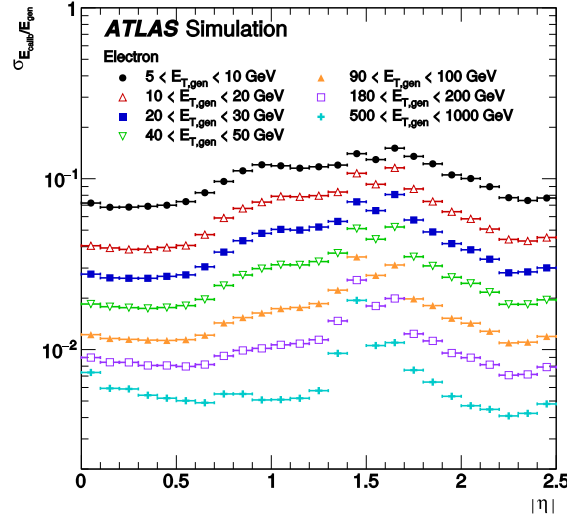


Figure 2.11 – Résolution sur l'énergie des électrons en fonction de leur énergie transverse et de $|\eta|$. Les valeurs sont obtenues grâce à une simulation Monte-Carlo, validée via une étude des résonances $J/\psi \rightarrow e^+e^-$ et $Z \rightarrow e^+e^-$ avec les données prises en 2015 et en 2016 [11].

Une partie cylindrique dans le tonneau entoure le tonneau du calorimètre électromagnétique dans la région $|\eta| < 1,0$ et est complétée par deux extensions cylindriques de part et d'autre dans la région $0,8 < |\eta| < 1,7$. Il s'agit d'un calorimètre à échantillonnage à absorbeur en acier, utilisant des tuiles scintillantes comme milieu actif. Le rayonnement émis dans une gerbe hadronique est collecté par des fibres optiques à décalage de longueur d'onde assurant la transmission à des photomultiplicateurs situés de chaque côté des tuiles. Ces tuiles sont perpendiculaires au faisceau. Le schéma d'un module de ce sous-détecteur est montré sur la figure 2.12.

Deux bouchons sont situés derrière les bouchons du calorimètre électromagnétique et couvrent la région $1,5 < |\eta| < 3,2$. L'absorbeur est ici du cuivre, et le milieu actif de l'argon liquide. Les électrodes sont perpendiculaires au faisceau.

Des calorimètres avant et arrière couvrent la région $3,1 < |\eta| < 4,9$. Ils sont constitués d'un module électromagnétique à absorbeur de cuivre et de deux modules hadroniques à absorbeur de tungstène. Le milieu actif est de l'argon liquide. Ces détecteurs permettent de détecter des traces additionnelles afin d'améliorer l'estimation de l'énergie transverse manquante.

La résolution du calorimètre pour l'énergie des jets s'exprime similairement à l'expression de l'équation 2.7. Le terme de bruit est de l'ordre de $1\% \text{ GeV}/E$ et domine à basse impulsion transverse $p_T \lesssim 30 \text{ GeV}$. Le terme constant quant à lui est de l'ordre de 3% et domine à haut p_T , au-delà de 400 GeV . Pour les valeurs intermédiaires de p_T , c'est le

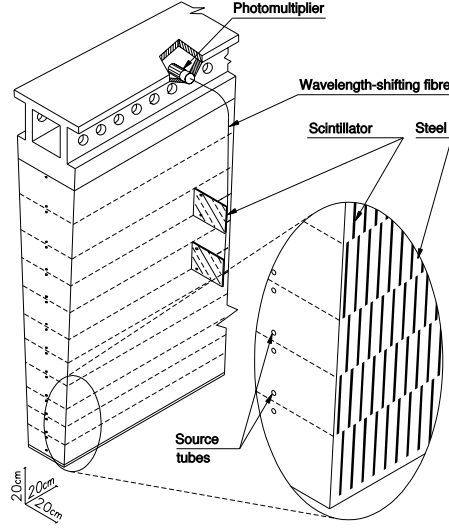


Figure 2.12 – Schéma d'un module du calorimètre hadronique à tuiles scintillantes, montrant l'alternance d'absorbeur en acier et de scintillateurs.

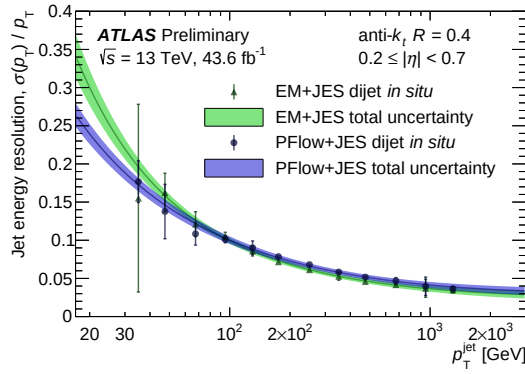


Figure 2.13 – Résolution sur l'impulsion transverse des jets dans les données prises en 2017 pour des jets dans la région $0,2 \leq |\eta| < 0,7$ [12, 13]. Les jets sont reconstruits avec l'algorithme anti- k_T et un paramètre de cône $R = 0,4$, à partir de *TopoClusters* pour les courbes vertes (voir section 2.3.3) ou via l'algorithme *particle flow* pour les courbes bleues [14].

terme stochastique qui est dominant, il est de l'ordre de $50\% \sqrt{\text{GeV}}/\sqrt{E}$. La figure 2.13 montre la résolution sur l'impulsion transverse des jets dans les données prises en 2017 pour les jets dans la région $0,2 \leq |\eta| < 0,7$. La résolution mesurée varie entre environ 4 et 16 %. Le rapport entre données et simulation varie entre 96 et 98 %.

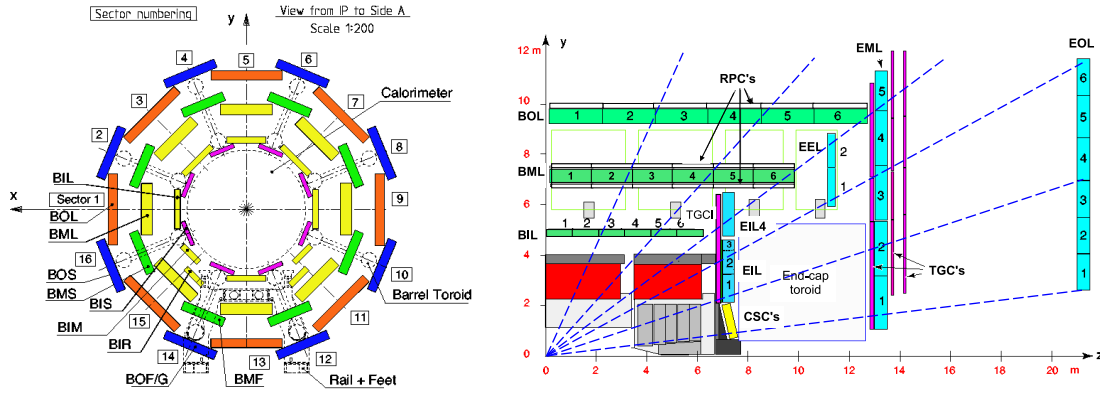


Figure 2.14 – Disposition des composants du spectromètre à muons en coupe transverse (gauche) et longitudinale (droite). Les éléments bleus et verts sur la figure de droite sont des détecteurs MDT, et l'élément jaune représente les détecteurs CSC. Les chambres MDT sont accompagnées de chambres RPC dans le tonneau, et de chambres TGC dans les bouchons.

2.2.4 Le spectromètre à muons

Le spectromètre à muons est situé autour des calorimètres et n'est traversé que par les muons et les neutrinos, ces derniers n'interagissant pas. Son rôle est de mesurer l'impulsion des muons, en tirant profit du champ magnétique généré par les aimants toroïdaux. Tandis que les lignes de champ solénoïdales sont parallèles au faisceau et dévient les trajectoires en ϕ dans le détecteur interne, les lignes de champ toroïdales sont circulaires autour du calorimètre et dévient les trajectoires en η dans le spectromètre. Ce détecteur contient également plusieurs chambres dédiées au déclenchement. La région de l'espace couverte par le spectromètre est $|\eta| < 2,7$. L'arrangement des sous-détecteurs constitutifs du spectromètre à muons est montré sur la figure 2.14.

Le spectromètre est constitué de trois cylindres dans le tonneau, ainsi que de roues dans les bouchons. Les mesures de précision sont assurées par des chambres à tubes à dérivation (MDT) et des chambres à pistes cathodiques (CSC) placées à grand $|\eta|$. Les CSC sont plus segmentées que les MDT pour maintenir le flux de particules très important vers l'avant. Des techniques précises d'assemblage mécanique, combinées à un système optique d'alignement des chambres et entre les chambres permettent un alignement relatif des composants du spectromètre à une précision d'environ $30 \mu\text{m}$.

Le système de déclenchement du spectromètre est constitué de chambres à plaques résistives (RPC) dans le tonneau et de chambres à intervalle fin (TGC) dans les bouchons. Il couvre la région $|\eta| < 2,4$. Les chambres utilisées pour le déclenchement complètent la mesure de l'impulsion réalisée par les MDT et CSC.

La résolution sur l'impulsion des muons est mesurée dans les données prises en 2015

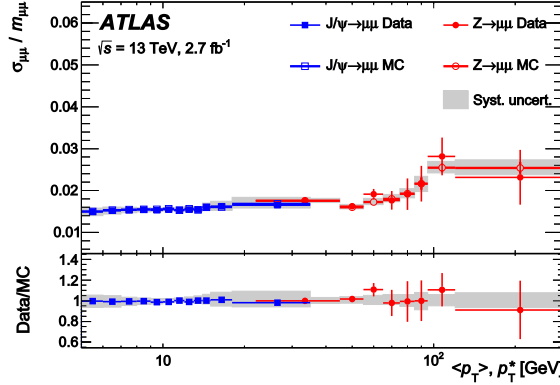


Figure 2.15 – Résolution relative sur la masse reconstruite de la paire de muons mesurée à partir des résonances $J/\psi \rightarrow \mu^+\mu^-$ et $Z \rightarrow \mu^+\mu^-$ dans les données prises en 2015. Les muons doivent passer la sélection CB (voir section 2.3.2). La résolution est présentée en fonction de la valeur moyenne des impulsions transverses des muons $\langle p_T \rangle$ pour $J/\psi \rightarrow \mu^+\mu^-$ et de la variable p_T^* définie dans le texte pour $Z \rightarrow \mu^+\mu^-$ [15].

à partir de paires de muons provenant des résonances $J/\psi \rightarrow \mu^+\mu^-$ et $Z \rightarrow \mu^+\mu^-$ [15]. Quand les deux muons ont une impulsion transverse similaire et qu'on néglige les effets angulaires, on peut exprimer la résolution relative sur l'impulsion transverse des muons σ_{p_μ}/p_μ en fonction de la résolution relative sur la masse reconstruite de la paire de muons $\sigma_{\mu\mu}/m_{\mu\mu}$:

$$\frac{\sigma_{\mu\mu}}{m_{\mu\mu}} = \frac{1}{\sqrt{2}} \frac{\sigma_{p_\mu}}{p_\mu}. \quad (2.8)$$

Les valeurs de $\sigma_{\mu\mu}/m_{\mu\mu}$ sont montrées sur la figure 2.15 en fonction de la valeur moyenne des impulsions transverses des muons $\langle p_T \rangle$ pour $J/\psi \rightarrow \mu^+\mu^-$ et de la variable p_T^* pour $Z \rightarrow \mu^+\mu^-$:

$$p_T^* = m_Z \sqrt{\frac{\sin \theta_1 \sin \theta_2}{2(1 - \cos \alpha_{12})}} \quad (2.9)$$

où m_Z est la masse du boson Z , θ_1 et θ_2 sont les angles polaires des deux muons, et α_{12} est l'angle d'ouverture de la paire de muons. La résolution sur la masse de la paire de muons est d'environ 1,2 % (1,6 %) pour la résonance J/ψ (Z) pour des faibles valeurs de $|\eta|$, et s'élève à environ 1,6 % (1,9 %) pour la résonance J/ψ (Z) dans les bouchons. Ceci correspond à une résolution sur l'impulsion des muons d'environ 1,7 % (2,3 %) au centre du détecteur, et d'environ 2,3 % (2,9 %) dans les bouchons pour la résonance J/ψ (Z).

La mise à jour du détecteur ATLAS haute luminosité ou « HL-LHC » [7] entraînera une augmentation significative de l'irradiation des détecteurs. Les détecteurs à muons situés dans les « petites roues » (voir figure 3.1) doivent être remplacés afin de supporter l'irradiation prévue au HL-LHC tout en maintenant les performances de résolution sur l'impulsion des muons, et surtout pour maintenir le taux de déclenchement L1 (voir

section 2.2.6) à des niveaux acceptables. Les nouveaux détecteurs sont appelés NSW et sont en partie construits par l'Irfu. Le chapitre 3 détaille cette problématique et présente le travail effectué dans le cadre de cette thèse sur le développement d'un code d'analyse d'un banc de test pour ces détecteurs.

2.2.5 Le détecteur LUCID

La calibration absolue de la mesure de luminosité dans le détecteur ATLAS est réalisée grâce à des scans Van de Meer, similairement à la méthode présentée dans la référence [16]. Le détecteur LUCID-2 [17] est un détecteur qui permet la mesure en ligne de luminosité instantanée dans ATLAS. Il est constitué de tubes situés de chaque côté du point d'interaction à environ 17 m, c'est-à-dire dans la région $5,6 < |\eta| < 6,0$. Chaque tube détecte la lumière Tcherenkov produite par les particules chargées qui le traversent. Le taux de particules ainsi détectées permet d'estimer la luminosité.

2.2.6 Le système de déclenchement et d'acquisition

Le taux de croisement de faisceaux au point d'interaction est de 40 MHz, avec plusieurs dizaines d'interactions proton-proton par croisement de faisceaux. Le taux d'interaction est donc de l'ordre du GHz, ce qui est bien trop élevé pour que toutes les collisions soient étudiées, et la quantité de données générées serait également trop grande. Afin de ne conserver que les événements les plus intéressants pour les analyses de physique, un système de déclenchement à trois niveaux est chargé de sélectionner les événements en vue d'enregistrer leurs informations pour une analyse plus poussée [18]. En sortie du système de déclenchement le taux d'événements est de quelques centaines de Hz. Plusieurs déclenchements différents sont disponibles selon les phénomènes qu'on souhaite étudier.

Le premier niveau de déclenchement (L1) utilise des informations de dépôts d'énergie dans le calorimètre avec une faible granularité, ainsi que les informations des chambres de déclenchement du spectromètre à muons (RPC et TGC). Seuls les événements dépassant certains seuils d'énergie ou d'énergie transverse manquante sont conservés. La valeur du seuil dépend du déclenchement choisi. La décision est prise directement au niveau de l'électronique sans traitement logiciel. Des régions d'intérêt sont identifiées et sont envoyées au deuxième niveau de déclenchement (L2). Ainsi les signaux de l'électronique de lecture des détecteurs autres que le calorimètre et les chambres RPC et TGC ne sont lus que si le déclenchement L1 a lieu. Le flux d'événements en sortie du niveau de déclenchement L1 est d'environ 100 kHz pour une latence d'environ 2,5 μ s.

Le second niveau de déclenchement (L2) est logiciel. Il effectue une reconstruction simplifiée dans les régions d'intérêt en utilisant toutes les informations des sous-détecteurs. Plusieurs déclenchements différents sont disponibles, ils sont fondés sur un ou plusieurs

objets physiques en particulier. Un événement sélectionné par le niveau de déclenchement L2 est analysé en entier, et toutes les informations des sous-détecteurs associées à cet événement sont envoyées au troisième niveau. Le flux de sortie est d'environ 3,5 kHz. Les événements sont traités en parallèle avec un temps de décision de l'ordre de 40 ms.

Le troisième et dernier niveau de déclenchement est logiciel et est appelé EF pour *event filter*. Les niveaux de déclenchement logiciels L2 et EF sont appelés HLT, pour *High-Level Trigger*. Au niveau du déclenchement EF une reconstruction complète des événements est réalisée. Les événements sont filtrés afin de réduire le taux à environ 1000 Hz. Les événements sélectionnés sont enregistrés de manière définitive pour pouvoir être analysés hors ligne. Plusieurs déclenchements sont disponibles et peuvent être utilisés hors ligne afin d'obtenir des jeux de données spécifiques, par exemple les événements contenant au moins un muon ayant une impulsion transverse d'au moins 50 GeV. Un événement contient environ 1,3 MB d'informations, ainsi le flux de données en sortie du système de déclenchement est d'environ 1 GB s^{-1} . Les événements sont traités en parallèle avec un temps de décision de l'ordre de la seconde.

2.3 Reconstruction des objets physiques

La combinaison des informations de coups ou de dépôts d'énergie dans les sous-détecteurs d'ATLAS permet l'identification d'objets physiques et la mesure de leurs propriétés. Un objet reconstruit est donc un candidat pour un type de particule ou un jet, plus ou moins crédible selon la qualité d'identification choisie. Par exemple une trace dans le détecteur interne compatible à un dépôt d'énergie dans le calorimètre électromagnétique peut être identifié comme un candidat électron, e^+ ou e^- selon la courbure de sa trace, tandis qu'un même dépôt sans trace associée peut être identifié comme un candidat photon. Les objets reconstruits sont utilisés par les analyses de physique pour identifier par exemple un état final.

La reconstruction commence avec les traces dans le détecteur interne, identifiées à partir des coups dans ses différents sous-détecteurs [19]. Chaque combinaison de traces compatible avec une origine commune génère un point d'interaction probable appelé vertex primaire [20, 21]. Parmi les vertex primaires ainsi générés, on choisit le vertex dont la somme quadratique des traces lui étant associées est la plus élevée.

Dans cette section, les méthodes de reconstruction des principaux objets utilisés dans les analyses présentées dans les chapitres 5 et 6 sont décrites, ainsi que les principaux critères d'identification et d'isolation.

2.3.1 Les électrons

Les candidats électron sont reconstruits à partir de dépôts d'énergie calorimétriques correspondant à une trace dans le détecteur interne. Ici, le terme « électron » fait référence

à un électron e^- ou un positron e^+ , la charge étant déterminée à partir de la courbure de la trace associée à l'objet. Les gerbes électromagnétiques issues des électrons déposent de l'énergie dans les cellules du calorimètre électromagnétique. Les cellules sont regroupées en amas grâce à un algorithme à fenêtre glissante [22] cherchant à maximiser l'énergie dans une fenêtre de taille 3×5 cellules en $\eta \times \phi$. Seuls les amas ayant une énergie transverse supérieure à 2,5 GeV sont conservés. La trace extrapolée au niveau de l'amas la plus proche en $\Delta R = \sqrt{(\eta_{amas} - \eta_{trace})^2 + (\phi_{amas} - \phi_{trace})^2}$ est ensuite sélectionnée comme trace de l'électron. Les électrons sont identifiés dans la région $|\eta| < 2,5$ puisque les informations du détecteur interne sont requises. La taille typique du cône d'un électron reconstruit est $\Delta R = 0,2$

Plusieurs niveaux d'identification sont définis pour les électrons, ils correspondent à des valeurs différentes d'efficacité de sélection du signal électron et de rejet du bruit de fond. Le bruit de fond est composé notamment de pions, de photons, d'électrons ne provenant pas du vertex primaire, de jets qui déposent majoritairement leur énergie dans le calorimètre électromagnétique, ou encore d'électrons voyant leur charge mal identifiée. Ce dernier type de bruit de fond fait l'objet d'une méthode spécifique d'estimation à partir des données présentée dans la section 5.4.2. Les trois niveaux d'identification sont *loose*, *medium*, et *tight* [23]. Ils sont listés par ordre croissant de pureté c'est-à-dire de rejet de bruit de fond, et donc également par ordre décroissant d'efficacité de signal. Les électrons *tight* sont inclus parmi les électrons *medium* qui sont inclus parmi les électrons *loose*. Ces critères d'identification reposent sur la forme de la gerbe électromagnétique et la correspondance entre amas calorimétrique et trace interne. L'efficacité de sélection des électrons est montrée sur la figure 2.16 en fonction de leur énergie transverse et des différents critères d'identification. Elle augmente avec l'énergie transverse de l'électron, et varie entre environ 55 % et 95 % selon le critère d'identification et la valeur d'énergie transverse.

Un électron provenant de bosons W ou Z est typiquement éloigné des autres objets physiques. Ces bosons sont souvent des particules d'intérêt dans les analyses de physique, c'est pourquoi des critères d'isolation sont définis à partir des signaux situés autour de l'électron. Ces critères peuvent permettre d'éliminer certains faux électrons provenant d'une conversion de photon en paire e^+e^- , d'une désintégration d'un hadron léger, ou encore d'un dépôt d'énergie électromagnétique d'un hadron. L'isolation peut être définie à partir des traces internes proches d'un électron reconstruit, ou à partir des dépôts d'énergie calorimétriques autour de lui.

Dans les simulations Monte-Carlo, les événements sont pondérés afin de corriger les différences observées entre données réelles et simulation au niveau de l'efficacité de reconstruction des électrons, de leur identification, de leur isolation, et du déclenchement. Les observations sont réalisées sur des désintégrations $J/\psi \rightarrow e^+e^-$ et $Z \rightarrow e^+e^-$ avec la méthode *tag-and-probe* qui consiste à identifier l'un des deux électrons afin de sonder le

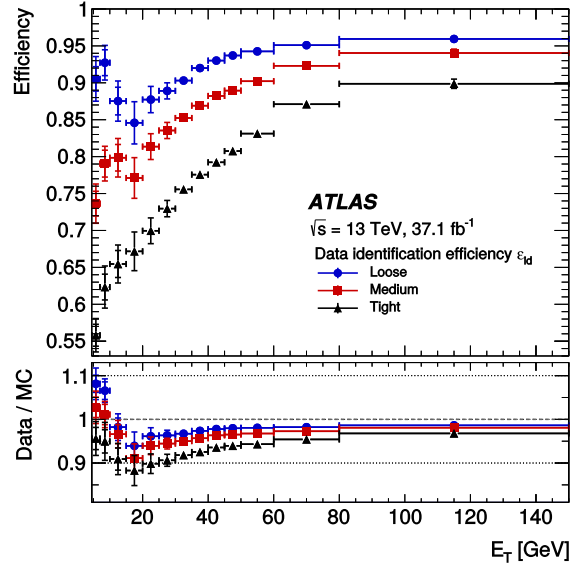


Figure 2.16 – Efficacité de sélection des électrons en fonction de leur énergie transverse et des différents critères d'identification. La mesure est effectuée sur les données prises en 2015 et en 2016 pour des événements $Z \rightarrow e^+e^-$. Le cadre en bas de la figure représente le rapport entre les efficacités mesurées dans les données et dans les simulations Monte-Carlo [23].

second. Les facteurs de correction ainsi obtenus dépendent des grandeurs η et E_T et des critères d'identification, ils se situent typiquement entre 2 et 10 %.

L'énergie de l'électron reconstruit est calculée à partir de l'énergie de l'amas calorimétrique correspondant, des propriétés de la gerbe électromagnétique et de la description précise de la matière en amont du calorimètre. Une calibration est nécessaire pour prendre en compte le débordement des gerbes en dehors de l'amas associé à l'électron reconstruit, ou en dehors du calorimètre, ou encore afin d'estimer l'énergie déposée avant l'entrée dans le calorimètre. Les échelles d'énergie des différentes couches du calorimètre électromagnétique sont ajustées entre elles, ainsi que l'échelle d'énergie globale grâce à des événements $Z \rightarrow e^+e^-$ [11].

2.3.2 Les muons

La reconstruction des candidats muon est réalisée à partir des informations de traces dans le détecteur interne ainsi que dans le spectromètre. Les muons sont d'abord reconstruits séparément dans le détecteur interne et dans le spectromètre puis un réajustement global de la trace dans les deux sous-détecteurs est effectué si cela est possible. Dans le détecteur interne les traces issues de muons sont reconstruites de manière standard, tandis que dans le spectromètre des segments de trace sont identifiés à partir des coups

dans les chambres MDT et CSC, puis les segments des différentes couches sont agrégés en traces.

Il existe quatre types de reconstruction des muons qui correspondent à différents niveaux de qualité de reconstruction [15].

Combined (CB) Un ré-ajustement global des coups dans le détecteur interne et dans le spectromètre est effectué. Le plus souvent, une trace reconstruite indépendamment dans le spectromètre est extrapolée au niveau du détecteur interne afin de la faire correspondre à une trace interne. Cette combinaison se fait avec un grand bras de levier, c'est pourquoi l'alignement des chambres du spectromètre doit être très bon.

Segment-tagged (ST) Une trace interne peut être reconstruite en tant que muon si son extrapolation au niveau du spectromètre est compatible avec au moins un segment de chambre MDT ou CSC. Ce type de reconstruction est utile pour les muons ne traversant qu'une couche du spectromètre, soit parce que leur impulsion transverse est faible, ou qu'ils traversent une région peu instrumentée du spectromètre.

Calorimeter-tagged (CT) Une trace interne peut être reconstruite en tant que muon si elle correspond à un dépôt d'énergie dans le calorimètre qui est compatible avec une particule au minimum d'ionisation. Il s'agit du type de reconstruction le moins fiable, mais il permet de reconstruire les muons traversant les régions très peu instrumentées dans lesquelles se situent des câbles par exemple. La reconstruction CT est optimisée les muons dans la région $|\eta| < 0,1$ et qui ont une impulsion transverse $15 < p_T < 100$ GeV.

Extrapolated (ME) Ce type de reconstruction n'utilise qu'une trace dans le spectromètre, qui doit être cohérente avec le vertex primaire. Au moins deux couches du spectromètre doivent avoir été touchées (trois dans les bouchons). Ce type de muons est principalement utilisé dans la région $2,5 < |\eta| < 2,7$ qui n'est pas couverte par le détecteur interne.

Les muons CB ont la plus grande pureté puisqu'ils sont reconstruits avec un maximum d'informations, toutefois les autres types sont utiles quand la reconstruction combinée n'est pas possible. Si plusieurs muons reconstruits se chevauchent, la préférence est donnée aux muons CB, puis ST, et enfin CT. Enfin les chevauchements restants avec les muons ME sont résolus en sélectionnant la trace de plus haute qualité et comportant le plus grand nombre de coups dans les sous-détecteurs.

Différents niveaux d'identification sont définis pour les muons : *loose*, *medium*, *tight*, *high- p_T* . Ils correspondent à différents compromis d'efficacité d'identification des muons, de mesure de leur impulsion, et de rejet du bruit de fond constitué principalement de muons secondaires provenant de désintégrations de pions et de kaons. La désintégration en vol de ces hadrons chargés présente une forte courbure au niveau du point où la désintégration a lieu, rendant incompatibles les impulsions mesurées séparément dans

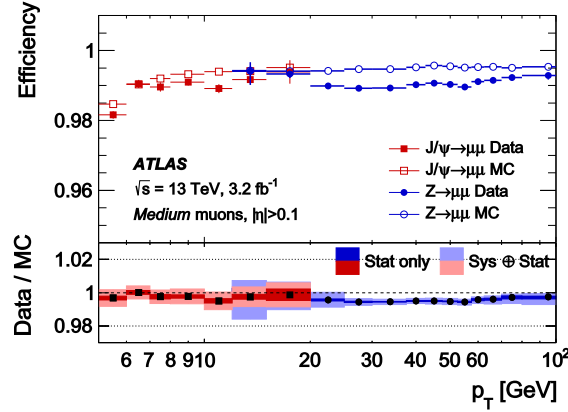


Figure 2.17 – Efficacité de sélection des muons *medium* en fonction de leur impulsion transverse. La mesure est effectuée sur les données prises en 2015 pour des événements $J/\psi \rightarrow \mu^+\mu^-$ et $Z \rightarrow \mu^+\mu^-$, dans la région $0,1 < |\eta| < 2,5$. Le cadre en bas de la figure représente le rapport entre les efficacités mesurées dans les données et dans les simulations Monte-Carlo [15].

le détecteur interne et dans le spectromètre. On s’attend donc à une faible qualité de trace globale reconstruite pour ces bruits de fond, comparé aux muons réels. Le point de fonctionnement standard dans ATLAS est *medium* puisqu’il minimise les incertitudes systématiques de reconstruction et de calibration. Seuls les muons CB et ME sont pris en compte. L’efficacité de sélection des muons *medium* est montrée sur la figure 2.17 en fonction de leur impulsion transverse. La mesure est réalisée grâce à la résonance $J/\psi \rightarrow \mu^+\mu^-$ à bas p_T et grâce à la résonance $Z \rightarrow \mu^+\mu^-$ à haut p_T . L’efficacité est supérieure à 98 %, et même supérieure à 99 % pour les muons vérifiant $p_T > 6$ GeV. Les efficacités sont stables pour des impulsions transverses allant de 6 à 100 GeV, et les valeurs issues des deux résonances sont compatibles dans leur région de recouvrement.

Plusieurs critères d’isolation sont également définis puisque les muons provenant de désintégrations de bosons W , Z , ou H sont typiquement isolés des autres particules. Les critères utilisent les informations de traces dans le détecteur interne ou de dépôts d’énergie calorimétriques autour du muon reconstruit.

Dans les simulations Monte-Carlo, les événements sont pondérés afin de corriger les différences observées entre données réelles et simulation au niveau de l’efficacité de reconstruction des muons, de leur identification, de leur isolation, et du déclenchement. La méthode est similaire à celle utilisée pour les électrons, c’est-à-dire qu’elle utilise les processus $J/\psi \rightarrow \mu^+\mu^-$ et $Z \rightarrow \mu^+\mu^-$ avec la méthode *tag-and-probe*. Les facteurs de correction ainsi obtenus dépendent des grandeurs η et E_T et des critères d’identification, il valent typiquement moins de 5 %, et s’élèvent à environ 10 % pour le muons ME à $|\eta| > 2,5$.

2.3.3 Les jets

La reconstruction des jets est très importante puisque beaucoup de particules hadroniques sont produites dans des collisions proton-proton. Des quarks ou gluons peuvent être émis puis s'hadroniser et ainsi générer un jet de particules collimatées. Ces particules doivent donc être reconstruites comme un unique objet physique. L'analyse d'un processus produisant plusieurs quarks dans l'état final, comme l'analyse présentée dans le chapitre 5, nécessite l'identification d'un grand nombre de jets.

Avant la reconstruction des jets à proprement parler, un premier regroupement de cellules du calorimètre est effectué afin d'identifier les amas d'énergie déposées par une potentielle gerbe hadronique. Les cellules ayant enregistré un dépôt d'énergie E supérieur à quatre déviations standard du bruit σ_{bruit} sont identifiées, et sont regroupées sous forme d'amas avec les cellules adjacentes vérifiant $E > 2\sigma_{\text{bruit}}$. Finalement, quand toutes les cellules voisines respectant la condition $E > 2\sigma_{\text{bruit}}$ ont été ajoutées à l'amas, les cellules formant le contour à l'extérieur de cet amas lui sont ajoutées sans condition sur E . Les amas de cellules ainsi générés sont appelés *TopoClusters* [22].

Plusieurs algorithmes permettent de rassembler les *TopoClusters* en véritables jets de manière itérative. L'algorithme standard utilisé dans ATLAS et donc dans cette thèse est l'algorithme anti- k_T [24]. Cet algorithme comporte un paramètre R représentant l'écart minimal entre deux jets distincts. On utilise souvent $R = 0,4$ dans ATLAS. L'algorithme est structuré en plusieurs étapes itératives.

- D'abord, les quantités $d_i = 1/p_{T,i}^2$ sont calculées pour chaque *TopoCluster* i , ainsi que les quantités $d_{i,j} = \min\left(1/p_{T,i}^2, 1/p_{T,j}^2\right) \times (\Delta R_{i,j}/R)$ pour chaque paire i, j de *TopoClusters*.
- Ensuite, on cherche le minimum parmi les d_i et les $d_{i,j}$.
- Si le minimum est un d_i , alors le *TopoCluster* correspondant est identifié en tant que jet et est supprimé de la liste de *TopoClusters*. Si le minimum est un $d_{i,j}$, alors les *TopoClusters* i et j sont agrégés, i et j sont supprimés de la liste et leur combinaison est ajoutée à la liste.
- Les grandeurs d_i et $d_{i,j}$ sont recalculées et le processus recommence jusqu'à ce que la liste soit vide.

L'intérêt d'un tel algorithme est que les *TopoClusters* les plus énergétiques sont agrégés en priorité. En effet la grandeur $d_{i,j}$ est élevée pour une paire d'amas peu énergétiques et faible pour une paire contenant au moins un amas très énergétique. Ainsi les dépôts les moins énergétiques sont agrégés aux dépôts les plus énergétiques, ce qui rend l'algorithme peu sensible aux émissions de partons à basse énergie. Tous les dépôts séparés d'une distance angulaire inférieure à $R = 0,4$ sont agrégés à la fin de l'algorithme, qui est donc également robuste face aux émissions de partons colinéaires. Ces deux caractéristiques sont intéressantes pour les comparaisons avec des calculs théoriques qui sont sensibles aux émissions de partons colinéaires et infrarouges c'est-à-dire à basse énergie.

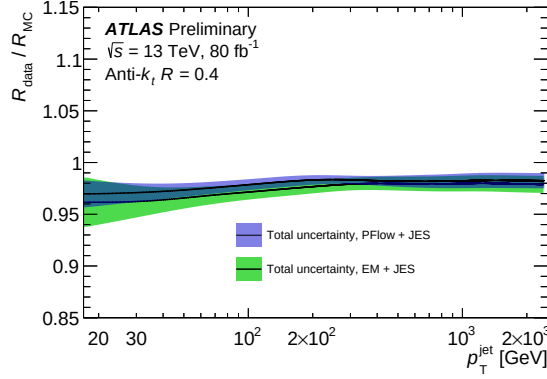


Figure 2.18 – Comparaison de l’échelle d’énergie des jets mesurées sur les données prises entre 2015 et 2017 et sur des simulations Monte-Carlo, en fonction de l’impulsion transverse des jets [26, 27]. Les jets sont reconstruits avec l’algorithme anti- k_T et un paramètre de cône $R = 0,4$, à partir de *TopoClusters* pour les courbes vertes ou via l’algorithme *particle flow* pour les courbes bleues [14].

Des coupures de qualité sont appliquées aux jets ainsi reconstruits car de nombreux phénomènes physiques peuvent générer de signaux ressemblant à des jets [25]. Une calibration est nécessaire pour prendre en compte le débordement des gerbes en dehors du cône associé au jet reconstruit, ou en dehors du calorimètre, ou encore afin d’estimer l’énergie déposée avant l’entrée dans le calorimètre. Des corrections sur l’échelle d’énergie des jets ainsi que leur résolution en énergie sont finalement appliquées. La figure 2.18 montre la comparaison de l’échelle d’énergie des jets mesurées sur les données prises entre 2015 et 2017 et sur des simulations Monte-Carlo, en fonction de l’impulsion transverse des jets.

Une observable appelée JVT pour *jet-vertex-tagger* [28] est définie à partir de la proportion d’impulsion transverse dans le jet qui provient du vertex primaire. Une coupure sur cette variable permet d’éliminer des jets issus de l’empilement. Une valeur de JVT proche de 0 correspond à un jet d’empilement, tandis qu’une valeur proche de 1 correspond à un jet provenant de l’interaction d’intérêt.

2.3.4 Les jets de quarks b

L’étude des quarks b est particulièrement importante pour les analyses de physique s’intéressant au quark top, puisque ce dernier n’est pas observable directement dans le détecteur et se désintègre quasi-exclusivement en quark b et en boson W . Les quarks b générés lors d’une interaction s’hadronisent en hadrons dont le temps de vie leur permet de se déplacer sur quelques mm avant de se désintégrer. Cette distance n’est pas suffisante pour permettre une mesure dans le détecteur interne, cependant un vertex secondaire

déplacé par rapport au vertex primaire peut être reconstruit. Ainsi, les jets issus de quarks b , appelés jets de quarks b , peuvent être identifiés. Des algorithmes d'identification de jets de quarks b , ou *b-tagging*, sont mis en place afin de séparer les jets provenant d'un quark b et les jets provenant de quarks plus légers [29]. En particulier, ces algorithmes doivent rejeter les jets de quarks c qui peuvent également générer un vertex secondaire déplacé mais à une distance typiquement inférieure comparé aux jets de quarks b .

L'identification des jets de quarks b dans le détecteur ATLAS est fondée sur trois algorithmes [30, 31]. L'algorithme IP3D s'intéresse aux déplacements des composants du jet par rapport au vertex primaire, tandis que les algorithmes SV et JetFitter reconstruisent un vertex secondaire déplacé. Dans la suite de cette section et sauf mention contraire, le terme « léger » qualifie un jet qui n'est ni un jet de quark b , ni un jet de quark c .

L'algorithme IP3D, pour *Impact Parameter 3D*, utilise les informations de paramètre d'impact pour distinguer les jets de quark b des autres jets. Le paramètre d'impact dans le plan transverse est appelé d_0 , et le paramètre d'impact dans le plan longitudinal z_0 . On utilise sa projection suivant l'axe z : $z_0 \sin \theta$. Les grandeurs d_0 et $z_0 \sin \theta$ sont plus élevées pour un jet de quark b que pour un jet de quark c ou un jet léger. De plus ces grandeurs sont corrélées pour un jet de quark b . C'est pourquoi l'algorithme utilise les grandeurs d_0/σ_{d_0} et $z_0 \sin \theta/\sigma_{z_0 \sin \theta}$ ainsi que leur corrélation.

L'algorithme SV, pour *Secondary Vertex*, reconstruit un vertex secondaire. Tous les vertex secondaires possibles sont générés à partir de toutes les paires de traces dans le jet. Des coupures sur la qualité des traces ainsi que sur les vertex générés permettent d'éliminer plusieurs possibilités. Notamment, les paires de deux traces dont la masse invariante est compatible avec un photon ou un hadron léger sont ignorées. S'il reste au moins une paire de traces, toutes les traces restantes sont combinées en un unique vertex puis une procédure itérative élimine les traces les moins compatibles.

Un quark b se désintègre principalement en quark c qui peut lui aussi produire un vertex déplacé, c'est pourquoi on attend un vertex tertiaire déplacé du vertex secondaire dû au hadron b . L'algorithme JetFitter reconstruit toute la chaîne de désintégration du hadron b , c'est-à-dire le vertex secondaire ainsi que le vertex tertiaire. La ligne de vol des hadrons b et c le long de laquelle les vertex primaire, secondaire et tertiaire sont alignés est identifiée, ainsi que la position des vertex sur cette ligne.

Finalement, les informations de sortie de ces trois algorithmes sont combinés via un classificateur appelé MV2 utilisant un arbre de décision boosté afin d'obtenir une meilleure discrimination que l'un des trois algorithmes seul. L'arbre de décision est entraîné sur des jets provenant d'événements simulés $pp \rightarrow t\bar{t}$. Il existe plusieurs versions de l'algorithme MV2 qui utilisent différentes proportions de jets de quarks c et de jets légers dans l'échantillon d'entraînement. Les variantes MV2c00, MV2c10 et MV2c20 utilisent respectivement 0, 7 et 15 % de jets de quarks c complétés par des jets légers.

Efficacité [%]	Facteur de rejet de jets de c	Facteur de rejet de jets légers
60	34	1540
70	12	380
77	6	130
85	3,1	33

Table 2.1 – Efficacité d’identification des jets de quarks b et facteurs de rejet des jets de quarks c et des jets légers pour les quatre points de fonctionnement de l’algorithme MV2c10 [31]. Ces valeurs sont calculées sur des événements simulés $pp \rightarrow t\bar{t}$ pour des jets vérifiant p_T .

L’algorithme MV2c20 obtient le meilleur facteur de rejet des jets de quarks c , mais celui obtenu par MV2c10 est très proche. De plus MV2c10 obtient un facteur de rejet des quarks légers significativement meilleur que MV2c20, c’est pourquoi la version de l’algorithme la plus utilisée est MV2c10. D’autres classificateurs alternatifs sont disponibles, utilisant par exemple des techniques d’apprentissage profond. Le classificateur fournit un score pour chaque jet qui permet son identification comme jet de quark b . Quatre points de fonctionnement sont définis à partir d’un seuil sur ce score, c’est-à-dire que tout jet de score supérieur au seuil est identifié en tant que jet de quark b . Les quatre points de fonctionnement sont définis selon l’efficacité d’identification des jets de quarks b obtenue. Le tableau 2.1 montre les efficacités d’identification de jets de quarks b pour ces quatre points de fonctionnement, ainsi que les facteurs de rejet de jets de quarks c et de jets légers associés, un facteur de rejet étant défini comme l’inverse d’une efficacité. Par exemple, le point de fonctionnement utilisé dans l’analyse présentée dans le chapitre 5 correspond à une efficacité de 77 %, c’est-à-dire que 77 % des jets de quarks b réels sont effectivement identifiés par l’algorithme. Pour ce point de fonctionnement, le facteur de rejet des jets de quarks c est d’environ 6 ce qui signifie qu’environ $1/6 \simeq 17\%$ des jets de quarks c sont identifiés à tort en tant que jets de quarks b . En ce qui concerne les jets légers, le facteur de rejet est d’environ 134, soit une probabilité de $1/134 \simeq 0,7\%$ d’identifier un jet léger en tant que jet de quark b .

2.3.5 L’énergie transverse manquante

Les neutrinos produits dans le détecteur ATLAS le traversent sans interagir. Afin d’estimer l’impulsion des neutrinos ou d’éventuelles autres particules n’interagissant pas dans le détecteur, l’énergie transverse manquante E_T^{miss} est construite.

Les deux protons entrant en collision se déplacent suivant l’axe du faisceau, ainsi

l'impulsion transverse avant collision est nulle :

$$\sum_{\text{particules}} \vec{p}_T = \vec{0}, \quad (2.10)$$

où la somme est effectuée sur tous les objets reconstruits dans l'événement. Si la somme des impulsions transverses des objets reconstruits n'est pas nulle, alors une partie de l'impulsion n'a pas été reconstruite. Il s'agit du terme dur de l'impulsion transverse manquante :

$$E_T^{\text{miss, hard}} = - \sum_{\text{objets}} \vec{p}_T. \quad (2.11)$$

Cette grandeur correspond à la somme vectorielle des impulsions qui n'ont pas été incluses dans des objets physiques. À ce terme s'ajoute le terme doux $E_T^{\text{miss, soft}}$ qui correspond aux signaux qui n'ont été associés à aucun objet. Il existe une version du terme doux utilisant les traces dans le détecteur interne associées au vertex primaire, et une seconde version utilisant les *TopoClusters* reconstruits dans le calorimètre. Le terme doux calorimétrique est toutefois sensible à l'empilement et moins performant que le terme doux basé sur les traces. C'est donc ce dernier qui est utilisé [32].

2.3.6 Effet de l'empilement

L'effet d'empilement, c'est-à-dire les collisions multiples dans un même croisement de faisceau, est très important comme montré sur la figure 2.5. Cet effet implique qu'il existe plusieurs vertex primaires dans un même événement. L'empilement génère des traces et des dépôts d'énergie parasites dans les sous-détecteurs, ce qui dégrade la reconstruction, la résolution des détecteurs, ou encore le calcul de l'énergie transverse manquante.

Dans les simulations Monte-Carlo, l'empilement est simulé avec PYTHIA et ajouté aux événements générés. Les événements résultant sont ensuite pondérés selon leur nombre de vertex primaires afin d'obtenir une bonne correspondance des distributions d'empilement μ entre données réelles et simulation.

2.3.7 Suppression des chevauchements

Une particule peut être reconstruite comme plusieurs objets à la fois, par exemple un électron est également reconstruit comme un jet. Il est donc nécessaire d'effectuer une suppression des objets trop proches entre eux selon des critères séquentiels, comme détaillé par exemple dans la section 5.2.

2.4 Les simulations

Le principe d'une analyse de physique centrée sur la recherche d'un nouveau phénomène est de simuler le signal recherché ainsi que les bruits de fond afin de les estimer. Ces

estimations sont ensuite comparées aux données récoltées par le détecteur ATLAS, et une analyse statistique est menée afin de calculer la vraisemblance de certaines hypothèses comme la présence d'un signal. Une analyse ayant pour but une mesure de précision d'un paramètre du modèle standard doit également confronter les observations expérimentales aux prédictions théoriques du modèle, qui sont simulées. La simulation est donc une étape cruciale de toute analyse de physique. Les méthodes de simulation Monte-Carlo (MC) sont utilisées via des programmes dédiés, appelés « générateurs », afin de simuler les phénomènes physiques événement par événement.

La première étape de la simulation MC d'un événement est la génération de particules issues de la diffusion inélastique proton-proton, simulé à partir du calcul de l'élément de matrice correspondant. Les restes des deux protons de la collision qui n'ont pas participé à l'interaction principale sont également pris en compte, on parle d'événement sous-jacent. Les partons dans l'événement sous-jacent ont typiquement une faible impulsion transverse et un faible angle par rapport à l'axe du faisceau. À ce stade de la simulation, une liste de particules produites par la collision proton-proton simulée est disponible. Les quarks et les gluons n'ont notamment pas encore subi d'étape d'hadronisation et sont enregistrés tels quels. Ces informations sont appelées informations « vraies », ou *truth*, on parle aussi d'informations au niveau partonique. Le calcul de l'élément de matrice est réalisé à un certain ordre en QED et QCD, par exemple à l'ordre dominant (LO) ou au second ordre dominant (NLO). Des radiations dans l'état initial ou final peuvent être également simulées. Une discussion plus détaillée de ces aspects au sujet du processus de production de quatre quarks top est disponible dans la section 4.1.1. Il est important de noter que la simulation est réalisée pour une collision proton-proton au LHC et que par conséquent elle nécessite la connaissance des fonctions de distribution de partons (PDF) dans le proton.

La seconde étape est l'étape de gerbe partonique, au cours de laquelle tous les partons (quarks et gluons) peuvent produire d'autres partons en cascade. Par exemple un gluon peut donner naissance à une paire quark-antiquark, ou encore un quark peut émettre un gluon. La gerbe partonique est supposée indépendante du processus physique initialement simulé, c'est l'hypothèse de factorisation. La séparation des deux régimes est effectuée via l'introduction d'une échelle d'énergie dite échelle de factorisation μ_F qui introduit une incertitude théorique. La gerbe partonique permet de simuler des radiations de partons, et ceci même pour un processus simulé à un ordre fixe LO ou NLO par exemple. Les partons sont finalement confinés dans des hadrons de couleur neutre, c'est l'étape d'hadronisation. Les hadrons ainsi générés peuvent se désintégrer avant d'atteindre le détecteur.

Ces deux premières étapes peuvent être menées par un même générateur MC ou par deux générateurs différents. Les générateurs MC principalement utilisés dans les analyses présentées dans les chapitres 5 et 6 sont MADGRAPH5_AMC@NLO [33, 34] pour la génération d'événements et PYTHIA [35] pour l'hadronisation. Certains processus

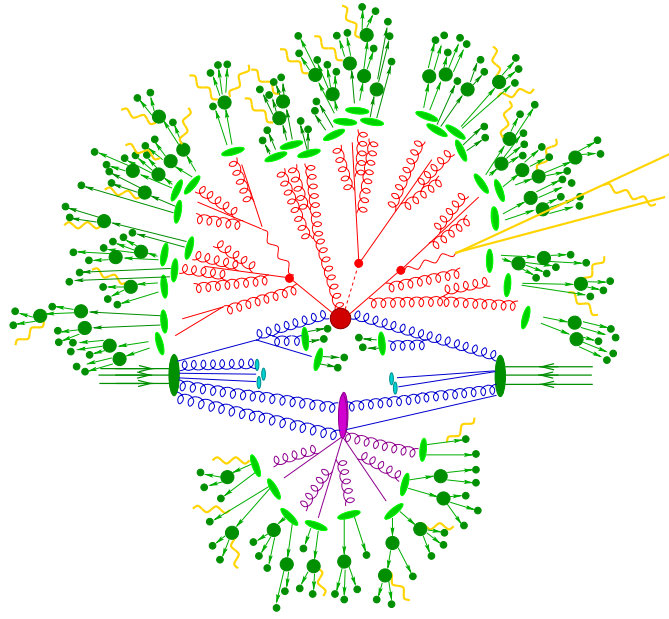


Figure 2.19 – Schéma représentant une collision proton-proton simulée par un générateur Monte-Carlo [37]. Les groupes de trois flèches à droite et à gauche de la figure représentent les deux protons entrant en collision. Le disque rouge central représente le processus de diffusion inélastique simulé, entouré de branches rouges représentant les partons générés dans la gerbe partonique. Les structures violettes représentent une diffusion inélastique secondaire. Les éléments ovales verts représentent l’hadronisation des partons, puis les disques verts les plus grands devenant des disques plus petits représentent la désintégration des hadrons. Enfin les lignes jaunes représentent les particules électromagnétiques émises dans le processus.

secondaires sont simulés entièrement avec PYTHIA, et d’autres avec SHERPA [36]. Une représentation graphique des différentes étapes de la simulation est montrée sur la figure 2.19.

Le trajet des particules créées par le générateur à travers le détecteur ATLAS est ensuite simulé. Cette étape se base sur une simulation du détecteur ATLAS réalisée avec le programme GEANT4 [38] au sein de l’infrastructure de simulation d’ATLAS [39]. Les interactions des particules dans la matière des sous-détecteurs sont simulées à partir de sections efficaces connues ou de modèles phénoménologiques calibrés sur des données expérimentales. Les coups et dépôts d’énergie dans tous les sous-détecteurs sont simulés, ainsi que les signaux en sortie de l’électronique de lecture et l’étape de numérisation, en prenant en compte le bruit de l’électronique. Finalement une reconstruction des objets est effectuée à partir de ces informations, de manière identique à la reconstruction réalisée sur les données réelles de collision. Ainsi les objets et les formats de données après

reconstruction sont identiques pour les simulations MC et pour les données prises par le détecteur, et on peut analyser les simulations comme on le ferait avec des données de collision.

La simulation du détecteur ne peut pas être parfaite. La géométrie des composants présente des imprécisions, notamment les câbles et autres services sont difficiles à inclure en totalité. De plus les interactions à très basse impulsion sont difficiles à modéliser. Ainsi les prédictions du nombre d'événements obtenus par simulation Monte-Carlo sont généralement supérieures aux nombres d'événements réellement observés dans les données de collision. Comme expliqué dans la section 2.3, des corrections de normalisation ou bien de forme de certaines distributions sont appliquées aux événements simulés, afin de corriger les désaccords observés entre simulation et données dans les efficacités de sélection, les efficacités d'identification des objets ainsi que leur résolution.

La simulation complète du détecteur ATLAS via GEANT4, appelée *full simulation*, est cependant très gourmande en temps de calcul, de l'ordre de quelques minutes par événement. La simulation du calorimètre est responsable d'environ 90 % de ce temps de calcul, dont environ 75 % uniquement pour la simulation des gerbes électromagnétiques. C'est pourquoi un mode de simulation approchée mais plus rapide ATLFAST-II est disponible, dans lequel on utilise une simulation du calorimètre simplifiée appelée FastCaloSim [40], ce qui permet de gagner un facteur 10 sur le temps de calcul.

Les événements MC simulés puis reconstruits sont générés en grand nombre, et stockés dans des fichiers appelés échantillons MC. Lors de l'utilisation d'échantillons MC dans une analyse de physique, il est nécessaire de normaliser le nombre d'événements sélectionnés conformément à la section efficace du processus correspondant, ainsi que la luminosité intégrée des données utilisées. Ainsi la prédiction obtenue est bien comparable au processus réel tel qu'il pourrait être observé dans les données de collision.

Références

- [1] Lyndon EVANS et Philip BRYANT. “LHC Machine”. In : *Journal of Instrumentation* 3.08 (août 2008), S08001-S08001. DOI : 10.1088/1748-0221/3/08/s08001.
- [2] ATLAS COLLABORATION. “The ATLAS Experiment at the CERN Large Hadron Collider”. In : *Journal of Instrumentation* 3.08 (août 2008), S08003-S08003. DOI : 10.1088/1748-0221/3/08/s08003.
- [3] CMS COLLABORATION. “The CMS experiment at the CERN LHC”. In : *Journal of Instrumentation* 3.08 (août 2008), S08004-S08004. DOI : 10.1088/1748-0221/3/08/s08004.
- [4] LHCb COLLABORATION. “The LHCb Detector at the LHC”. In : *Journal of Instrumentation* 3.08 (août 2008), S08005-S08005. DOI : 10.1088/1748-0221/3/08/s08005.
- [5] ALICE COLLABORATION. “The ALICE experiment at the CERN LHC”. In : *Journal of Instrumentation* 3.08 (août 2008), S08002-S08002. DOI : 10.1088/1748-0221/3/08/s08002.
- [6] CERN’S ACCELERATOR COMPLEX. URL : http://cern60-downloads.web.cern.ch/cern60-downloads/downloads/EXHIBITION%20CONTENT/3_LHC_Accelerators/3_2_Posters/3_2_1_Overview/CERN_Accelerators/3_PDF-text/CERN_Accelerators.pdf.
- [7] G APOLLINARI et al. *High-Luminosity Large Hadron Collider (HL-LHC) Preliminary Design Report*. 2015. DOI : 10.5170/cern-2015-005.
- [8] ATLAS COLLABORATION. “The ATLAS Inner Detector commissioning and calibration”. In : *The European Physical Journal C* 70.3 (août 2010), p. 787-821. DOI : 10.1140/epjc/s10052-010-1366-7.
- [9] M CAPEANS et al. *ATLAS Insertable B-Layer Technical Design Report*. Rapp. tech. CERN-LHCC-2010-013. ATLAS-TDR-19. Sept. 2010. URL : <https://cds.cern.ch/record/1291633>.
- [10] ATLAS COLLABORATION. *Track Reconstruction Performance of the ATLAS Inner Detector at $\sqrt{s} = 13$ TeV*. Rapp. tech. ATL-PHYS-PUB-2015-018. Geneva : CERN, juil. 2015. URL : <https://cds.cern.ch/record/2037683>.
- [11] ATLAS COLLABORATION. “Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data”. In : *Journal of Instrumentation* 14.03 (mar. 2019), P03017-P03017. DOI : 10.1088/1748-0221/14/03/p03017.
- [12] ATLAS COLLABORATION. *Jet energy resolution in 2017 data and simulation*. URL : <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-005/>.

- [13] ATLAS COLLABORATION. *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*. Rapp. tech. ATLAS-CONF-2015-037. Geneva : CERN, août 2015. URL : <https://cds.cern.ch/record/2044941>.
- [14] ATLAS COLLABORATION. “Jet reconstruction and performance using particle flow with the ATLAS Detector”. In : *The European Physical Journal C* 77.7 (juil. 2017). DOI : 10.1140/epjc/s10052-017-5031-2.
- [15] ATLAS COLLABORATION. “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV”. In : *The European Physical Journal C* 76.5 (mai 2016). DOI : 10.1140/epjc/s10052-016-4120-y.
- [16] ATLAS COLLABORATION. “Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC”. In : *The European Physical Journal C* 76.12 (nov. 2016). DOI : 10.1140/epjc/s10052-016-4466-1.
- [17] G. AVONI et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In : *Journal of Instrumentation* 13.07 (juil. 2018), P07017-P07017. DOI : 10.1088/1748-0221/13/07/p07017.
- [18] ATLAS COLLABORATION. “Performance of the ATLAS trigger system in 2015”. In : *The European Physical Journal C* 77.5 (mai 2017). DOI : 10.1140/epjc/s10052-017-4852-3.
- [19] ATLAS COLLABORATION. “Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2”. In : *The European Physical Journal C* 77.10 (oct. 2017). DOI : 10.1140/epjc/s10052-017-5225-7.
- [20] ATLAS COLLABORATION. “Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton-proton collisions at the LHC”. In : *The European Physical Journal C* 77.5 (mai 2017). DOI : 10.1140/epjc/s10052-017-4887-5.
- [21] ATLAS COLLABORATION. *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*. Rapp. tech. ATL-PHYS-PUB-2015-026. Geneva : CERN, juil. 2015. URL : <https://cds.cern.ch/record/2037717>.
- [22] W LAMPL et al. *Calorimeter Clustering Algorithms : Description and Performance*. Rapp. tech. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva : CERN, avr. 2008. URL : <https://cds.cern.ch/record/1099735>.
- [23] ATLAS COLLABORATION. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV”. In : (11 fév. 2019). arXiv : 1902.04655v1 [physics.ins-det].
- [24] Matteo CACCIARI, Gavin P SALAM et Gregory SOYEZ. “The anti-kt jet clustering algorithm”. In : *Journal of High Energy Physics* 2008.04 (avr. 2008), p. 063-063. DOI : 10.1088/1126-6708/2008/04/063.

-
- [25] ATLAS COLLABORATION. *Selection of jets produced in proton-proton collisions with the ATLAS detector using 2011 data*. Rapp. tech. ATLAS-CONF-2012-020. Geneva : CERN, mar. 2012. URL : <https://cds.cern.ch/record/1430034>.
 - [26] ATLAS COLLABORATION. *Jet energy scale and uncertainties in 2015-2017 data and simulation*. URL : <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-006/>.
 - [27] ATLAS COLLABORATION. “Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Physical Review D* 96.7 (oct. 2017). DOI : 10.1103/physrevd.96.072002.
 - [28] ATLAS COLLABORATION. *Tagging and suppression of pileup jets with the ATLAS detector*. Rapp. tech. ATLAS-CONF-2014-018. Geneva : CERN, mai 2014. URL : <https://cds.cern.ch/record/1700870>.
 - [29] ATLAS COLLABORATION. “Performance of b -jet identification in the ATLAS experiment”. In : *Journal of Instrumentation* 11.04 (avr. 2016), P04008-P04008. DOI : 10.1088/1748-0221/11/04/p04008.
 - [30] ATLAS COLLABORATION. *Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data*. Rapp. tech. ATLAS-CONF-2011-102. Geneva : CERN, juil. 2011. URL : <https://cds.cern.ch/record/1369219>.
 - [31] ATLAS COLLABORATION. *Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run*. Rapp. tech. ATL-PHYS-PUB-2016-012. Geneva : CERN, juin 2016. URL : <https://cds.cern.ch/record/2160731>.
 - [32] ATLAS COLLABORATION. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV”. In : *The European Physical Journal C* 78.11 (nov. 2018). DOI : 10.1140/epjc/s10052-018-6288-9.
 - [33] Stefano FRIXIONE, Paolo NASON et Bryan R WEBBER. “Matching NLO QCD and parton showers in heavy flavour production”. In : *Journal of High Energy Physics* 2003.08 (août 2003), p. 007-007. DOI : 10.1088/1126-6708/2003/08/007.
 - [34] J. ALWALL et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In : *Journal of High Energy Physics* 2014.7 (juil. 2014). DOI : 10.1007/jhep07(2014)079.
 - [35] Torbjörn SJÖSTRAND et al. “An introduction to PYTHIA 8.2”. In : *Computer Physics Communications* 191 (juin 2015), p. 159-177. DOI : 10.1016/j.cpc.2015.01.024.

- [36] T GLEISBERG et al. “Event generation with SHERPA 1.1”. In : *Journal of High Energy Physics* 2009.02 (fév. 2009), p. 007-007. DOI : 10.1088/1126-6708/2009/02/007.
- [37] Stefan HÖCHE. “Introduction to Parton-Shower Event Generators”. In : *Journeys Through the Precision Frontier : Amplitudes for Colliders*. WORLD SCIENTIFIC, sept. 2015. DOI : 10.1142/9789814678766_0005. URL : <https://arxiv.org/abs/1411.4085>.
- [38] S. AGOSTINELLI et al. “Geant4-a simulation toolkit”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (juil. 2003), p. 250-303. DOI : 10.1016/s0168-9002(03)01368-8.
- [39] ATLAS COLLABORATION. “The ATLAS Simulation Infrastructure”. In : *The European Physical Journal C* 70.3 (sept. 2010), p. 823-874. DOI : 10.1140/epjc/s10052-010-1429-9.
- [40] Takashi Yamanaka AND. “The ATLAS calorimeter simulation FastCaloSim”. In : *Journal of Physics : Conference Series* 331.3 (déc. 2011), p. 032053. DOI : 10.1088/1742-6596/331/3/032053.

3

Caractérisation de grands détecteurs Micromegas pour le projet *New Small Wheel*

3.1 Introduction

L'amélioration du détecteur ATLAS pour la phase haute luminosité du LHC ou « HL-LHC » [1] devrait permettre d'obtenir de nombreux résultats de physique mais sa mise en place présente certains défis. L'augmentation de la luminosité va en effet de pair avec l'augmentation de l'irradiation des détecteurs, et l'augmentation globale du nombre d'interactions par croisement de faisceaux du LHC.

Plusieurs parties du spectromètre à muons doivent être remplacées, notamment les détecteurs vers l'avant les plus proches du point d'interaction, appelés *Small Wheels* ou « petites roues ». Ils devront supporter une irradiation pouvant aller jusqu'à 15 kHz cm^{-2} [2], ce qui n'est pas envisageable pour les détecteurs en place, basés sur la technologie des chambres à pistes cathodiques (CSC). Les signaux interprétés à tort comme des muons seraient notamment beaucoup trop fréquents avec l'équipement actuel. Le taux de déclenchement serait de l'ordre de 60 kHz soit trois fois plus que le taux cible d'environ 20 kHz [2].

Les nouveaux détecteurs qui seront installés au niveau des petites roues en remplacement des détecteurs existants sont appelés *New Small Wheel* (NSW). Leur position en termes de pseudo-rapacité est $1,3 < |\eta| < 2,7$. La figure 3.1 illustre leur disposition au sein du détecteur ATLAS. Leur installation est prévue en 2019 et 2020.

Cette mise à niveau des petites roues doit permettre de maintenir voire même d'améliorer les performances de reconstruction des muons dans l'environnement très irradié du HL-LHC. L'analyse hors ligne requiert une résolution spatiale d'environ $100 \mu\text{m}$ pour la reconstruction de traces, essentielle pour maintenir la précision actuelle sur la résolution de l'impulsion des muons. Les segments de traces utilisés par le système de déclenchement L1 quant à eux doivent être reconstruits en ligne, avec une résolution angulaire d'environ 1 mrad, afin de rejeter les signaux de déclenchement ne provenant pas du point d'interaction [2]. La figure 3.2 illustre ces performances pour les petites roues actuelles « SW » dans différentes conditions de luminosité instantanée, ainsi que pour les nouvelles roues NSW.

La figure 3.2 (a) à gauche montre l'efficacité de reconstruction de segments de traces

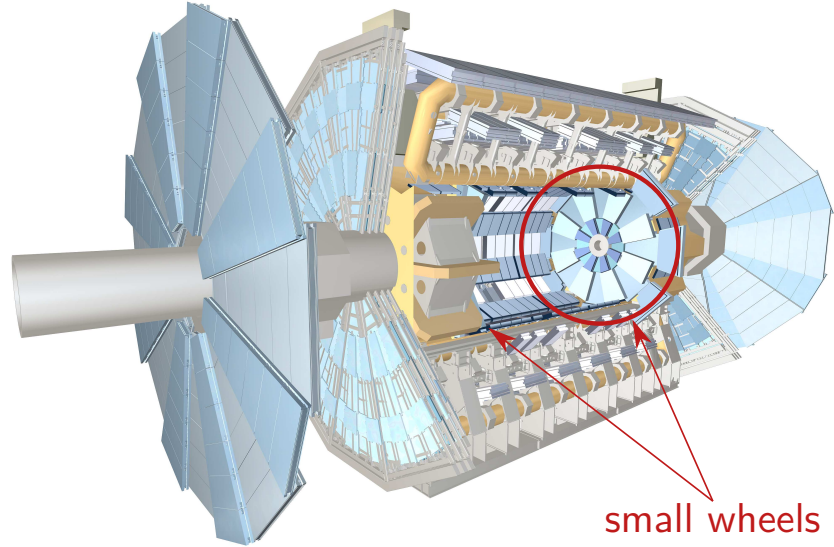


Figure 3.1 – Vue en coupe des détecteurs à muons au sein du détecteur ATLAS. La zone entourée en rouge et désignée par une flèche correspond à une petite roue (ou *Small Wheel*). La seconde flèche montre la deuxième roue, symétrique de la première par rapport au point d’interaction au centre du détecteur. Elle n’est pas visible sur cette image.

en fonction de leur position radiale, pour des événements simulés $Z' \rightarrow \mu^+ \mu^-$ avec une masse du boson hypothétique Z' égale à 2 TeV. Ce modèle est un signal typique de physique au-delà du modèle standard. On peut notamment observer les points bleus (resp. rouges) représentant l’efficacité sur des événements simulés de la petite roue NSW (resp. SW) à une luminosité instantanée typique des conditions HL-LHC. L’efficacité NSW est bien meilleure que l’efficacité avec les roues standard SW dans ces conditions, et même meilleure que l’efficacité actuelle des petites roues (en noir)

La figure 3.2 (b) à droite présente le taux de reconstruction de « faux » segments, c’est-à-dire n’étant pas associés à un muon réel, en fonction de leur position radiale. Ce taux dans le cas NSW aux conditions HL-LHC (en bleu) est bien inférieur au taux SW dans les mêmes conditions (en rouge), ce qui est un très bon point. Il est même comparable au taux SW dans les conditions actuelles (en noir). Les taux en bleu et en noir sont situés à l’extrémité basse de la figure, sous l’ordonnée 0,5.

Le projet NSW repose sur l’utilisation de deux technologies de détecteurs gazeux : les *small-strip Thin Gap Chambers* ou sTGC [3, 4], et les « Micromegas » ou MM. Un détecteur sTGC est constitué de fils parallèles disposés entre deux cathodes ayant une très grande résistivité. Derrière les cathodes se trouvent des pistes perpendiculaires aux fils d’un côté, et des électrodes de l’autre côté. Il est opéré en mode saturé. La description

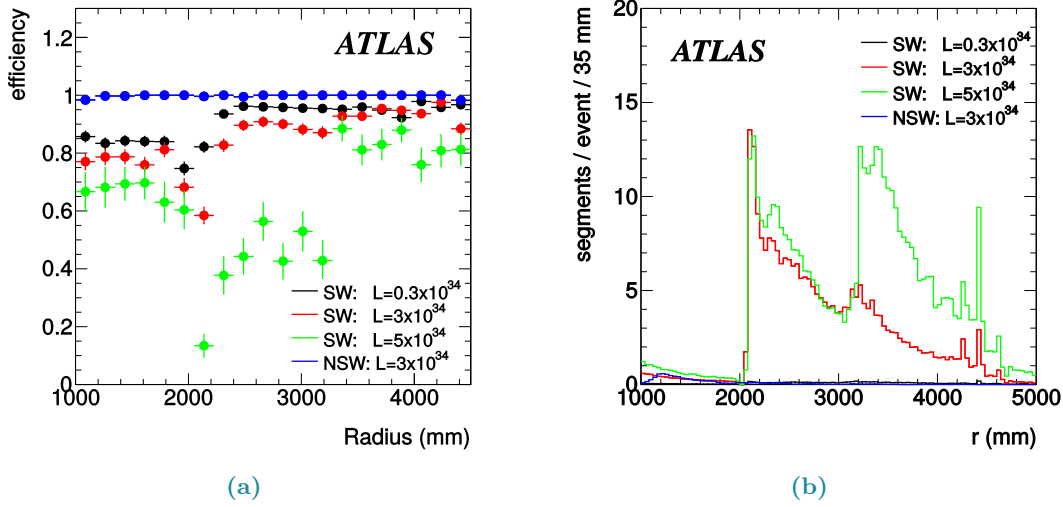


Figure 3.2 – (a) Efficacité de reconstruction de segments de traces en fonction de leur position radiale, pour des événements simulés $Z' \rightarrow \mu^+ \mu^-$ avec une masse du boson hypothétique Z' égale à 2 TeV [2]. La couleur bleue correspond aux petites roues NSW, les autres couleurs aux roues actuelles. Le noir représente les conditions actuelles en termes de luminosité instantanée, et les couleurs rouge et verte des conditions possibles pour le HL-LHC. (b) Taux de reconstruction de « faux » segments, c'est-à-dire n'étant pas associés à un muon réel, en fonction de leur position radiale [2]. Mêmes couleurs que pour (a).

de la technologie MM est réalisée dans la section 3.2.1.

Les sTGC sont principalement utilisés pour le déclenchement Niveau-1 et les Micromegas pour reconstruire les muons hors ligne avec leur très bonne précision de reconstruction des traces. La combinaison de ces deux technologies permet donc de la redondance et de très bonnes performances.

Deux roues de détection vont donc être construites, en remplacement des petites roues actuelles du détecteur ATLAS. Chaque roue se compose de seize secteurs : huit grands secteurs et huit petits secteurs. Chacun de ces secteurs est constitué de quatre multiplets de détection, dans l'ordre suivant : sTGC - MM - MM - sTGC. Chaque multiplet, sTGC ou MM, contient quatre plans de détection sTGC ou MM respectivement. Cette disposition est schématisée sur les figures 3.3 et 3.4.

Ces figures montrent que chaque quadruplet Micromegas est divisé en deux parties, notées M1 et M2. On a donc quatre types de quadruplets MM différents : SM1, SM2, LM1 et LM2. Le « S » ou le « L » au début du nom désigne une partie d'un petit (*Small*) ou d'un grand (*Large*) secteur.

Les multiplets sTGC sont disposés à l'extérieur des secteurs afin d'avoir le plus grand

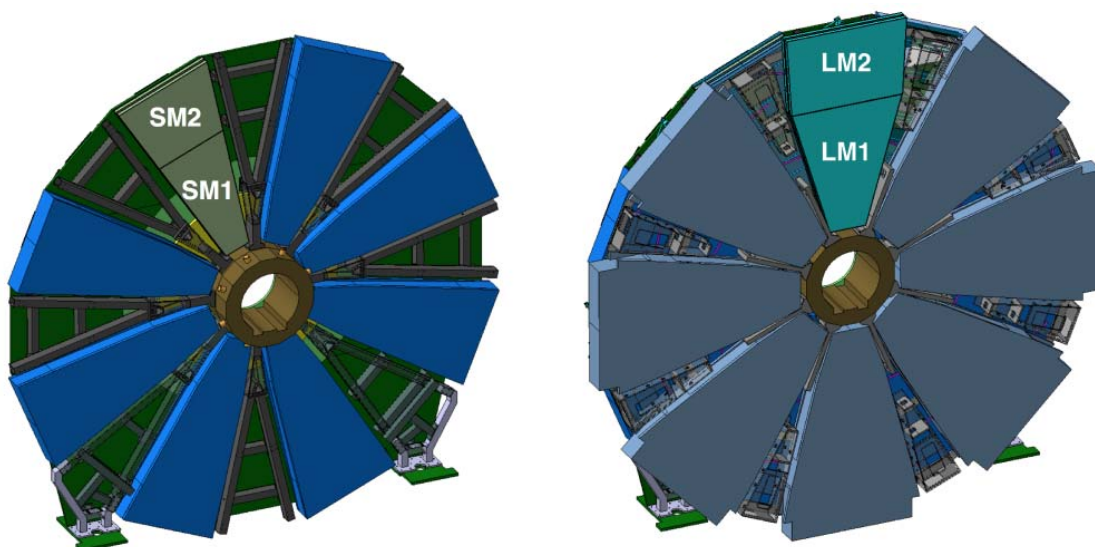


Figure 3.3 – Disposition des détecteurs NSW sur une roue [5]. Les deux côtés de la roue sont montrés : à gauche le côté faisant face au point d’interaction composé des huit petits secteurs ; à droite le côté opposé au point d’interaction composé des huit grands secteurs. Chaque petit (respectivement grand) secteur comporte deux parties : SM1 et SM2 (respectivement LM1 et LM2). Les petits et grands secteurs se recouvrent légèrement de sorte que la couverture angulaire soit totale.

bras de levier possible pour améliorer la résolution angulaire des traces reconstruites. Les détecteurs MM ont quant à eux une résolution intrinsèque suffisante.

Le choix d’assembler huit détecteurs de chaque technologie dans un seul secteur est motivé par le risque d’usure des détecteurs tout au long du déroulement de l’expérience, ainsi que l’éventualité du fonctionnement de certains secteurs en mode dégradé. Le système doit permettre une efficacité proche de 100 %, même en considérant ces risques.

Chaque type de module SM1, SM2, LM1, LM2 est construit par un ou plusieurs laboratoires différents. L’assemblage final des secteurs sera effectué au Cern. L’Irfu assure la construction complète de tous les modules Micromegas LM1, au nombre de 32. Les modules LM2 sont produits en Grèce (Thessalonique, Athènes) et en Russie (Dubna) ; les modules SM1 en Italie (Pavia, Frascati, Rome) ; et les modules SM2 en Allemagne (Freiburg, Würzburg, Munich).

3.2 Les détecteurs Micromegas pour *New Small Wheel*

Le travail présenté ici concerne les modules Micromegas LM1 construits à l’Irfu. Des détails sur la technologie utilisée ainsi que les spécificités des détecteurs produits par

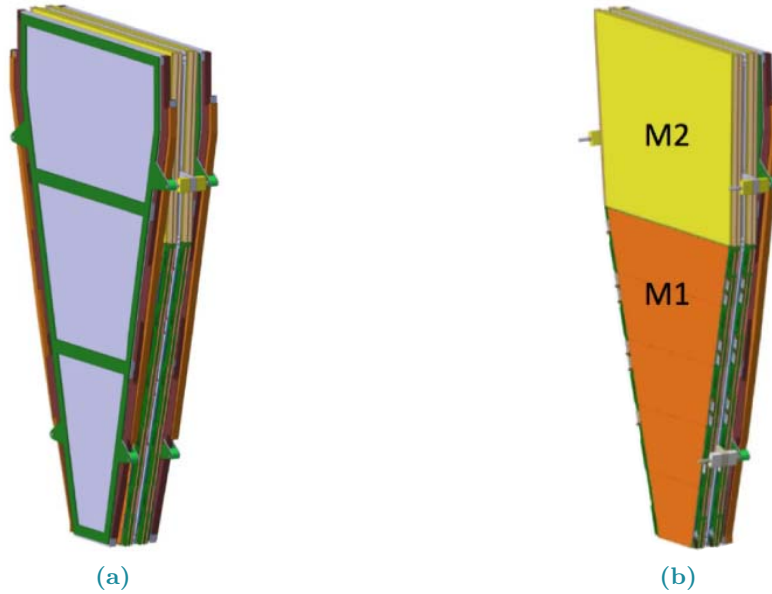


Figure 3.4 – (a) Vue d'un secteur NSW [5]. Il est composé de quatre quadruplets sTGC-MM-MM-sTGC. Les deux quadruplets MM ont une tranche de couleur verte sur l'image. (b) Vue des deux quadruplets MM d'un secteur NSW [5]. Chaque quadruplet est constitué de quatre couches de détection, voir figure 3.6 (a). Il est séparé en deux parties M1 et M2.

l'Irfu vont être présentés.

3.2.1 La technologie Micromegas

Le principe de la technologie Micromegas (pour MICRO-MEsh-Gaseous Structure) [6] repose sur deux espaces situés entre une électrode et un circuit imprimé, électriquement séparés par une grille métallique. Il s'agit de l'espace de dérive (ou de conversion) entre l'électrode et la grille, et de l'espace d'amplification entre la grille et le circuit imprimé. Ils sont très asymétriques en termes de taille et de champ électrique. Des pistes métalliques parallèles sont présentes sur le circuit imprimé. L'électrode et les pistes (ou parfois la grille) sont portées au potentiel électrique nécessaire pour générer les deux champs électriques.

Le détecteur contient un gaz ou un mélange de gaz permettant la conversion de l'énergie incidente des particules chargées en signal électrique. Il s'agit souvent d'un gaz noble puisqu'un tel gaz n'est pas réactif et possède une couche électronique externe pleine, il est donc facilement ionisé. Il peut être mélangé à un gaz polyatomique dit *quencher* qui permet la dissipation d'une grande partie de l'énergie sans création d'ions ni de photons,

permettant de limiter la divergence des avalanches électroniques et ainsi d'opérer de manière stable à des gains élevés [7].

Les particules chargées pénétrant le détecteur interagissent avec le gaz et génèrent des électrons dits primaires qui dérivent jusqu'à l'espace d'amplification grâce au champ de dérive. L'espace de dérive mesure en général quelques mm et le champ de dérive est suffisamment faible pour ne pas générer d'avalanche électronique. Il permet de contrôler la vitesse de dérive des électrons primaires. L'espace d'amplification mesure quelques centaines de μm et contient un champ de 40 à 50 kV cm^{-1} . Un électron traversant cet espace est fortement accéléré et peut ioniser le gaz alentour qui génère alors des photons UV capables d'arracher de nouveaux électrons au gaz. Le processus se répète avec les nouveaux électrons, on parle d'avalanche électronique. Les électrons acheminés en direction des pistes ainsi que les ions se déplaçant en sens inverse induisent un signal mesuré par l'électronique de lecture.

3.2.2 Spécificités des détecteurs Micromegas pour le projet NSW

Le gaz proposé pour les détecteurs Micromegas du projet NSW est un mélange composé de 93 % d'argon et 7 % de CO_2 , le même mélange que pour les tubes à dérive (MDT) actuels. Les chambres CSC actuelles utilisent également un mélange d'argon et de CO_2 , dans les proportions 80 % Ar, 20 % CO_2 . L'argon est choisi pour sa qualité de gaz noble, et le CO_2 a été retenu en tant que *quencher* car il est stable, peu cher, il est déjà utilisé dans ATLAS, et est inerte chimiquement. L'isobutane a également été utilisé en tant que *quencher* au cours de ce travail.

Cependant à forte irradiation les avalanches dans l'espace d'amplification peuvent s'emballer et provoquer des claquages. Ceux-ci risquent d'endommager le détecteur, ainsi que causer des chutes de tension entraînant des temps morts et donc des inefficacités. Pour pallier ce problème, les pistes de lecture des détecteurs sont séparées de l'espace d'amplification par un isolant. Des pistes résistives sont déposées sur ce dernier, celles-ci étant couplées capacitivement aux pistes de lecture. Cette configuration permet de protéger le détecteur en réduisant grandement le courant de décharge en cas de claquage, de sorte que les baisses de tension engendrées soient négligeables [8, 9]. Un schéma de détecteur Micromegas pour NSW est montré sur la figure 3.5.

La figure montre également la grille (« Micromesh ») mais ne représente pas les piliers isolants qui soutiennent la grille et la maintiennent à une distance constante des pistes. Le pas entre deux pistes consécutives est de 450 μm , permettant une très bonne résolution dans la direction transverse aux pistes, de l'ordre de 100 μm .

La cathode de dérive est alimentée en haute tension à une valeur de -300 V qui correspond au point de fonctionnement standard choisi. Les pistes résistives quant à elles sont alimentées à des tensions entre $+550$ et $+600\text{ V}$. La grille est placée à la masse.

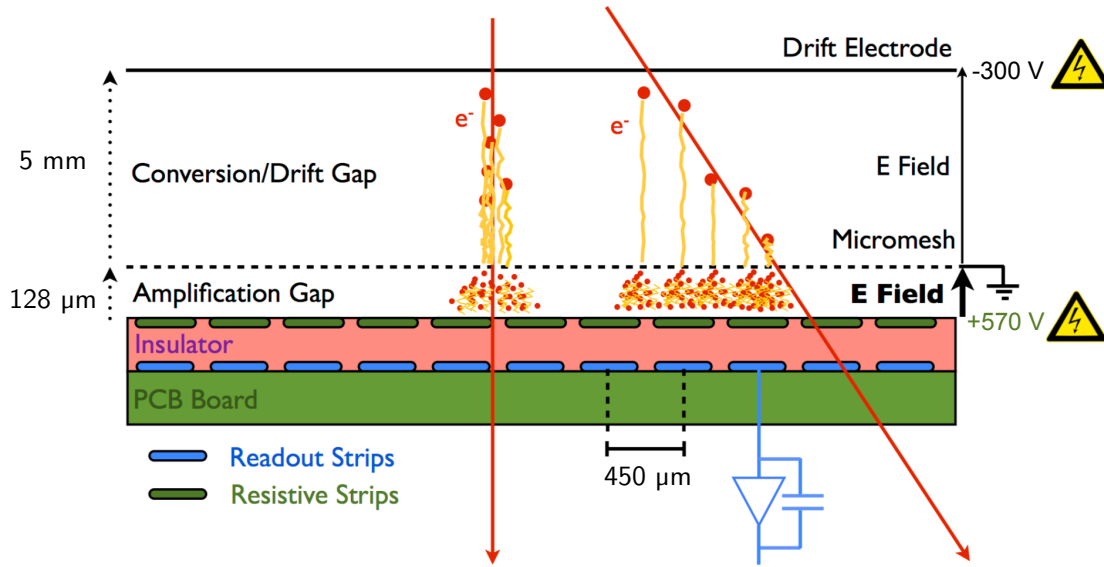


Figure 3.5 – Schéma du concept d'un détecteur Micromegas, les dimensions ne sont ici pas à l'échelle [10]. Les flèches rouges représentent le trajet possible de particules chargées à travers le détecteur, créant des électrons dans l'espace de dérive (« Conversion/Drift gap »), puis générant une avalanche dans l'espace d'amplification (« Amplification gap »). Les électrons sont schématisés par des points rouges se déplaçant le long de trajectoires en jaune. Les pistes résistives sont représentées en vert foncé, et les pistes de lecture en bleu.

3.2.3 Les modules LM1

Les détecteurs Micromegas produits à l'Irfu sont les quadruplets LM1. Les points de fonctionnement cités dans cette section correspondent aux valeurs nominales, requises pour un fonctionnement correspondant aux caractéristiques attendues.

La structure de chaque couche Micromegas de détection suit le schéma de la figure 3.5. Le géométrie trapézoïdale des quatre couches d'un multiplet est représentée sur la figure 3.6 (a).

Sur ce schéma les quatre couches sont représentées sous le nom « layer », avec un numéro allant de 1 à 4. Chaque couche est divisée en cinq « PCB » ou circuits imprimés, numérotés de 1 à 5. Ces circuits imprimés sont produits séparément puis assemblés pour former une couche complète. La direction des pistes est également représentée, elles sont parallèles aux bords de la petite et de la grande base pour les couches 1 et 2, et présentent un angle de $\pm 1,5^\circ$ pour les couches 3 et 4. Les couches 1 et 2 sont appelées « ETA », et les couches 3 et 4 « STEREO ».

Il y a environ 5 120 pistes de lecture sur chaque couche, soit 1 024 par circuit imprimé. Les pistes sont connectées à l'électronique par groupe de 512, un groupe de chaque côté

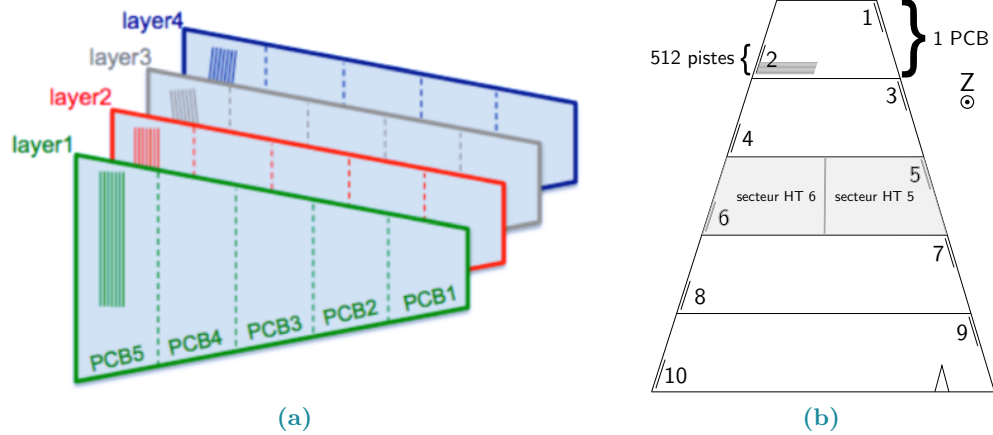


Figure 3.6 – (a) Représentation des quatre couches d’un quadruplet MM LM1 [10]. Chaque couche est constituée de circuits imprimés ou « PCB », numérotés de 1 à 5. Les couches 1 et 2 (en vert et en rouge) ou « ETA » ont des pistes de lecture parallèles aux bords des circuits imprimés. Les couches 3 et 4 (en gris et en bleu) ou « STEREO » ont des pistes de lecture faisant un angle de $\pm 1,5^\circ$ avec bords des circuits imprimés. (b) Représentation d’une couche d’un quadruplet MM LM1 vue de dessus. Les numéros de 1 à 10 correspondent aux zones de connexion avec l’électronique, par groupe de 512 pistes. Il y a également 10 secteurs d’alimentation en haute tension, correspondant à des demi circuit imprimé. Les secteurs 5 et 6 sont montrés ici en gris.

de chaque circuit imprimé, voir figure 3.6 (b).

Les pistes résistives s’étendent sur toute la longueur d’un circuit imprimé, avec une coupure au milieu. Elles sont donc alimentées en haute tension indépendamment de chaque côté, et sont interconnectées via un motif à embranchements. Il y a ainsi dix alimentations en haute tension par couche de détection, un secteur de haute tension correspondant à un demi circuit imprimé, voir figure 3.6 (b). Quatre alimentations supplémentaires permettent d’établir les champs de dérive dans les quatre couches, pour un total de 44 hautes tensions par quadruplet.

Il ne sera pas possible d’installer un canal d’alimentation en haute tension indépendant par demi circuit imprimé lors de l’installation de la petite roue dans ATLAS, pour des raisons d’encombrement. Le schéma retenu consiste à connecter 10 secteurs de haute tension au même canal, répartis sur deux chambres différentes. Ceci pourrait poser problème puisqu’il faudra aligner le réglage de tension sur le secteur le moins performant, réduisant potentiellement l’efficacité des autres secteurs.

Chaque couche de détection est également alimentée en gaz par les deux côtés de la grande base, c’est-à-dire par le circuit imprimé n°5, et la gaz est évacué par les deux bords du circuit imprimé n°1. L’alimentation en gaz dans la petite roue complétée se fera

par l'extérieur, à travers le quadruplet LM2 avant d'entrer dans le quadruplet LM1 puis de ressortir par l'intérieur de la petite roue.

Les détecteurs sont construits et assemblés dans un environnement contrôlé afin d'éviter la présence d'impuretés. Ils sont donc produits dans une salle blanche avant d'en sortir pour le contrôle qualité. Une partie du contrôle qualité est effectué directement en salle blanche, et seuls les détecteurs assemblés peuvent en sortir. L'assemblage de panneaux légers est réalisé en collant une structure rigide en nid d'abeille sur un plan de cinq circuits imprimés, puis en collant un second plan de cinq circuits imprimés de l'autre côté. Cette technique permet de fabriquer des panneaux très plans. Cinq panneaux sont finalement assemblés afin de former un module complet, comme décrit dans la référence [10]. Des goupilles permettent d'aligner avec précision les panneaux assemblés.

3.3 Le banc cosmique de test des modules

Les modules LM1 construits à l'Irfu doivent être testés pour s'assurer que les critères de qualité détaillés dans la section 3.4.4 sont respectés. Il est également nécessaire d'avoir un retour sur les processus de fabrication, en particulier puisque le travail décrit dans cette thèse s'inscrit dans la période de démarrage de la production des modules complets.

Un banc de test a donc été mis en place, afin de pouvoir récolter des informations sur la qualité des modules construits. Ce banc de test tire parti de l'interaction de particules « cosmiques » qui interagissent dans l'atmosphère et génèrent un flux de muons qui atteint la surface terrestre. On parle donc de « banc cosmique » dans la suite de ce chapitre.

Ce travail de thèse s'est concentré sur le développement d'un code d'analyse des données issues de ce banc cosmique. Le travail d'installation, de branchement et d'exploitation des détecteurs placés sur le banc cosmique a principalement été effectué par des membres du Département d'Électronique des Détecteurs et d'Informatique pour la Physique.

3.3.1 Mise en place

Le banc cosmique a été installé près de la salle blanche dans laquelle les modules sont assemblés. Chaque module est déposé sur une structure roulante permettant l'installation en dehors du banc. Ainsi les tubes de gaz et les différents câbles haute tension peuvent être installés à l'extérieur, sans encombre. Une photo du dispositif est montrée sur la figure 3.7. On peut observer le trajectographe externe appelé « Mcube » à l'intérieur duquel le module est placé pour les prises de données. Celui-ci possède deux plans de détection au-dessus du module, et un plan en-dessous.

Une fois mis en place, le détecteur doit être alimenté en gaz. Un mélangeur de gaz permet d'obtenir le mélange requis de 93 % d'argon et 7 % de CO₂ en entrée du module. Un dispositif analyse le gaz sortant par les extrémités des petites bases. L'humidité du

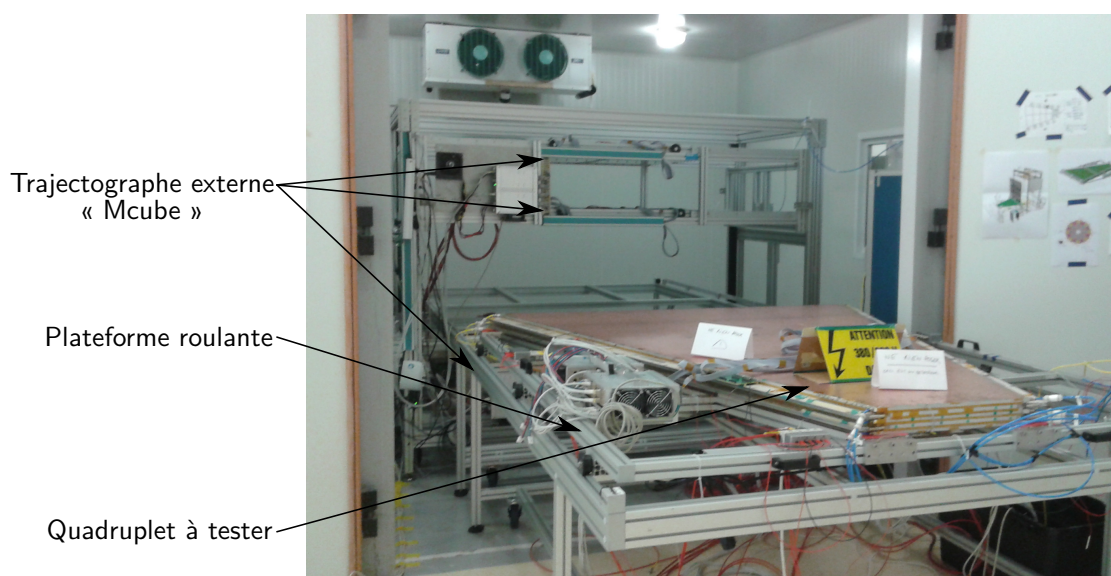


Figure 3.7 – Photo du banc cosmique mis en place à l'Irfu. Le quadruplet à tester est déposé sur une plateforme roulante, à l'extérieur de la salle. Il est ainsi possible de le faire rouler jusqu'à l'intérieur du banc cosmique, au niveau du trajectographe externe « Mcube ». Deux plans de ce dernier sont visibles en haut du dispositif, et un troisième plan est situé plus bas, comme indiqué par les flèches sur l'image.

mélange en sortie est particulièrement surveillée puisqu'elle dégrade les performances du détecteur.

Le détecteur est ensuite alimenté en haute tension au niveau de chaque moitié de circuit imprimé. La tension maximale applicable à chaque zone est déterminée avant toute prise de données. Elle correspond à la tension stable la plus élevée qu'il a été possible d'atteindre dans la zone considérée, avec un courant débité par l'alimentation inférieur à 20 nA. Cette cartographie du détecteur permet de vérifier si tous les secteurs supportent la tension nominale et avec quelle marge de manœuvre.

Il est possible qu'une zone de haute tension ne supporte pas la tension utilisée lors d'une étude donnée à cause d'un défaut de fabrication, par exemple des piliers manquants ou une mauvaise planéité des composants. Selon les cas cette zone sera alimentée à la tension maximale qu'elle supporte ou bien sera complètement éteinte pour la durée de l'étude. On ignorera souvent ces zones lors de calculs d'efficacité, afin de comparer uniquement les zones effectivement portées à la tension désirée.

3.3.2 L'électronique de lecture et le système de déclenchement

L'électronique de lecture VMM3a [11] prévue pour le projet NSW n'est pas encore disponible. Un autre type de carte de lecture est donc utilisé pour les tests effectués à l'Irfu. Il s'agit de cartes d'électronique « DREAM » développées à l'Irfu, pour le projet CLAS 12 [12, 13].

Une carte d'électronique DREAM est composée de 8 puces contrôlant 64 canaux chacune, ce qui correspond bien aux 512 pistes qu'on retrouve sur un côté de circuit imprimé comme montré sur la figure 3.6 (b). Il faut donc deux cartes pour un circuit imprimé : une carte de chaque côté.

La liaison électrique entre les pistes de lecture et celles des cartes d'électronique est assurée grâce à un connecteur amovible, appelé « Zebra » [14]. Ce connecteur est constitué de fils d'or incrustés dans une bande en silicone, comprimée entre la carte de lecture et les pistes du circuit imprimé, comme indiqué sur la figure 3.8. Le serrage du dispositif est réalisé via une barre de compression, et la pression est appliquée en plusieurs points à l'aide d'un tournevis dynamométrique afin de contrôler la force de compression.

Une puce DREAM traite le signal analogique : pré-amplification, mise en forme, échantillonnage et enregistrement. Les 64 canaux fonctionnent en parallèle, à une fréquence d'échantillonnage pouvant aller jusqu'à 40 MHz. L'électronique DREAM peut donc s'adapter à des configurations très différentes.

Chaque canal d'une puce DREAM possède une mémoire analogique constituée de 512 cellules, elle peut ainsi stocker les informations sur 512 intervalles de temps. Elle efface les données les plus anciennes au fur et à mesure que les plus récentes sont écrites, et ceci tant qu'aucun ordre de lecture ne lui parvient. Quand un ordre d'enregistrement est reçu, l'écriture est stoppée et la mémoire présente dans une certaine fenêtre de temps est alors sauvegardée. Cette fenêtre est configurable et commence typiquement quelques 100 ns avant l'ordre reçu, afin de bien lire le signal déclencheur de l'ordre. Cette fenêtre est d'une taille d'au moins 30 intervalles de temps pour bien capturer tout le signal, chaque intervalle de temps correspondant à 60 ns.

Une photo de carte d'adaptation est montrée sur la figure 3.8. Les huit connecteurs DREAM sont visibles, ils sont reliés via les câbles gris à une carte DREAM non visible sur l'image. On peut également voir la barre de compression au bord de la carte en haut de l'image.

Toutes les cartes DREAM peuvent envoyer un signal de déclenchement à un système qui les combine et qui donne un ordre de déclenchement global si plusieurs signaux sont en coïncidence. La mémoire de chaque puce DREAM est alors lue de manière synchronisée. Cependant sauf mention contraire dans la suite, les signaux des cartes DREAM sont ignorés et le déclenchement utilisé est externe aux détecteurs NSW.

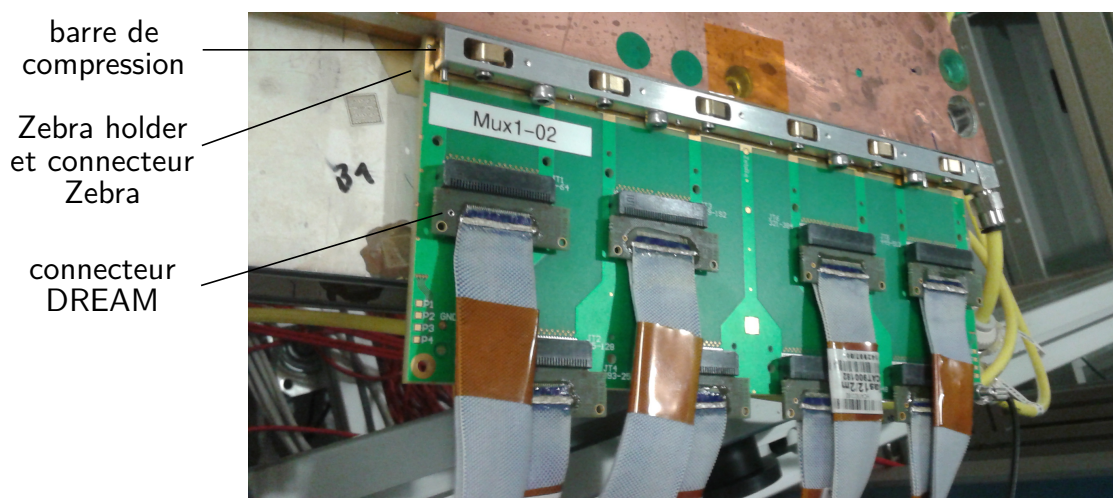


Figure 3.8 – Photo d’une carte d’adaptation avec ses huit connecteurs. Ceux-ci sont reliés à une carte DREAM non représentée sur l’image. Une barre de compression est visible. Elle comprime un connecteur Zebra contre le détecteur, afin d’assurer la liaison électrique entre les canaux de la carte et les pistes du module. Une pièce mécanique appelée « Zebra holder » sur le schéma contient le connecteur Zebra.

3.3.3 Le trajectographe externe « Mcube »

Un trajectographe externe, appelée Mcube [15] est utilisé sur le banc de test comme référence, afin d’estimer notamment l’efficacité des détecteurs. Il est lui-même composée de détecteurs Micromegas disposés selon trois plans de détection : deux au-dessus du module à tester, et un en-dessous. L’ensemble mesure environ 1 m de haut sur environ 1 m de large, et est visible sur la figure 3.7. Seule une partie de module peut être testée à la fois, puisque la taille des plans de détection de Mcube est bien inférieure à celle du module.

Chaque plan de détection est constitué de quatre détecteurs Micromegas de dimensions $0,5\text{ m} \times 0,5\text{ m}$, assemblés en carré. Chaque détecteur comporte des pistes selon deux axes orthogonaux, ce qui lui permet de donner une information de position en deux dimensions dans le plan [16]. L’alimentation en haute tension est effectuée avec un dispositif similaire à celui du détecteur à tester. L’alimentation en gaz est également similaire.

L’ensemble de détecteurs qui composent Mcube a été mécaniquement aligné lors de l’installation du banc cosmique. Les positions de chaque sous-détecteur ne sont toutefois pas parfaitement connues, c’est pourquoi elles doivent être gérées logiciellement. Il est ainsi possible de régler finement la position et l’orientation relative des différents sous-détecteurs afin de prendre en compte les défauts d’alignement mécanique. Plusieurs études exploratoires d’amélioration de l’alignement de Mcube ont été effectuées dans ce

travail de thèse mais sans effet notable sur les résultats. C'est pourquoi la configuration utilisée n'inclut pas de différences de position ni d'angle relatif entre sous-détecteurs d'un même plan. Seule la position verticale des plans a été réglée au millimètre près, en minimisant les résidus des signaux observés par rapport à la prédiction des autres plans. Il existe par conséquent plusieurs pistes d'amélioration de cet alignement, notamment la prise en compte des angles relatifs entre sous-détecteurs qui n'a pas été étudiée à ce jour.

Les détecteurs composant Mcube sont contrôlés par des cartes DREAM qui sont reliées au même système de déclenchement que les cartes branchées au module NSW. Toutes les cartes DREAM sont donc synchronisées par le système de déclenchement, et toutes les mémoires analogiques lues en même temps. Dans les études présentées par la suite, le déclenchement prend en compte uniquement les signaux provenant de Mcube, sauf mention contraire. Ainsi la lecture des événements se fait sans biais provenant du détecteur à tester, Mcube ayant la fonction de déclencheur externe. Un événement n'est enregistré que si des signaux coïncidents sont observés dans les trois plans de Mcube.

3.4 Analyse des données du banc cosmique

Les données prises par le banc cosmique sont traitées hors ligne. À chaque déclenchement du système d'enregistrement correspond un numéro d'événement unique. La trame générale de l'analyse des données est présentée dans cette introduction, puis chaque point est détaillé dans une sous-section dédiée.

Tout d'abord les fichiers binaires générés par la prise de données sont lus : ils contiennent les amplitudes numérisées par événement et par canal électronique. Ces amplitudes sont ensuite comparées à un seuil calculé à partir des informations de piédestaux : si le seuil est dépassé pendant plusieurs intervalles de temps consécutifs, un « coup » est enregistré. Les coups adjacents sont analysés et regroupés en « amas ». Enfin, la position de chaque amas est comparée à une prédiction du trajectographe externe pour déterminer l'efficacité. Un résumé des coupures et autres actions appliquées aux différents objets lors de l'analyse des données est présenté dans le tableau 3.1.

3.4.1 Pré-traitement des données

À chaque ordre d'enregistrement de données déclenché par le passage d'un muon correspond un événement. Les données de chaque événement sont enregistrées dans différents fichiers en format binaire spécifique à l'électronique DREAM. Ces fichiers sont lus et les informations décodées : numéro d'événement unique, temps de l'événement, liste d'amplitudes numérisées par canal électronique.

Objet	Action ou coupure
Amplitude	Soustraction de piédestal. Suppression du mode commun.
Coups	Amplitude $> 5 \times \text{RMS}_{\text{canal}}$, Pendant ≥ 180 ns.
Amas	Pistes physiquement voisines (trous de largeur 1 piste acceptés). Taille ≥ 1 coup. Temps dans une fenêtre de 600 ns autour du réglage (fenêtre agrandie si tension de dérive diminuée).
Amas (calcul d'efficacité)	$ \text{résidu} \leq 10$ mm. 10 % de la taille du détecteur ignorés sur chaque bord.

Table 3.1 – Résumé des coupures et autres actions appliquées aux différents objets lors de l’analyse des données.

Soustraction de piédestal

Des données spécifiques, dites de piédestal, sont prises via des déclenchements aléatoires du système d’enregistrement. La tension de dérive est fixée à 0 V pour cette prise de données afin de ne pas détecter de signaux physiques, tandis que la tension d’amplification est fixée à la valeur désirée. Les amplitudes ainsi obtenues correspondent au bruit de l’électronique puisqu’aucune particule chargée n’a déclenché la mesure et donc aucun signal d’intérêt n’est présent.

La valeur moyenne de ces amplitudes, ou « piédestal », est alors calculée par canal d’électronique. Ces valeurs de piédestal sont soustraites de toutes les données prises dans les mêmes conditions, afin de ramener le niveau moyen d’amplitude à zéro en l’absence de signal d’intérêt. Cette opération est réalisée indépendamment pour chaque canal. Il est nécessaire de recalculer les piédestaux et donc de prendre à nouveau des données de piédestal dès que les conditions de prise de données sont modifiées : modification de la tension d’amplification, des branchements, ou encore du gaz par exemple.

Suppression du mode commun

Après soustraction de piédestal, il subsiste encore une forme de bruit cohérent observable simultanément sur plusieurs canaux, appelé « mode commun ». Ce bruit se manifeste sous la forme de déplacement dans le temps de la ligne de base des amplitudes de tout ou partie des canaux d’une puce DREAM, de manière cohérente.

Une estimation de ce bruit est calculée pour chaque intervalle de temps d’un événement donné, à partir des amplitudes des 64 canaux d’une puce DREAM donnée. La médiane

de cette liste d'amplitudes est sélectionnée, afin d'éviter de moyenner un éventuel signal dans l'estimation de ce bruit. Un signal réel est en effet de taille bien inférieure à 32 pistes et donc ne perturbe pas la médiane des amplitudes.

La valeur du bruit cohérent ainsi obtenue est alors soustraite des amplitudes pour l'événement et l'intervalle de temps considérés, et ceci pour les amplitudes de tous les canaux associés à la puce DREAM considérée.

Il est important de noter que cette suppression du mode commun est effectuée sur les données de piédestal, avant de calculer la moyenne quadratique des piédestaux par canal qui sera utilisée par la suite.

Effet des corrections

L'effet des deux corrections décrites plus tôt est présenté sur la figure 3.9. On se place en l'absence de signal. Deux séries d'amplitudes sont montrées, l'une avant les corrections (a) et l'autre après (b). À chaque canal est associé une couleur différente. L'axe des ordonnées mesure les amplitudes numérisées avec l'unité « ADC », comportant 4096 valeurs. En abscisse est indiqué le temps en termes de numéro d'échantillon, ici de 0 à 49. Un échantillon dure 60 ns.

La moyenne des amplitudes se situe initialement autour de 300 ADC et est centrée autour de zéro après les corrections. De plus, il est visible sur la figure (a) que les amplitudes des différents canaux se déplacent de manière simultanée avant les corrections. Il s'agit de bruit cohérent. Ce bruit n'est plus visible sur la figure (b) grâce à la suppression du mode commun. La dispersion des valeurs d'amplitude passe d'environ 200 ADC à environ 40 ADC grâce aux corrections.

La figure 3.10 présente l'effet de ces mêmes corrections en présence d'un signal. Les numéros d'échantillon s'étendent dans ce cas de 0 à 31. La soustraction de piédestal est à nouveau bien visible puisque la moyenne des amplitudes passe d'environ 300 ADC à 0. Il est possible de distinguer quatre canaux différents montrant un signal de plusieurs centaines d'ADC dont le maximum se situe au niveau de l'échantillon de temps numéro 7.

3.4.2 Sélection des événements

Une fois les données pré-traitées et la correction de piédestal et de mode commun effectuées, les amplitudes corrigées sont lues à nouveau, les signaux d'intérêt sont détectés, et les événements ne contenant aucun signal sont éliminés.

Détection des coups sur les pistes

Pour chaque événement enregistré, les amplitudes corrigées du piédestal et du bruit cohérent sont lues et comparées à un seuil. Celui-ci est calculé à partir de la moyenne quadratique des piédestaux pour le canal considéré. Les valeurs de piédestal sont corrigées

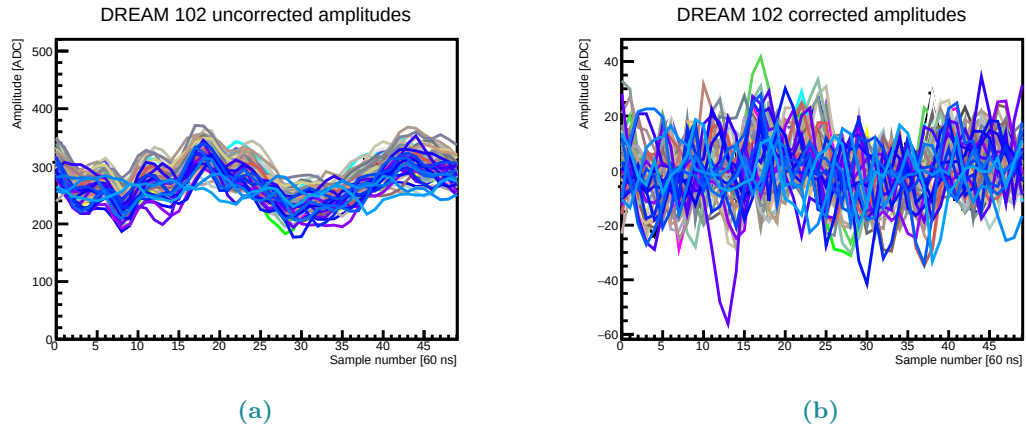


Figure 3.9 – (a) Amplitudes en l’absence de signal avant correction de piédestal et suppression du bruit cohérent. (b) Amplitudes en l’absence de signal après correction de piédestal et suppression du bruit cohérent. Les amplitudes sont exprimées en « ADC », soit leur valeur numérisée. L’abscisse représente le temps exprimé en numéro d’échantillon, ici de 0 à 49. Chaque couleur représente un canal distinct, il y en a 64.

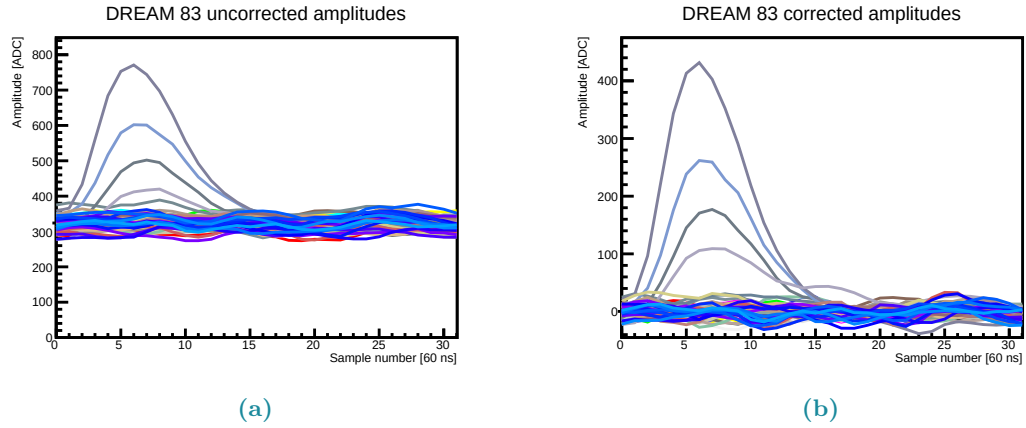


Figure 3.10 – (a) Amplitudes en présence d’un signal avant correction de piédestal et suppression du bruit cohérent. (b) Amplitudes en présence d’un signal après correction de piédestal et suppression du bruit cohérent. Les amplitudes sont exprimées en « ADC », soit leur valeur numérisée. L’abscisse représente le temps exprimé en numéro d’échantillon, ici de 0 à 31. Chaque couleur représente un canal distinct, il y en a 64.

du bruit cohérent avant détermination de cette moyenne quadratique. Le seuil de sélection est donc calculé séparément pour chaque canal et fixé à cinq fois la moyenne quadratique des piédestaux dans ce canal. Cette valeur a été choisie suffisamment basse pour ne pas éliminer trop de signaux, mais suffisamment élevée pour rejeter les bruits de fond.

Si les amplitudes lues dépassent le seuil pendant suffisamment d'intervalles de temps consécutifs, on enregistre un « coup ». L'objectif de cette coupure est d'éliminer des bruits de fond correspondant à une montée en amplitude très brève, en conservant un maximum de signal. Le temps au-dessus du seuil requis par la suite est d'au moins trois intervalles de temps consécutifs soit 180 ns, sauf mention contraire. Un signal typique dure une dizaine de pas de temps soit 600 ns. Ces valeurs ont été vérifiées sur des signaux observés.

L'amplitude maximale parmi celles qui constituent le coup est sauvegardée. On l'appellera amplitude du coup. Elle est proportionnelle au nombre d'électrons collectés le long de la piste considérée. En effet, ce nombre est proportionnel à l'intégrale de l'amplitude dans le temps, qui est proportionnelle à son maximum grâce à la fonction de mise en forme de la puce DREAM. Cette correspondance nécessite un temps de mise en forme supérieur à la durée du signal. La position en temps de ce maximum est également enregistrée.

Les distributions de nombres de coups sont utilisées pour le contrôle qualité comme décrit dans la section 3.4.4. Deux exemples sont représentés sur la figure 3.18.

Regroupement des coups en amas

Les coups ainsi détectés sont rassemblés en un « amas » s'ils se situent sur des pistes physiquement voisines. Néanmoins certaines pistes peuvent être déconnectées ou bruyantes et peuvent donc mal fonctionner. C'est pourquoi deux pistes séparées de deux fois la distance inter-pistes sont considérées comme voisines. Un amas est donc reconstruit même s'il lui manque une ou plusieurs pistes adjacentes, tant que les trous sont de la taille d'une piste au maximum.

La position reconstruite de l'amas est calculée en moyennant les positions des coups le constituant, pondérées par l'amplitude de chaque coup. Il s'agit de la méthode dite du centroïde. De la même façon on calcule la position de l'amas en temps. L'amplitude de l'amas reconstruit est simplement la somme des amplitudes de ses coups constitutifs.

Cette méthode pose problème pour des muons pénétrant le détecteur avec un grand angle par rapport à la verticale. En effet un tel muon dépose son énergie tout au long de son trajet dans l'espace de dérive, ce qui engendre un amas plus large que dans le cas d'un muon vertical. Dans ce cas de figure la position reconstruite par méthode du centroïde ne correspond plus au point de passage du muon au niveau des pistes du détecteur et la résolution spatiale diminue. Par exemple, la résolution spatiale attendue avec la méthode

du centroïde pour des traces inclinées d'environ 30° est typiquement de $500\text{ }\mu\text{m}$, bien trop élevée par rapport à l'objectif de $100\text{ }\mu\text{m}$.

Une solution à ce problème est d'utiliser l'information du temps d'arrivée des différents coups constitutifs de l'amas afin de reconstruire la coordonnée orthogonale au plan des pistes, la vitesse de dérive des électrons étant connue. Cette méthode dite micro-Time Projection Chamber ou μTPC [17] nécessite toutefois une résolution en temps suffisante au niveau de l'électronique. Elle permet d'obtenir une meilleure résolution spatiale que la méthode du centroïde pour des traces d'un angle de 10° ou plus par rapport à la normale au plan des pistes. Après installation des détecteurs dans ATLAS, les traces les traversant se situeront entre 8 et 30° . Le mode de reconstruction μTPC sera donc nécessaire pour assurer une résolution de l'ordre de $100\text{ }\mu\text{m}$. L'algorithme μTPC n'a pas été implémenté dans le travail présenté puisque les principaux objectifs étaient l'étude du gain et de l'efficacité des détecteurs, et non de leur résolution.

Les amas doivent avoir une taille minimale pour être acceptés. La taille de l'amas est définie comme le nombre de coups le constituant. Ce minimum est de un coup dans les études présentées par la suite, si bien qu'aucun amas n'est éliminé par ce critère. En effet de par la variété des conditions de gaz et de haute tension lors des études réalisées, les signaux n'étaient pas toujours de taille supérieure à un coup, notamment pour des faibles valeurs de tension d'amplification.

Un autre critère d'acceptation des amas est leur position en temps. En effet pour un véritable signal, cette valeur doit correspondre aux réglages de l'électronique vis-à-vis du déclenchement. Le temps d'un amas est calculé à partir de la position temporelle du maximum d'amplitude de chaque coup le constituant. Ces valeurs de temps sont pondérées par l'amplitude de l'amas (méthode du centroïde), comme dans le cas de sa position. Une fenêtre de 600 ns autour de la valeur attendue est définie, et tout amas se situant en dehors de cette fenêtre est éliminé. Lors des études impliquant une modification de la tension de dérive, la fenêtre doit être agrandie du fait de l'étalement en temps des signaux. Dans le cas le plus extrême d'une tension de dérive fixée à 25 V , la fenêtre choisie est de 1800 ns .

Une distribution du temps calculé des amas reconstruits dans un circuit imprimé est montrée sur la figure 3.11. La distribution présente un pic et s'étend sur environ 10 intervalles de temps soit 600 ns , ce qui correspond bien à la fenêtre choisie.

Un exemple de distribution des positions reconstruites des amas dans un circuit imprimé est présenté sur la figure 3.12. Pour comparaison, la distribution des positions des traces provenant de Mcube dans le plan du détecteur est montrée sur la figure 3.13. Ces deux figures ont été générées à partir des événements provenant de la même prise de données. Les coordonnées X et Y sont calculées à partir des informations de traces provenant de Mcube. Les dimensions de Mcube (forme carrée) et du circuit imprimé considéré (forme trapézoïdale) sont rappelées sur les figures en traits noirs. On peut

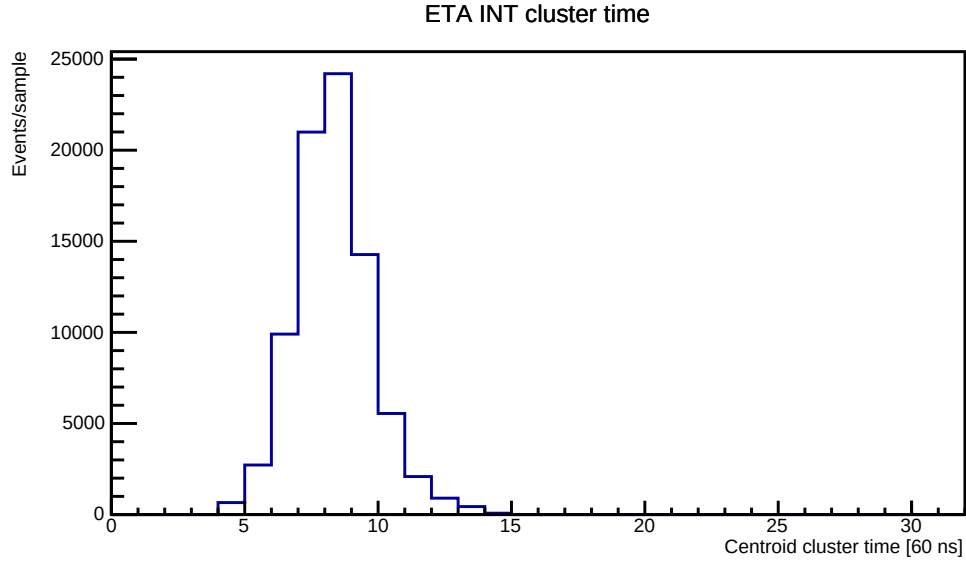


Figure 3.11 – Distribution du temps calculé des amas reconstruits dans un circuit imprimé. La distribution est piquée et s’étend sur environ 10 intervalles de temps soit 600 ns.

notamment observer que le circuit imprimé étudié est plus long que Mcube, et qu’on ne peut donc pas observer ses bords. Pour chaque trace Mcube traversant la couche étudiée, on vérifie qu’un amas a bien été reconstruit autour du point de rencontre entre la trace et la couche. La couleur sur la figure 3.12 désigne la proportion de traces correspondant à un amas reconstruit à ± 10 mm de la prédiction. Cette coupure est utilisée pour les calculs d’efficacité, comme expliqué dans la section 3.4.3. Seule la distance orthogonale avec la piste de l’amas en question est considérée puisqu’il n’existe aucune information sur la position de l’amas le long de sa piste.

La figure 3.12 laisse apparaître deux zones, l’une majoritairement de couleur jaune et l’autre se situant entre le vert et l’orange. La première correspond à une zone de grande efficacité, où la plupart des traces détectées par Mcube correspondent à un amas. L’autre correspond à une zone d’efficacité inférieure à 80 % où plusieurs traces sont observées par Mcube mais aucun n’amas ne peut être reconstruit par le module étudié. Cette faible efficacité provient de la tension d’amplification utilisée. En effet durant la prise de données en question le secteur de haute tension du haut de la figure était porté à une tension de 580 V, contre 510 V pour le secteur du bas qui ne supportait pas de tension plus élevée.

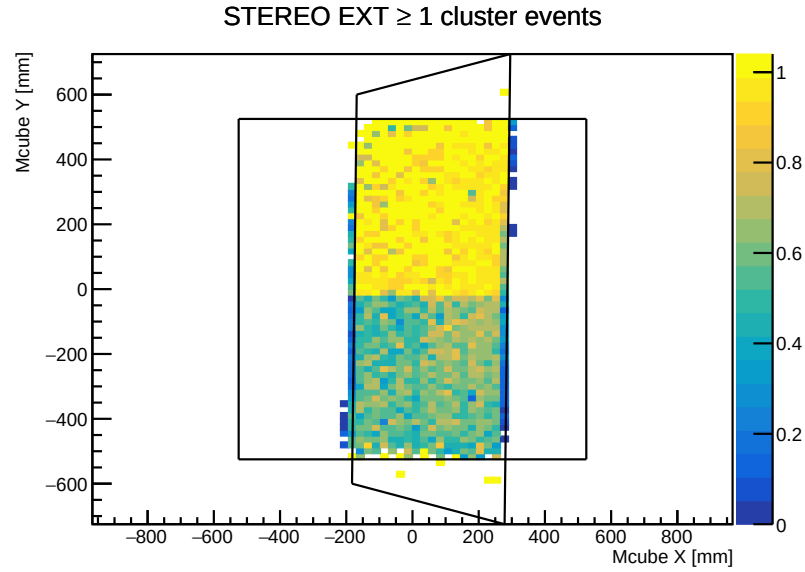


Figure 3.12 – Distribution des amas reconstruits. Les coordonnées X et Y sont calculées à partir des informations de traces provenant de Mcube. Les dimensions de Mcube (forme carrée) et du circuit imprimé étudié (forme trapézoïdale) sont rappelées en traits noirs. La couleur désigne la proportion de traces correspondant à un amas reconstruit à ± 10 mm de la prédiction.

3.4.3 Caractérisation de la qualité des modules

Cette section présente les différentes caractéristiques d'un module utilisées lors des études présentées dans la section 3.5 et section 3.6, à savoir le gain et l'efficacité. Les méthodes de calcul de ces grandeurs sont décrites. La définition des différents critères de qualité sont quant à eux présentés en section 3.4.4.

Calculs de gain

Une caractéristique essentielle d'un détecteur gazeux est son gain, qui représente l'amplification du signal initié par la particule chargée qu'on cherche à détecter. En effet, une particule chargée traversant un gaz ionise certaines molécules de ce gaz et ainsi détache des électrons qu'on appelle électrons primaires. Ces électrons migrent depuis la zone de dérive vers la zone d'amplification du détecteur grâce au champ de dérive. Le nombre d'électrons primaires créés lors du passage d'une particule au minimum d'ionisation dépend de la composition du gaz et de la distance traversée [18].

Quand ces électrons primaires arrivent dans l'espace d'amplification, ils rencontrent un champ électrique beaucoup plus important que le champ de dérive. Ils vont alors

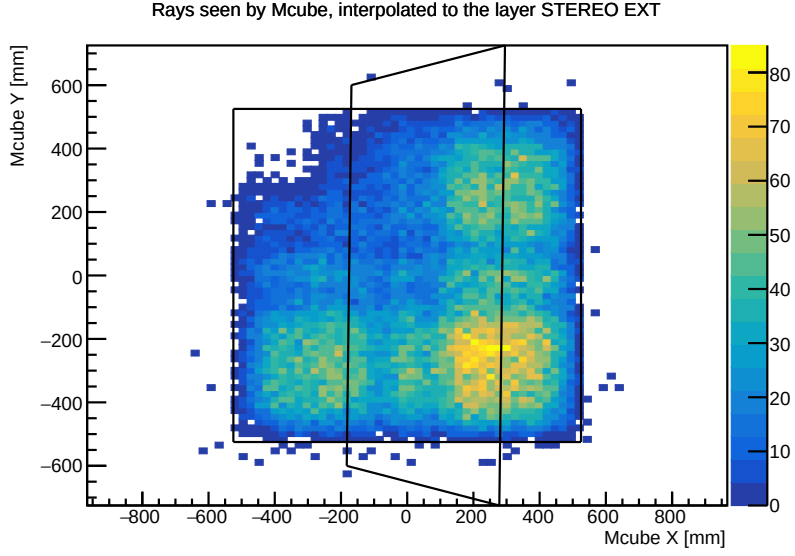


Figure 3.13 – Distribution des positions des traces provenant de Mcube dans le plan d’un détecteur. Les dimensions de Mcube (forme carrée) et du circuit imprimé étudié (forme trapézoïdale) sont rappelées en traits noirs. Le quart en haut à gauche de l’un des plans de détection de Mcube n’était pas opérationnel pour cette prise de données.

ré-ioniser d’autres molécules du gaz et produire une cascade, ce qui va multiplier le nombre d’électrons acheminés en direction des pistes. Le gain est alors défini comme le facteur de proportionnalité entre la charge totale induite Q (électrons et ions) et le nombre d’électrons primaires :

$$\text{gain} = \frac{Q}{N_{e^-}^{\text{primaire}}}. \quad (3.1)$$

Pour estimer le gain dans cette analyse, on utilise les caractéristiques des signaux de l’électronique DREAM. En effet, le maximum d’amplitude électronique d’un coup est proportionnel à l’intégrale de l’amplitude en fonction du temps, c’est-à-dire de la charge induite sur la piste correspondante :

$$Q = \frac{\text{amplitude max}}{2^n} \times \frac{\text{plage de gain}}{e} \quad (3.2)$$

où n est le nombre de bits utilisés pour la conversion du signal analogique en signal numérique, 2^n est le nombre de valeurs d’amplitude possibles, la plage de gain représente la charge correspondant à une valeur d’amplitude maximale exprimée en fC, et e est la charge élémentaire.

Par exemple, avec le réglage de plage de gain le plus souvent utilisé dans cette étude : 100 fC, et en considérant le codage de l'électronique sur 12 bits, donc $2^{12} = 4096$ valeurs possibles, on obtient :

$$N_{e^-}^{\text{secondaire}} = \frac{\text{amplitude max}}{4096} \times \frac{100 \text{ fC}}{1,60 \times 10^{-4} \text{ fC}}. \quad (3.3)$$

Le nombre d'électrons primaires générés par particule incidente quant à lui est calculé à partir de la taille de l'espace de dérive et du nombre d'électrons primaires générés par unité de distance pour le gaz considéré. On ne prend pas en compte l'effet de l'angle des trajectoires par rapport à la verticale qui allonge la distance de parcours :

$$N_{e^-}^{\text{primaire}} = n_{e^-}^{\text{primaire}} \times \text{distance de dérive}. \quad (3.4)$$

En combinant les Équations 3.1 et 3.2, on obtient :

$$\text{gain} = \frac{1}{n_{e^-}^{\text{primaire}} \times \text{distance de dérive}} \times \frac{\text{amplitude max}}{2^n} \times \frac{\text{plage de gain}}{e}. \quad (3.5)$$

La taille de l'espace de dérive des détecteurs utilisés est 5 mm, on considère la plage de gain usuelle de 100 fC et un codage sur 12 bits, on obtient alors :

$$\text{gain} = \frac{1}{n_{e^-}^{\text{primaire}} [\text{cm}] \times 0,5} \times \frac{\text{amplitude max}}{4096} \times \frac{100 \text{ fC}}{1,60 \times 10^{-4} \text{ fC}}. \quad (3.6)$$

Le nombre d'électrons secondaires n'est pas toujours le même pour un nombre d'électrons primaires donné. Ce dernier n'est pas constant non plus puisque le processus de dépôt d'énergie dans le gaz est Poissonnien, et les muons cosmiques ne sont pas mono-énergétiques. La charge collectée suit donc une distribution de Landau. Il faut ajuster une fonction de Landau sur la distribution observée des amplitudes des amas reconstruits et sélectionner la valeur la plus probable pour obtenir l'amplitude à utiliser dans les Équations 3.2, 3.3, 3.5 et 3.6. Un exemple d'ajustement sur une distribution d'amplitudes d'amas est présentée sur la figure 3.14. La valeur la plus probable se situe ici autour de 650 ADC.

Le nombre d'électrons primaires générés par unité de distance $n_{e^-}^{\text{primaire}}$ pour des gaz usuels est pris de la référence [18]. Les gaz utilisés avec les détecteurs NSW sont l'argon, le CO_2 et l'isobutane. Les valeurs retenues pour ces gaz sont respectivement $n_{e^-}^{\text{primaire, Ar}} = 94 \text{ cm}^{-1}$, $n_{e^-}^{\text{primaire, CO}_2} = 91 \text{ cm}^{-1}$ et $n_{e^-}^{\text{primaire, HC(CH}_3)_3} = 195 \text{ cm}^{-1}$.

Par exemple, pour le mélange de gaz nominal 93 % Ar et 7 % CO_2 , le nombre d'électrons primaires générés par unité de longueur est $n_{e^-}^{\text{primaire}} = 0,93 \times 94 + 0,07 \times 91 = 93,8 \text{ cm}^{-1}$. Dans ce cas de figure, en utilisant l'équation 3.6, on trouve :

$$\text{gain} = \frac{1}{46,9} \times \frac{\text{amplitude max}}{4096} \times \frac{100 \text{ fC}}{1,60 \times 10^{-4} \text{ fC}}. \quad (3.7)$$

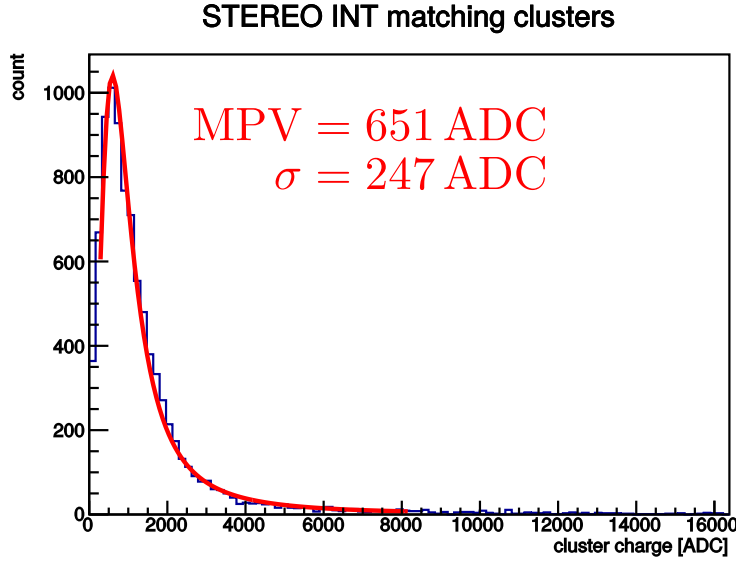


Figure 3.14 – Distribution des amplitudes des amas reconstruits sur un circuit imprimé. Seuls les amas se situant à ± 10 mm d’une trace Mcube sont considérés. Un ajustement de fonction de Landau est réalisé et la fonction obtenue est présentée en rouge. L’écart-type σ ainsi que l’abscisse du pic de la fonction appelé MPV sont donnés sur la figure.

À titre d’illustration, la figure 3.15 montre l’évolution du gain d’une chambre Micromegas en fonction du champ d’amplification, pour différents mélanges de gaz. La méthode d’estimation du gain dans ce cas de figure est similaire mais différente de celle présentée plus haut et utilisée dans ce travail de thèse. Les auteurs de la figure ont en effet utilisé des sources mono-énergétiques pour irradier leur détecteur et non des muons cosmiques [19]. Ceci rend le calcul du gain plus simple puisque l’énergie incidente de chaque particule traversant le détecteur est parfaitement connue, et l’amplitude apparaissant dans l’équation 3.2 par exemple correspond au centre d’un pic gaussien dans la distribution d’amplitude des amas, et non à la valeur la plus probable d’une distribution de Landau.

La figure montre que le gain augmente exponentiellement avec le champ d’amplification, et que les mélanges de gaz contenant de l’isobutane, appelé « Iso » sur la figure, ont un gain typiquement plus élevé que les mélanges contenant du CO_2 , tout en nécessitant un champ d’amplification moins important. Cependant l’isobutane est inflammable et entraîne un vieillissement des détecteurs. Si on s’intéresse aux mélanges d’argon et de CO_2 , on peut observer que le gain maximal pouvant être obtenu ainsi que la tension maximale applicable augmentent avec la proportion de CO_2 .

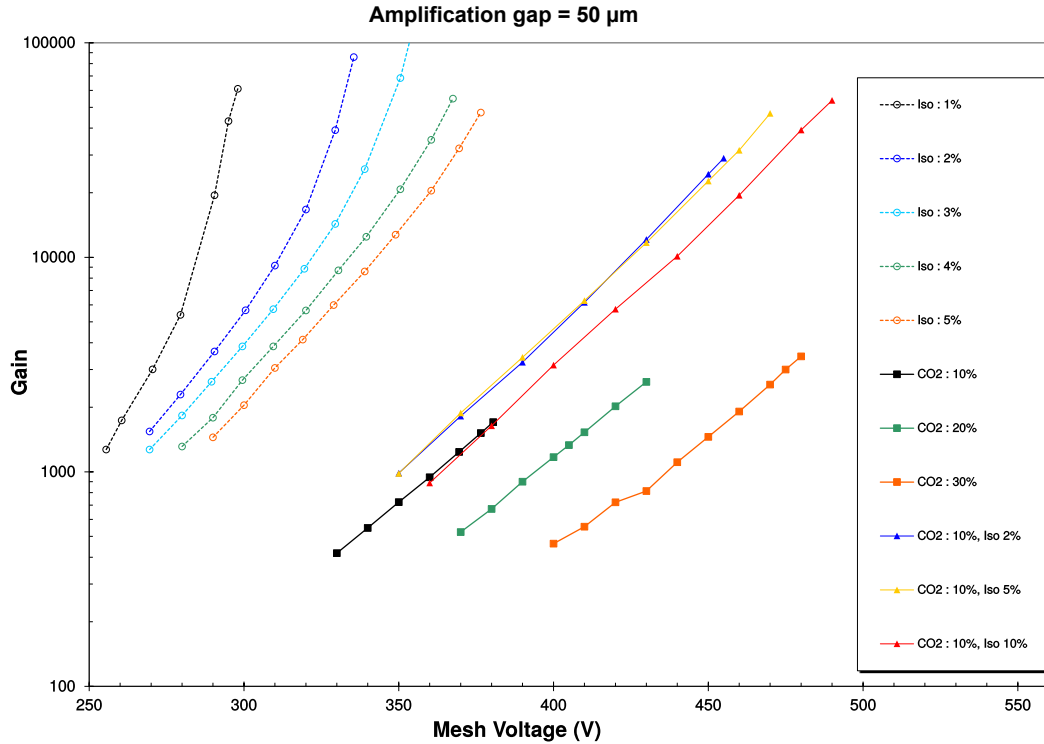


Figure 3.15 – Évolution du gain d’une chambre Micromegas en fonction du champ d’amplification, pour différents mélanges de gaz. La taille de l’espace d’amplification est de 50 µm. Figure adaptée de [19].

Calculs d'efficacité

La caractéristique essentielle des modules étudiés est leur efficacité. L'efficacité représente la capacité d'un détecteur à observer un signal valide pour chaque particule d'intérêt le traversant. Dans le cas du banc cosmique, il s'agira de détecter avec succès les muons atmosphériques traversant les différentes couches du détecteur.

Il y a plusieurs façon de procéder pour estimer cette efficacité. Le module étudié étant composé de quatre couches, il est possible de tester chaque couche en utilisant comme référence les trois autres. Autrement dit, on reconstruit la trace observée d'un muon traversant les trois couches de référence, et on vérifie l'existence d'un signal valide dans le quatrième plan autour de la position prédite. L'avantage de cette méthode est de pouvoir tester un module sans nécessiter d'appareillage externe. Néanmoins pour que cette estimation soit valide, il faut que les trois couches de référence fonctionnent conformément aux spécifications.

Puisque les modules étudiés ici sont les premiers de la série, il a été choisi d'utiliser

une référence externe plutôt que de se baser sur une reconstruction en interne. Un essai préliminaire de la méthode interne a toutefois pu être effectué, et est présenté dans la section 3.6.4.

La méthode externe utilisée dans ce travail consiste à utiliser un trajectographe externe comme référence. Ce trajectographe est constitué de plans de détection situés en-dessous et au-dessus de l'objet étudié. L'appareil utilisé, appelé Mcube, a été détaillé dans la section 3.3.3. La méthode et le code de reconstruction de traces associés à cet appareil ont déjà été validés sur d'autres projets. Le fonctionnement de la méthode externe est détaillé dans la suite de cette section.

Le principe de reconstruction des traces Mcube est similaire à celui présenté dans les sections précédentes. Des coups puis des amas sont reconstruits sur les pistes des différentes sous-détecteurs. Une différence notable provient du multiplexage des pistes de lecture. Les pistes sont connectées par groupe de 16 pour un même canal électronique. Ainsi un coup détecté dans ce canal peut provenir de n'importe laquelle des pistes appartenant au groupe. Le multiplexage est réalisé de sorte qu'à partir d'au moins deux coups voisins il est possible de déterminer la position des coups de manière unique.

L'autre différence majeure est l'existence de pistes selon deux axes orthogonaux X et Y dans les sous-détecteurs de Mcube. Les amas sont reconstruits indépendamment sur les pistes X et les pistes Y . Les informations de géométrie sont alors utilisées pour reconstruire des traces bi-dimensionnelles dans les plans XZ et YZ . Toutes les combinaisons de ces traces en deux dimensions sont considérées pour générer des traces tri-dimensionnelles, et les paramètres d'ajustement des traces dans les deux plans χ_X^2 et χ_Y^2 sont sauvegardés.

Le trajectographe externe fournit des coordonnées en deux dimensions des traces qu'il reconstruit. Les détecteurs à tester quant à eux ne donnent qu'une information unidimensionnelle puisque toutes les pistes d'une même couche sont parallèles. L'information de position d'un amas le long d'une piste est donc inaccessible. C'est pourquoi la distance entre une trace et une amas, ou « résidu », est calculée suivant l'axe orthogonal aux pistes.

La distribution des résidus permet d'ajuster l'altitude Z du détecteur étudié par rapport à l'altitude des plans de détection Mcube. Z correspond au milieu de l'épaisseur de l'espace de conversion. On choisit l'altitude qui minimise la largeur de la distribution. Les traces non verticales sont en effet sensibles à la valeur d'altitude choisie puisque cette dernière permet de calculer la position de l'intersection de la trace avec le plan du détecteur considéré. C'est pourquoi la distribution des résidus s'élargit dès que l'altitude choisie s'éloigne de sa valeur réelle. Une fois cet ajustement de Z effectué, le détecteur peut être aligné avec Mcube en centrant la distribution des résidus à zéro.

La résolution attendue de Mcube est de l'ordre de 1 mm, ce qui est bien supérieur à la résolution attendue des modules MM. Une distribution de résidus est présentée sur la figure 3.16. On observe des résidus gaussiens ayant un écart-type d'environ 1,8 mm.

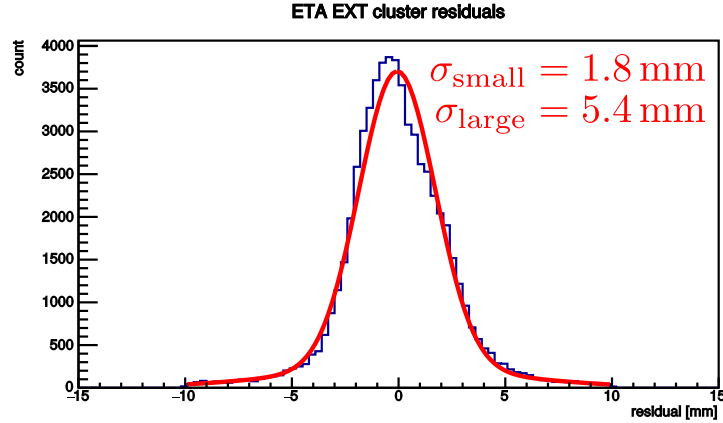


Figure 3.16 – Distribution des résidus entre les traces de référence et les pistes des amas sélectionnés. Chaque entrée de l’histogramme correspond à un unique événement. La distribution est ajustée avec une somme de deux gaussiennes centrées à la même valeur. Les écarts-types de ces deux gaussiennes sont indiqués sur la figure.

Cette valeur est dominée par la résolution de Mcube et s’explique du fait de l’alignement imparfait du trajectographe externe, ainsi que de la grande acceptance angulaire.

Il est nécessaire d’appliquer une coupure sur les paramètres d’ajustement des traces Mcube, afin d’éliminer les traces les moins plausibles. Dans le cas d’un événement contenant plusieurs traces, une seule d’entre elles est sélectionnée. La trace qui est en accord avec au moins un amas sur le plus grand nombre de couches est choisie. On considère que la trace est en accord avec un amas si le résidu associé est inférieur à 10 mm, comme expliqué dans la suite de cette section. En cas d’égalité sur le nombre de couches, on choisit la trace qui possède la valeur la plus faible de χ^2_X . En effet l’axe Y est parallèle aux pistes « ETA », autrement dit à l’axe selon lequel aucune information de position n’est disponible. La remarque est la même pour les pistes « STEREO » avec un faible angle. La valeur de χ^2_Y n’a donc que peu d’influence dans les études effectuées comme celles présentées dans la section 3.6.3.

L’efficacité d’un plan de détection est calculée par piste en faisant le rapport du nombre d’événements contenant un amas valide sur la piste, et le nombre d’événements avec une trace de référence touchant la piste. Un amas est valide s’il est situé à ± 10 mm de la prédiction de Mcube, le long de l’axe orthogonal aux pistes. Cette valeur a été choisie puisqu’elle correspond à environ cinq fois la largeur de la distribution des résidus. Cette coupure diminue les valeurs d’efficacités d’environ 2 % par rapport à une absence de coupure sur la position des amas. Elle est donc suffisamment grande pour ne pas rejeter trop de signaux valides et suffisamment petite pour ne pas générer de coïncidences fortuites.

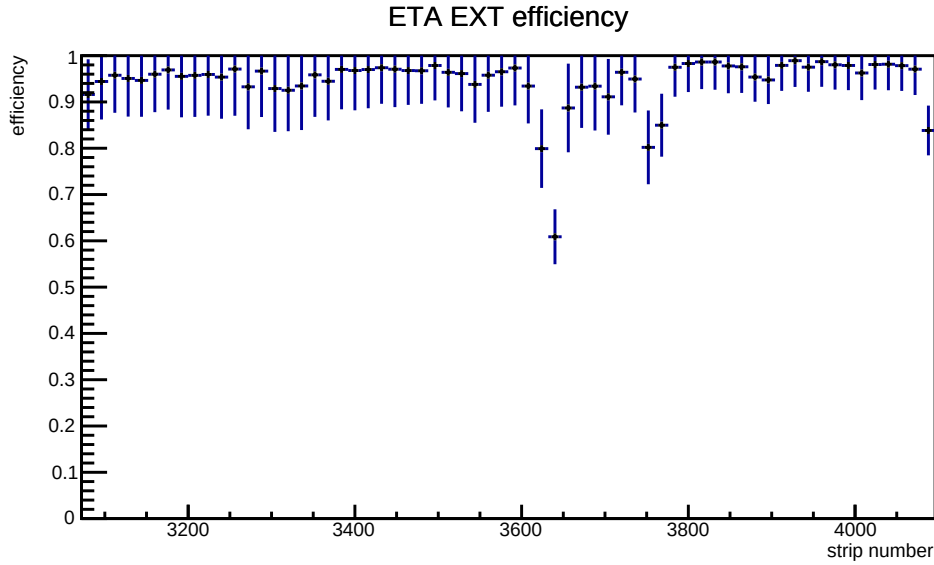


Figure 3.17 – Graphe d’efficacité en fonction du numéro de piste.

Une coupure d’angle maximal par rapport à la verticale autorisé pour une trace Mcube est également définie. En effet, il a été expliqué dans la section 3.4.2 que la méthode du centroïde utilisée pour calculer la position des amas perd en résolution avec un angle grandissant. Cette résolution dégradée pourrait entraîner le rejet à tort d’amas reconstruits trop loin des traces Mcube de référence. La coupure d’angle maximal permet donc de réduire le biais du chiffre d’efficacité obtenu par rapport à la distribution angulaire des traces Mcube. Une étude de l’influence de cette coupure ainsi que des coupures sur les paramètres d’ajustement des traces Mcube est présentée dans la section 3.6.3.

La figure 3.17 montre un graphe d’efficacité en fonction du numéro de piste. Les pistes sont regroupées par seize pour améliorer la lisibilité. L’efficacité moyenne est d’environ 96 %, avec quelques inefficacités locales au niveau des pistes 3600 à 3800 notamment, dues à des pistes bruyantes dans cette zone.

Le chiffre d’efficacité moyenne par couche est utilisé dans les études présentées dans la section 3.5 et section 3.6. Ce chiffre ne prend donc pas en compte les différences entre les pistes. Il est calculé par couche en faisant le rapport du nombre d’événements avec un amas valide sur n’importe quelle piste de la couche, et du nombre d’événements contenant une trace de référence. Cette valeur est obtenue en ignorant 10 % de la taille du détecteur sur chacun de ses bords. D’éventuelles baisses d’efficacité aux extrémités des détecteurs dues à des effets de bord sont ainsi ignorées.

L’efficacité d’un détecteur Micromegas dépend du champ d’amplification mis en place. Plus ce champ est grand et plus le signal est susceptible de se démarquer du bruit, et donc

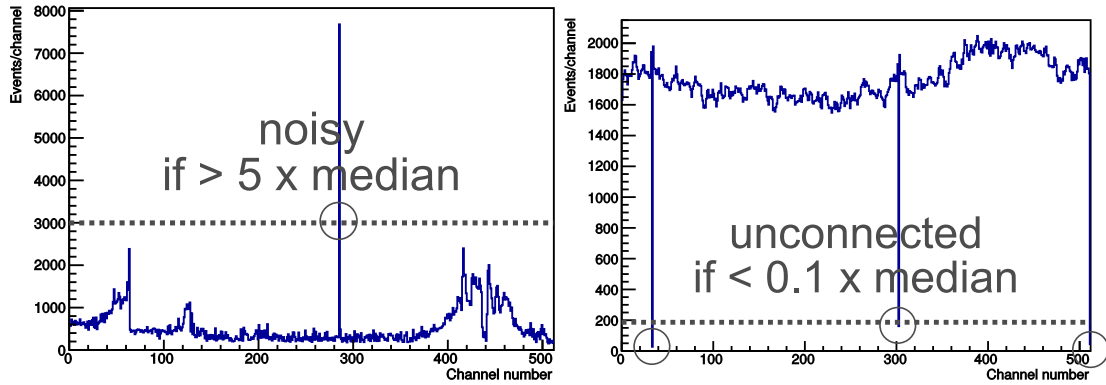


Figure 3.18 – Représentation des critères de détection des pistes bruyantes (gauche) et déconnectées (droite).

plus l'efficacité est grande. Il existe un plateau d'efficacité maximale pour les grandes valeurs de champ d'amplification, et il est souvent souhaitable de choisir un point de fonctionnement sur ce plateau. Il s'agit d'une caractéristique primordiale d'un détecteur.

3.4.4 Critères requis pour le contrôle qualité des modules

Les modules construits à l'Irfu doivent subir un contrôle qualité poussé afin de vérifier le bon respect des spécifications techniques.

Le premier point est la détection de zones bruitées et de pistes non fonctionnelles, par exemple si elles sont coupées. La proportion de pistes bruitées ou déconnectées doit être inférieure à 1 %. Les critères utilisés sont représentés sur la figure 3.18, montrant des distributions de coups le long de 512 pistes. Une piste qui détecte plus de cinq fois plus de coups que la médiane calculée le long des 512 pistes est identifiée comme bruyante, comme montré sur la figure 3.18 à gauche. De manière similaire une piste qui détecte moins d'un dixième du nombre de coups médian est identifiée comme déconnectée, comme montré sur la figure 3.18 à droite.

Le second point est la valeur d'efficacité moyenne par secteur de haute tension (voir schéma figure 3.6 (b)), en fonction de la valeur de haute tension. Ce point est particulièrement important puisqu'il correspond à l'observation du plateau d'efficacité et à la détermination du point de fonctionnement. Pour qualifier un secteur de haute tension comme étant efficace, il faut observer un plateau d'efficacité à une valeur supérieure à 98 %. Le point de fonctionnement est alors la tension d'amplification pour laquelle l'efficacité est de 97 %. En pratique, la plupart des modules NSW produits par les différents instituts jusqu'ici présentent des problèmes de tenue en tension, c'est pourquoi le point de fonctionnement désormais préconisé est de 570 V.

Une carte des efficacités en deux dimensions est également requise. Une efficacité de

plus de 90 % sur plus de 85 % de la surface du détecteur est requise.

Enfin, le gain doit être homogène sur toute la surface du module. Il faut que les valeurs de gain soient comprises dans un intervalle de $\pm 50\%$ de la valeur moyenne, et ceci par secteur haute tension.

Les modules produits sont encore en phase de prototypage, ainsi les tests se sont concentrés sur l'observation du plateau d'efficacité, et l'étude du gain des détecteurs. Ces points semblaient particulièrement importants pour le lancement de la production des modules afin de s'assurer de leur bon fonctionnement. D'autres tests plus extensifs et plus systématiques seront mis en place quand la production entrera en régime permanent.

3.5 Caractérisation du Module 1

Le premier quadruplet de détection assemblé à l'Irfu est appelé « Module 1 ». Les techniques de nettoyage avant assemblage n'étant pas parfaitement définies lors de sa construction, le Module 1 était susceptible de contenir des impuretés. Après les premiers tests sur le banc cosmique il a été décidé de nettoyer à nouveau le module, puis de tester un doublet constitué des deux couches « ETA » nettoyées.

Cette section présente dans un premier temps les performances du Module 1 lors de sa première installation sur le banc cosmique. Puis les études effectuées sur le doublet « ETA » après nettoyage sont présentées. Lors de ces études, seul un PCB c'est-à-dire un cinquième du détecteur a été mis en fonctionnement et a été placé entre les plans de détection du trajectographe externe. Il s'agit du PCB le plus petit lors de l'étude avant nettoyage, et du PCB le plus grand après nettoyage. Ces zones ont été choisies comme étant les plus stables lors de tests préliminaires de tenue en tension.

3.5.1 Caractérisation du Module 1 avant nettoyage

L'objectif avec le Module 1 a été d'observer un plateau d'efficacité pour s'assurer qu'il pouvait fonctionner dans un régime respectant le cahier des charges.

Plusieurs prises de données ont donc été effectuées, en faisant varier tantôt la tension d'amplification, tantôt la tension de dérive, tout en gardant l'autre à une valeur fixée. Les couches « ETA » ont présenté des problèmes de tenue en tension, c'est pourquoi les tensions d'amplification des couches « ETA » étaient systématiquement inférieures à celles des couches « STEREO ».

La figure 3.19 montre les efficacités obtenues en modifiant les tensions d'amplification. La tension de dérive est fixée à 400 V. Les points les plus à droite de la figure représentent les tensions d'amplification maximales supportées par les différentes couches.

L'efficacité de chaque couche est supposée augmenter avec la tension d'amplification jusqu'à atteindre un plateau, comme expliqué dans la section 3.4.3. Dans notre cas on s'attend à un plateau au dessus de 95 %. La figure montre que les couches « ETA » ne

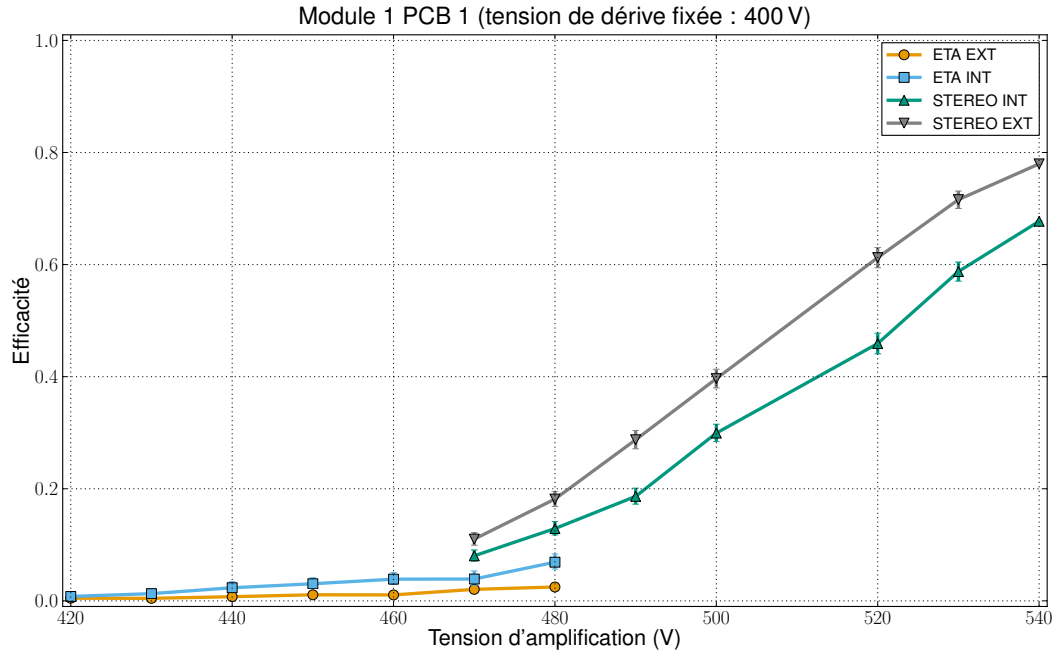


Figure 3.19 – Courbes d'efficacité des différentes couches du Module 1 avant nettoyage, en fonction de la tension d'amplification. La tension de dérive est fixée à 400 V. Seule l'incertitude statistique sur l'efficacité est présente.

sont pas efficaces aux tensions accessibles, mais les couches « STEREO » présentent un comportement attendu, sans toutefois atteindre un plateau. Il faudrait pouvoir appliquer une tension d'amplification plus élevée pour vérifier son existence.

L'étude suivante consiste à faire varier la tension de dérive, tout en gardant la tension d'amplification fixée à 520 V pour les couches « STEREO », et 470 V pour les couches « ETA ». Les valeurs d'efficacité ainsi obtenues sont représentées sur la figure 3.20.

On s'intéresse ici à l'efficacité en transmission électronique du module. En effet, pour un champ d'amplification donné :

- si le champ de dérive est très faible, des électrons sont capturés par les atomes du gaz et ne peuvent pas être acheminés vers l'espace de dérive ;
- si le champ de dérive est très élevé, les lignes de champ sont trop droites et ne contournent pas suffisamment le maillage métallique séparant la zone de dérive de la zone d'amplification. Des électrons seront perdus par collision avec le maillage.

La grandeur qui conditionne ce comportement est le rapport entre les champs de dérive et d'amplification. C'est pourquoi dans cette étude la tension d'amplification est fixée et la tension de dérive varie.

On observe bien le comportement attendu, c'est-à-dire un maximum et des baisses d'efficacité pour les petites et grandes valeurs de tension de dérive. La tension de dérive

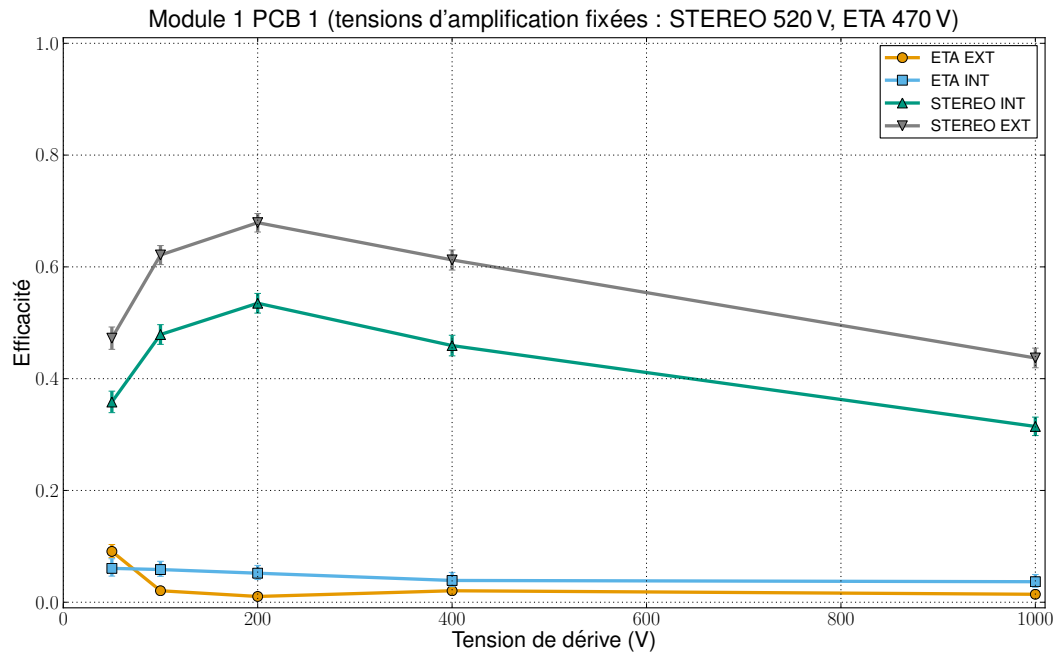


Figure 3.20 – Courbes d'efficacité des différentes couches du Module 1 avant nettoyage, en fonction de la tension de dérive. La tension d'amplification est fixée à 520 V pour les couches « STEREO », et 470 V pour les couches « ETA ». Seule l'incertitude statistique sur l'efficacité est présente.

optimale pour une tension d'amplification de 520 V semble se situer autour de 200 V. Si la tension d'amplification était nominale, soit environ 570 V, la tension de dérive optimale serait décalée vers la droite de la figure afin de garder un rapport des champs de dérive et d'amplification constant. La tension de dérive préconisée pour les modules est de 300 V, ce qui semble donc être un bon choix au vu de ce résultat.

3.5.2 Caractérisation du doublet « ETA » après nettoyage

Les premiers résultats obtenus après l'assemblage du Module 1 montrent un comportement attendu pour les couches « STEREO », toutefois la tension d'amplification maximale atteignable ne permet pas un gain suffisant pour atteindre le plateau d'efficacité éventuel. Les couches « ETA » qui présentent une efficacité quasi nulle ont été moins nettoyées lors de la fabrication, ce qui pousse à s'intéresser à la procédure de nettoyage.

Il a donc été décidé de démonter le Module 1, de nettoyer les deux couches « ETA », et des les ré-assembler pour en faire un doublet. Ce doublet a pu être installé et testé sur le banc cosmique. On peut donc bien apprécier l'impact de ce nettoyage sur les caractéristiques des couches « ETA », non fonctionnelles avant nettoyage.

Afin d’observer le plateau d’efficacité, plusieurs prises de données ont été effectuées avec différents mélanges de gaz circulant à l’intérieur du module. Tous ces mélanges ont pour composant principal l’argon, qui est complété par du CO_2 et/ou de l’isobutane $\text{HC}(\text{CH}_3)_3$. L’avantage de l’isobutane par rapport au CO_2 est qu’il génère plus d’électrons primaires au passage d’une particule chargée, permettant ainsi d’obtenir un signal plus élevé pour un même champ d’amplification. Il est toutefois inflammable et présente donc un danger à l’utilisation, en particulier pour de grands détecteurs. Un mélange d’argon et d’isobutane à un taux inférieur ou égal à 2,3 % n’est cependant pas inflammable et peut présenter une alternative intéressante à un mélange argon/ CO_2 .

La liste des mélanges de gaz qui ont été testés est la suivante :

- 93 % Ar, 7 % CO_2 (nominal) ;
- 70 % Ar, 30 % CO_2 ;
- 91 % Ar, 7 % CO_2 , 2 % $\text{HC}(\text{CH}_3)_3$;
- 95 % Ar, 5 % $\text{HC}(\text{CH}_3)_3$;
- 98 % Ar, 2 % $\text{HC}(\text{CH}_3)_3$.

Études de gain

Les différents gaz utilisés lors de cette étude permettent d’atteindre des valeurs de gain très différentes. Ce gain peut être estimé à l’aide de l’équation 3.5. Les estimations du gain ainsi obtenues sont représentées en fonction de la tension d’amplification sur la figure 3.21. La prise de données avec 30 % de CO_2 n’est pas représentée ici car les efficacités obtenues avec ce mélange sont trop faibles, rendant l’ajustement de la loi de Landau impossible par manque d’amas reconstruits.

Les valeurs de gain correspondant au mélange de gaz nominal s’étendent d’environ 1 300 à environ 3 400. Ce résultat est nettement meilleur que celui obtenu avant nettoyage grâce à l’augmentation de la tension d’amplification maximale applicable. Le gain maximal obtenu avant nettoyage est d’environ 1 000 pour une tension d’amplification de 540 V.

Les valeurs de gain obtenues avec les mélanges contenant de l’isobutane sont plus élevées que dans le cas du mélange avec CO_2 uniquement. De plus, selon la proportion d’isobutane utilisée, ainsi que l’ajout ou non de CO_2 , des gains de l’ordre de 5 000 peuvent être obtenus à des tensions d’amplification inférieures à la valeur nominale de 570 V. Par exemple avec le mélange contenant 7 % de CO_2 et 2 % d’isobutane, on peut obtenir un gain similaire à celui du gaz nominal avec une tension de 550 V au lieu de 610 V.

On peut qualitativement comparer ces résultats de gain avec la figure 3.15. Les méthodes d’estimation du gain sont différentes dans les deux cas, ainsi que les détecteurs utilisés (voir section 3.4.3). Certaines caractéristiques sont toutefois comparables. L’augmentation exponentielle du gain avec la tension d’amplification est bien vérifiée dans les deux cas, ainsi que les valeurs de gain plus élevées pour les mélanges contenant de l’isobu-

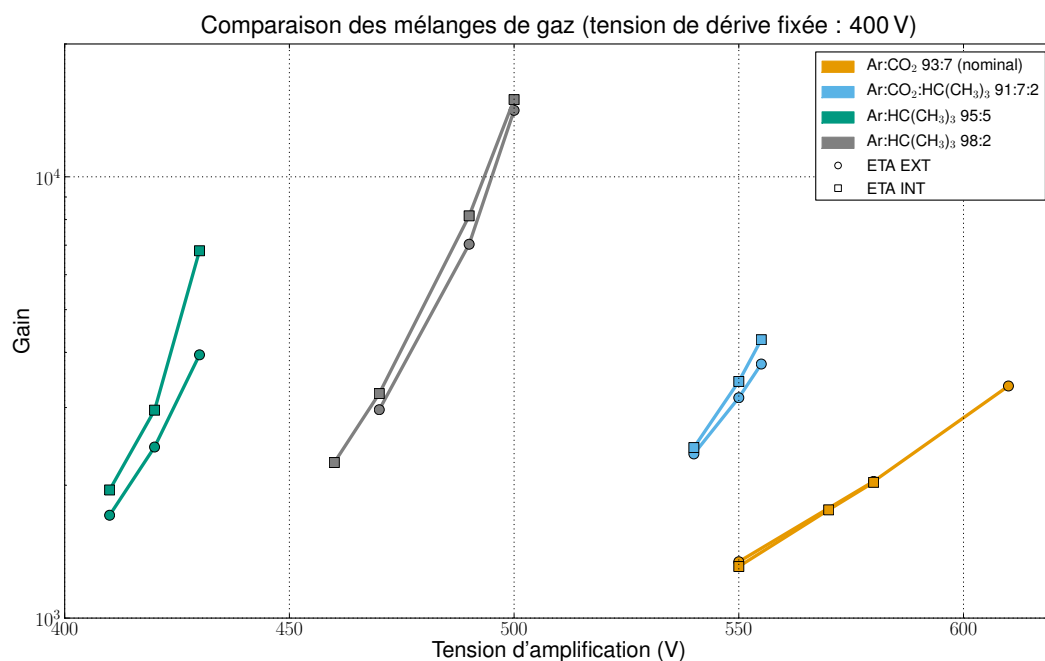


Figure 3.21 – Comparaison du gain obtenu avec le doublet ETA du Module 1 en utilisant différents mélanges de gaz, en fonction de la tension d'amplification. La tension de dérive est fixée à 400 V. Chaque couleur représente un mélange différent, et on peut observer les deux couches de détection ETA pour chacune de ces couleurs.

tane par rapport aux mélanges contenant du CO₂, et ceci à des tensions d'amplification inférieures.

Études d'efficacité

L'efficacité globale des différentes couches a été calculée pour chacune des prises de données avec le mélange de gaz utilisé. La méthode présentée dans la section 3.4.3 est utilisée. L'efficacité obtenue est représentée sur la figure 3.22.

On observe une zone d'efficacité constante pour chacun des gaz testés, sauf le mélange à 30 % de CO₂ pour lequel l'efficacité maximale obtenue est de 20 ou 30 % selon la couche considérée. Il semble donc qu'un plateau d'efficacité soit atteint, à une tension d'amplification d'environ 570 V dans le cas du gaz nominal, et à des tensions inférieures dans le cas des mélanges contenant de l'isobutane. L'efficacité atteinte sur le plateau est d'environ 96 %. Dans le cas du mélange à 30 % de CO₂, une tension d'amplification supérieure à 610 V aurait dû être appliquée afin d'être en mesure d'observer le plateau.

La figure montre également certains points avec des efficacités entre 45 et 50 % pour la couche « ETA EXT ». Chaque piste est en effet alimentée en haute tension à chaque

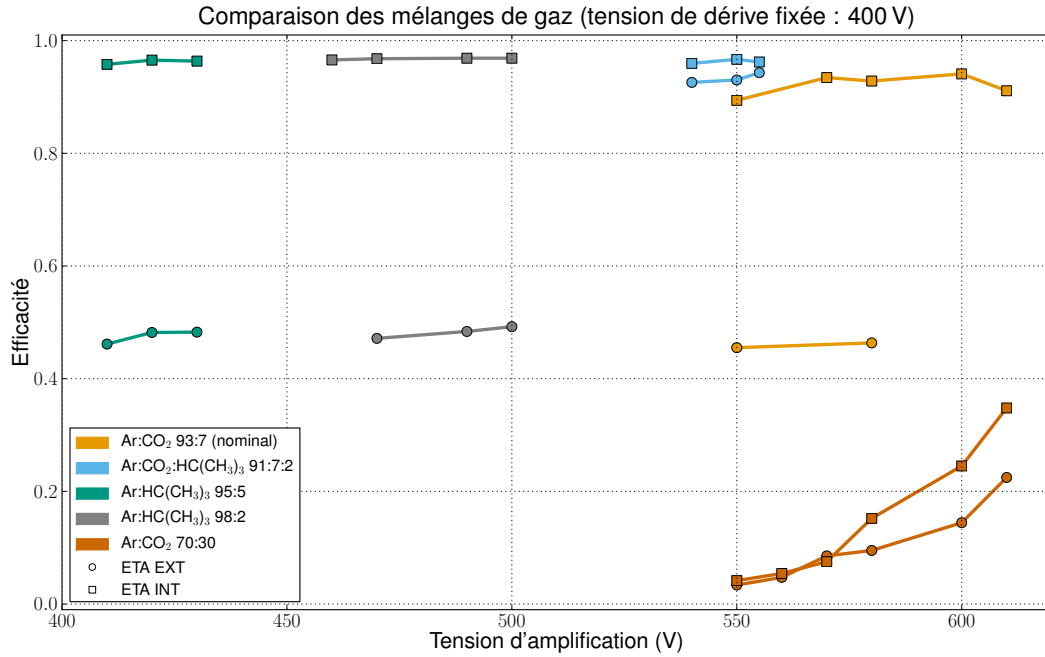


Figure 3.22 – Comparaison de l’efficacité obtenue avec le doublet ETA du Module 1 après nettoyage en utilisant différents mélanges de gaz, en fonction de la tension d’amplification. La tension de dérive est fixée à 400 V. Chaque couleur représente un mélange différent, et on peut observer les deux couches de détection ETA pour chacune de ces couleurs.

extrémité de manière indépendante, et dans le cas de ces points, l’un des deux côtés n’a pu maintenir la tension appliquée, rendant une moitié de détecteur inefficace. En ignorant ces moitiés de détecteur inefficaces suite à un problème d’alimentation en haute tension, on obtient la figure 3.23. Cette figure montre bien que les couches alimentées à la même tension ont des efficacités similaires.

La figure 3.19 montre qu’une efficacité maximale d’environ 80 % est atteinte pour une tension d’amplification de 540 V avant nettoyage. Après nettoyage l’efficacité atteinte est supérieure grâce à l’augmentation de la tension d’amplification maximale applicable. De plus les couches « ETA » n’étaient pas fonctionnelles avant nettoyage, et atteignent le plateau d’efficacité après avoir été nettoyées, ce qui est une réussite.

Plusieurs zones de haute tension des modules produits montrent des problèmes de tenue en tension pour des valeurs inférieures à 600 V. Puisque la valeur minimale de tension dans le cas nominal semble se situer autour de 570 V, la marge de fonctionnement est très réduite. L’addition d’isobutane dans le gaz utilisé pourrait dégager une marge de manœuvre de plusieurs dizaines de volt. Ces résultats montrent qu’une modification du gaz nominal pourrait être envisagée par la collaboration, s’il est prouvé que cela n’engendrera pas de problèmes de vieillissement des chambres par exemple, ou encore si

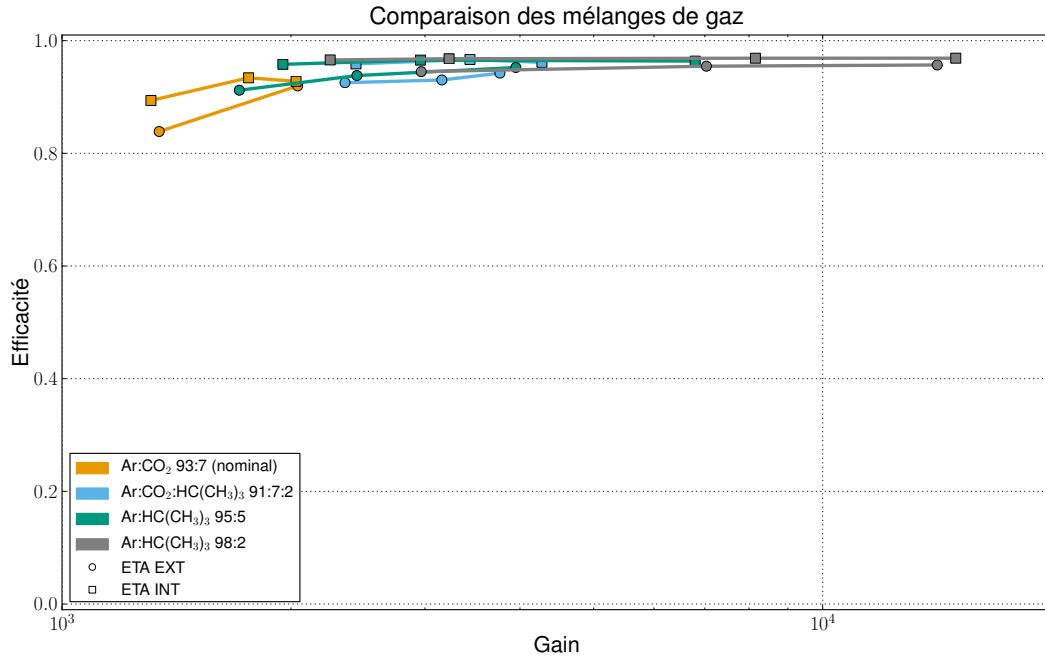


Figure 3.24 – Efficacité en fonction du gain pour les couches du doublet « ETA » du Module 1 après nettoyage (zones des haute tension trop basse ignorées). Chaque point représente une prise de données avec une tension d’amplification différente. La tension de dérive est fixée à 400 V. Pour chaque valeur de tension, les zones du détecteur ne pouvant la supporter ont été ignorées. Chaque couleur représente un mélange de gaz différent, et on peut observer les deux couches de détection ETA pour chacune de ces couleurs.

gain supérieur à 2 000 à une tension d’amplification de 540 V seulement. L’hypothèse de modification du mélange de gaz est étudiée par la collaboration. Des études similaires avec d’autres mélanges de gaz sont en cours de réalisation, et semblent favoriser les mélanges d’argon et de CO₂ avec une proportion de CO₂ d’environ 15 %. Ce type de mélange réduit le gain obtenu à une tension d’amplification donnée, mais permet d’augmenter la tension maximale applicable au détecteur, permettant in fine d’obtenir un gain maximal plus élevé.

3.6 Caractérisation du Module 2

Le deuxième module de pré-production « Module 2 » a été testé sur le banc cosmique. Des études d’efficacité et de gain en fonction des tensions d’amplification et de dérive ont été menées sur ce nouveau module, ainsi qu’une étude poussée de l’influence du trajectographe externe. Seule une surface limitée du Module a été testée, les résultats présentés dans cette section ne concernent donc pas le détecteur en entier.

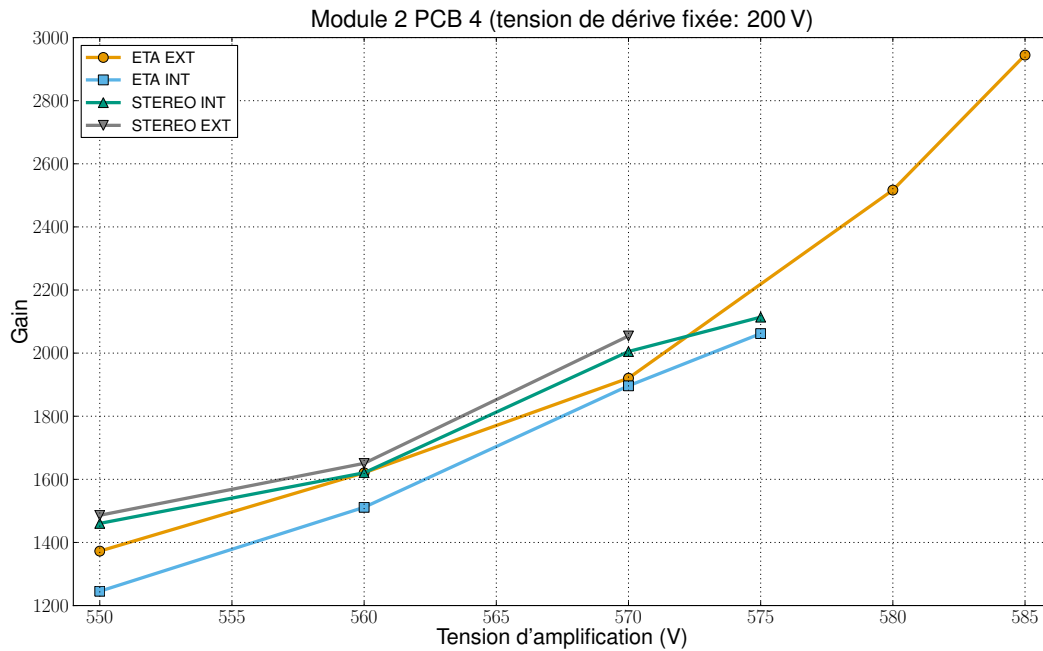


Figure 3.25 – Courbes de gain des différentes couches du Module 2, en fonction de la tension d'amplification. La tension de dérive est fixée à 200 V.

3.6.1 Influence de la tension d'amplification

La tension d'amplification maximale supportée par les couches du Module 2 se situe entre 575 V et 585 V selon la couche considérée. Il n'a pas été possible de poursuivre les études à des tensions plus élevées qui induisent des courants trop importants et donc des instabilités. Toutefois le point de fonctionnement de 570 V est bien atteignable par toutes les couches. On rappelle que si une moitié de circuit imprimé ne supporte pas la tension choisie pour une prise de données, elle est ignorée dans l'analyse. La figure 3.25 présente la valeurs de gain obtenues dans les couches du Module 2 en fonction de la tension d'amplification appliquée. Toutes les couches atteignent un gain supérieur à 2000, valeur qui semblait conditionner l'accès au plateau d'efficacité pour le Module 1 (voir section 3.5.2). Cette valeur de gain est atteinte pour une tension d'environ 575 V, ce qui est cohérent avec les observations effectuées sur le Module 1 sur la figure 3.21.

La figure 3.26 montre les valeurs d'efficacité obtenues lors des mêmes prises de données. Elle ne permet pas d'affirmer avec certitude qu'un plateau d'efficacité est atteint pour les différentes couches, même au-delà de 570 V. Il faudrait pour cela poursuivre l'étude avec des tensions plus élevées. Une augmentation de l'efficacité avec la tension est cependant bien observable, avec un maximum se situant autour de 96 % pour les couches « ETA EXT » et « STEREO EXT ».

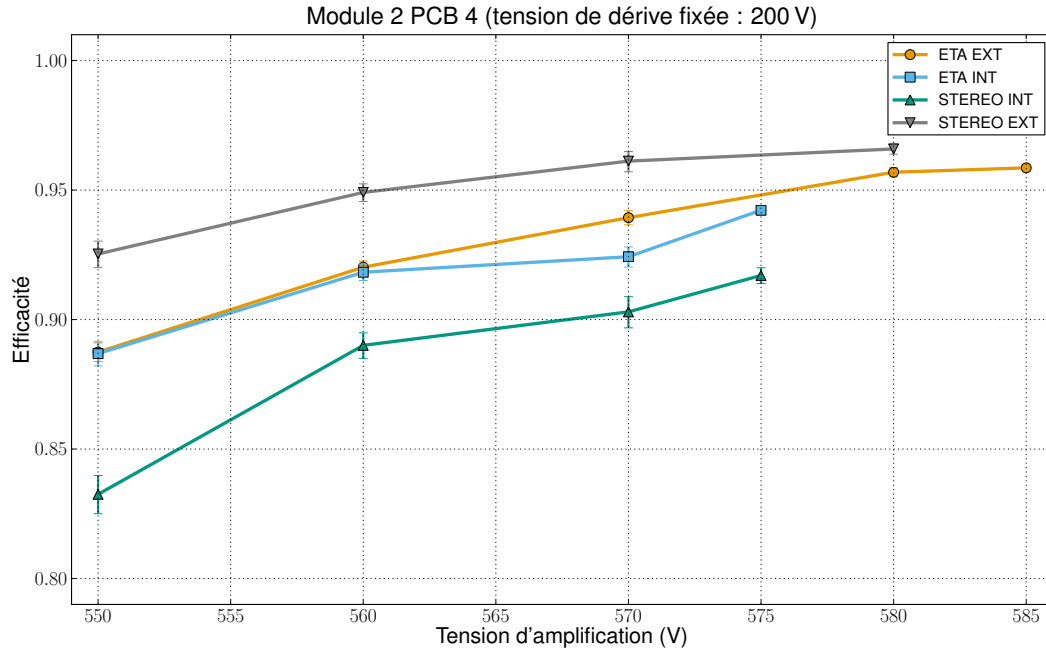


Figure 3.26 – Courbes d'efficacité des différentes couches du Module 2, en fonction de la tension d'amplification. La tension de dérivation est fixée à 200 V. Seule l'incertitude statistique sur l'efficacité est présente.

La valeur d'efficacité au point de fonctionnement quant à elle varie entre 90 % et 96 % selon la couche considérée. Cette différence s'explique du fait de défauts de fabrication, notamment du nombre de pistes non connectées différent pour chaque couche, ou encore du fait d'une mauvaise connexion au niveau du Zebra. La couche « STEREO EXT » par exemple ne possède que très peu de pistes non connectées. L'écart entre les courbes semble se réduire avec l'augmentation de la tension d'amplification donc avec un gain grandissant. En effet une augmentation du gain engendre des amas d'amplitude supérieure, leur permettant de mieux se démarquer du bruit de fond. La probabilité qu'un amas ne se démarque du bruit de fond sur aucune piste fonctionnelle diminue donc avec le gain.

Il faut donc noter que la tension d'amplification nominale de 570 V ne semble pas suffisante pour atteindre le plateau d'efficacité, bien qu'une efficacité de 96 % comme pour la couche « STEREO EXT » soit tout à fait acceptable. Par conséquent chaque couche devrait supporter une tension sensiblement plus élevée pour avoir une marge de sécurité raisonnable. C'est un point important d'amélioration pour les modules futurs, qui pourrait être obtenue en modifiant le mélange de gaz utilisé, comme présenté dans la section 3.5.2.

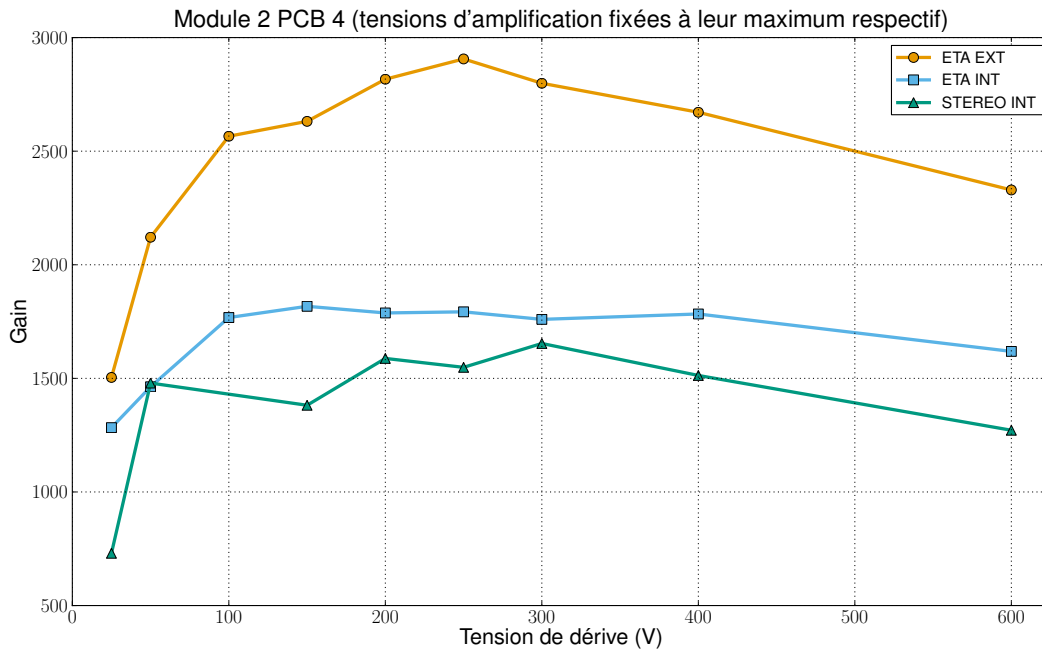


Figure 3.27 – Courbes de gain des différentes couches du Module 2, en fonction de la tension de dérivation. La tension d'amplification est fixe et dépend de la couche considérée : 585 V pour « ETA EXT », 570 V pour « ETA INT », et 560 V pour « STEREO INT ». Seule l'incertitude statistique sur l'efficacité est présente.

3.6.2 Influence de la tension de dérivation

D'autres prises de données ont été effectuées en faisant varier la tension de dérivation à tension d'amplification constante. Cette dernière varie entre les différentes couches puisqu'elle correspond à la tension maximale atteignable : 585 V pour la couche « ETA EXT », 570 V pour la couche « ETA INT », 560 V pour la couche « STEREO INT », et 510 V pour la couche « STEREO EXT ». Dans le cas des couches « ETA INT » et « STEREO INT », un demi circuit imprimé ne supportant pas la tension a été ignoré.

Les valeurs de gain obtenues lors de ces prises de données sont présentées sur la figure 3.27. La couche « STEREO EXT » est absente de la figure du fait de sa faible tension d'amplification. Les valeurs trop basses des amplitudes des amas ne permettent pas un ajustement de la distribution par une courbe de Landau. La forme des courbes présente un optimum entre 200 et 300 V, ce qui est attendu. Le gain d'une couche suit bien la valeur de sa tension d'amplification, le gain le plus élevé étant notamment atteint par « ETA EXT ».

Les efficacités correspondant à toutes ces prises de données sont montrées sur la figure 3.28. La différence entre les courbes provient toujours des valeurs différentes de

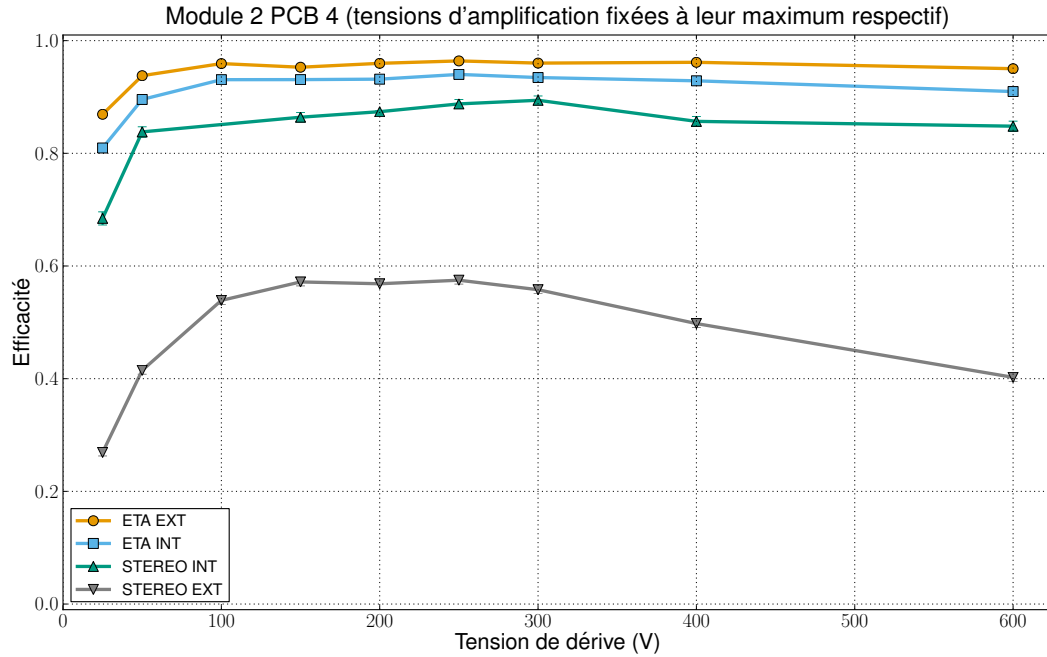


Figure 3.28 – Courbes d'efficacité des différentes couches du Module 2, en fonction de la tension de dérive. La tension d'amplification est fixe et dépend de la couche considérée : 585 V pour « ETA EXT », 570 V pour « ETA INT », 560 V pour « STEREO INT », et 510 V pour « STEREO EXT ». Seule l'incertitude statistique sur l'efficacité est présente.

tension d'amplification selon la couche considérée, la valeur d'efficacité augmentant avec celle de la tension d'amplification. En particulier, la couche « STEREO EXT » n'a pu être alimentée qu'à une tension d'amplification de 510 V ce qui explique les faibles valeurs d'efficacité associées. On observe toutefois bien la forme attendue des courbes avec un maximum situé autour du point de fonctionnement. Cette étude confirme la pertinence du point de fonctionnement standard pour une tension de dérive de 300 V, comme indiqué dans la section 3.2.2.

3.6.3 Étude de l'influence de la référence externe Mcube

Il a été mentionné dans la section 3.4.3 que plusieurs coupures sont appliquées sur les traces fournies par le trajectographe de référence externe Mcube. Ces coupures permettent d'ignorer certains événements pour le calcul de l'efficacité, ceux-ci contenant une trace de mauvaise qualité ou encore présentant un angle trop grand par rapport à l'horizontale. Les coupures ont donc pour but de s'assurer que les traces de référence sont fiables afin de biaiser les résultats le moins possible. L'influence de ces paramètres a été étudié et quelques résultats sont présentés dans cette section. Les coupures et leurs effets sont

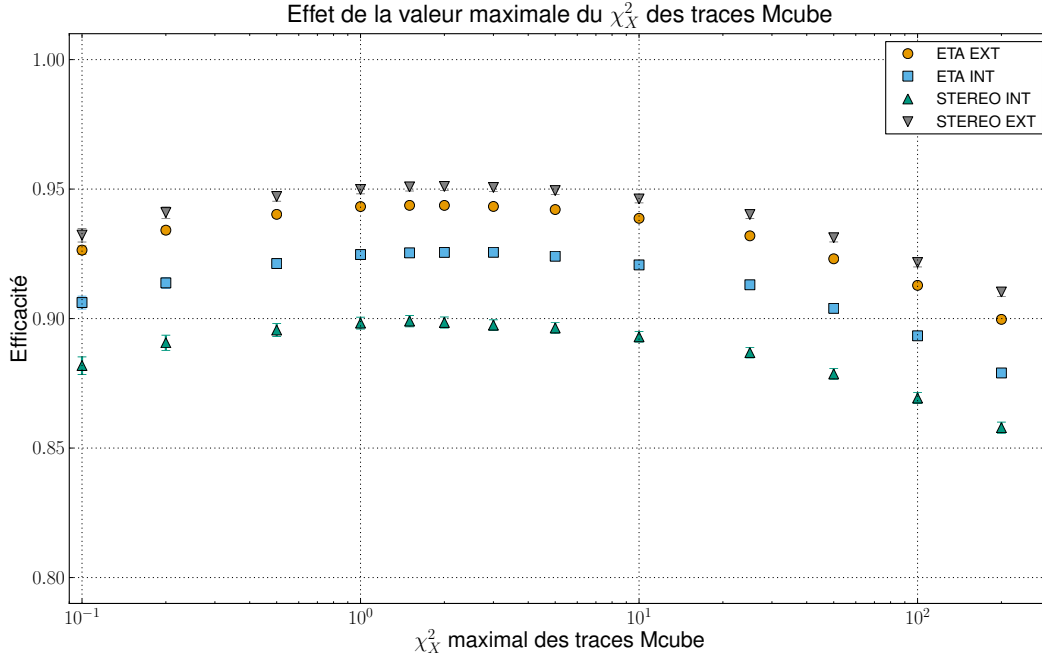


Figure 3.29 – Efficacité des couches du Module 2 en fonction de la valeur maximale de la variable χ^2_X . La tension d’amplification de chaque couche est la tension maximale atteignable. Seule l’incertitude statistique sur l’efficacité est présente.

résumés à la fin de cette section.

Les traces Mcube sont construites à partir de coups sur des pistes disposées selon deux axes orthogonaux. On a donc une information selon chacun de ces axes, nommés X et Y . Les coupures appliquées sur ces traces correspondent aux paramètres d’ajustement de ces dernières selon les axes X et Y , appelés χ^2_X et χ^2_Y , comme présenté dans la section 3.4.3.

Plusieurs estimations d’efficacité ont été menées pour une même prise de données, ceci en faisant varier la valeur maximale autorisée du paramètre d’ajustement χ^2_X . La prise de données choisie ici correspond aux tensions atteintes les plus élevées dans les différentes couches. La figure 3.29 montre l’efficacité des différentes couches du Module 2 en fonction de la valeur maximale de χ^2_X . La tension d’amplification est différente selon la couche considérée, ce qui explique l’écart entre les courbes.

L’efficacité baisse avec l’augmentation de la valeur maximale autorisée, puisque les traces les moins fiables ont une valeur élevée de χ^2_X . La valeur retenue pour la coupure est $\chi^2_X < 2.5$, valeur proche de l’optimum d’efficacité mais permettant également de conserver un maximum de traces pour impacter le moins possible la précision statistique des études effectuées avec ce réglage.

Une étude similaire est effectuée avec la valeur maximale autorisée du paramètre χ^2_Y . Les résultats sont présentés sur la figure 3.30. Il s’agit de la même prise de données

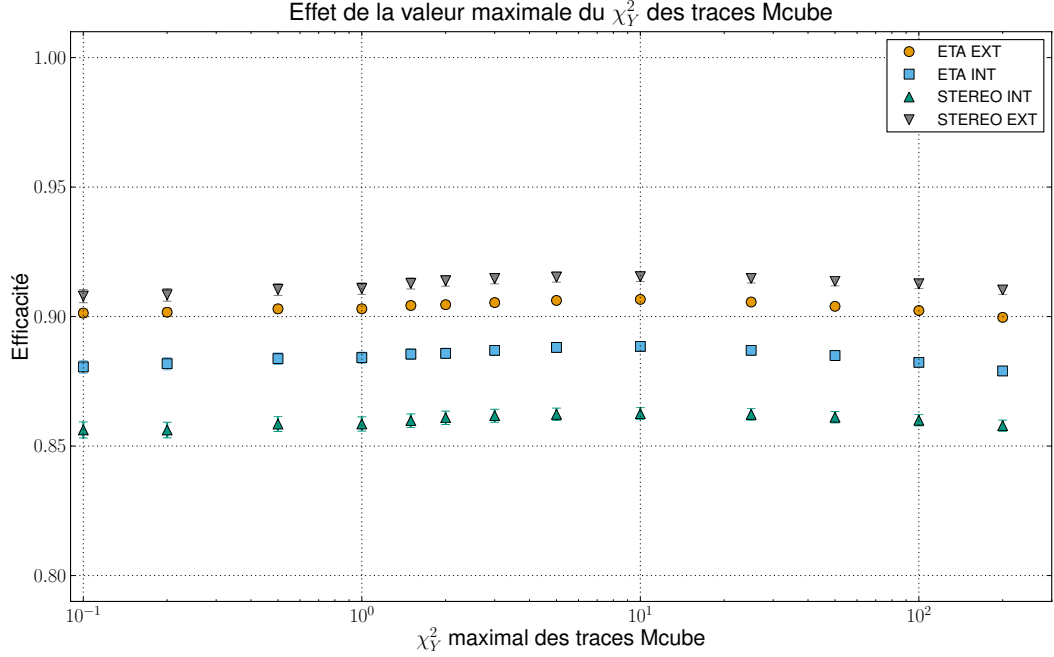


Figure 3.30 – Efficacité des couches du Module 2 en fonction de la valeur maximale de la variable χ^2_Y . La tension d’amplification de chaque couche est la tension maximale atteignable. Seule l’incertitude statistique sur l’efficacité est présente.

que pour la figure 3.29. Le choix de la valeur maximale du paramètre χ^2_Y n’a que peu d’influence sur l’efficacité des différentes couches. En effet la variation d’efficacité maximale observée est d’environ 1 %. La coupure standard $\chi^2_Y < 10$ est ainsi définie pour n’éliminer qu’un faible proportion des traces correspondant aux traces les moins fiables.

Le dernier paramètre d’influence est l’angle de la trace par rapport à la verticale, appelé α_z . Il est défini entre 0 et $+\frac{\pi}{2}$ puisqu’aucune particule chargée ne provient du bas du détecteur. La majorité des traces se situe dans l’intervalle $0 < \alpha_z < 0,8$, avec un maximum aux alentours de 0,3. Le maximum n’est pas situé à 0 car les traces très verticales génèrent moins de charges lors de leur traversée du trajectographe, et les amas d’une taille d’une piste seulement sont difficilement utilisables par Mcube, les pistes étant connectées par groupe de 16 avant de rejoindre l’électronique de lecture. La figure 3.31 montre l’efficacité des différentes couches du Module 2 en fonction de la coupure sur l’angle α_z des traces Mcube. Il s’agit de la même prise de données que pour la figure 3.29.

La figure montre une baisse d’efficacité aux faibles et aux grandes valeurs maximales de α_z . Ceci est attendu puisque les traces très verticales génèrent des amas plus petits donc plus difficilement détectables, et les traces très penchées souffrent d’une résolution dégradée du détecteur comme il a été expliqué dans la section 3.4.3. On observe également une augmentation de l’incertitude statistique vers la gauche de la figure, due à un nombre

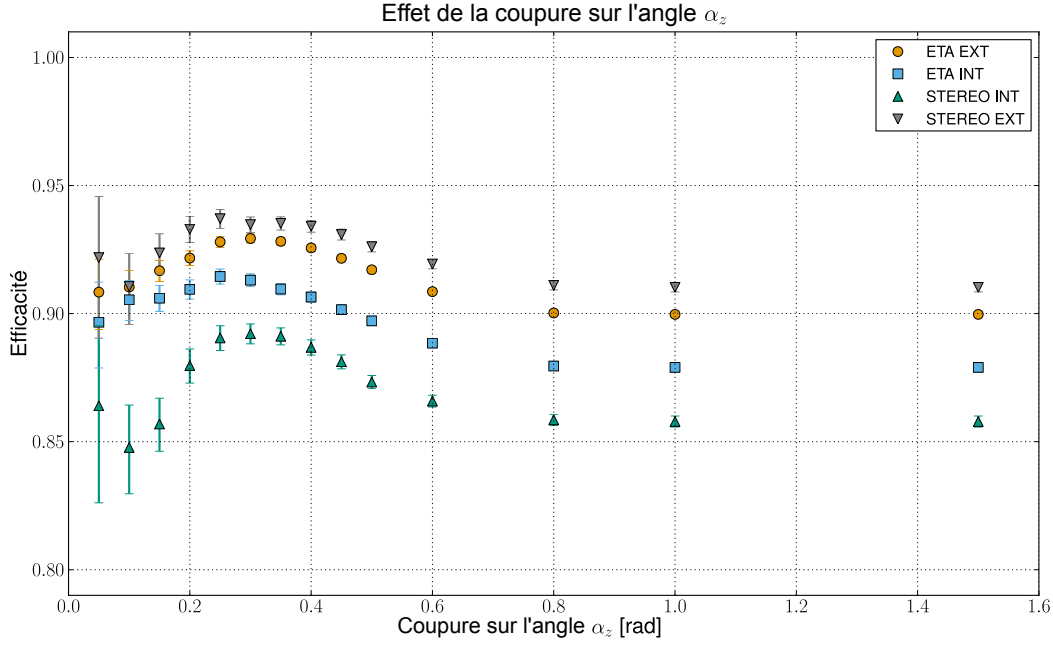


Figure 3.31 – Efficacité des couches du Module 2 en fonction de la coupure sur l'angle des traces par rapport à la verticale α_z . La tension d'amplification de chaque couche est la tension maximale atteignable. Seule l'incertitude statistique sur l'efficacité est présente.

d'événements rejeté d'autant plus grand que la coupure sur α_z est stricte. Il existe donc un optimum qui semble se situer aux alentours de la valeur 0,3. La coupure retenue est cependant $\alpha_z < 0,4$, puisque cette valeur permet d'atteindre une efficacité presque identique et de conserver davantage d'événements, réduisant ainsi l'incertitude globale des études effectuées avec ce réglage.

En résumé, les coupures effectuées sont les suivantes :

- $\chi_X^2 < 2.5$ qui élimine typiquement 20 à 40 % des événements ;
- $\chi_Y^2 < 10$ qui élimine typiquement 10 à 30 % des événements ;
- $\alpha_z < 0,4$ qui élimine typiquement 40 à 50 % des événements.

Les trois coupures combinées éliminent environ 60 à 70 % des événements. Il s'agit donc de coupures assez strictes, mais le nombre d'événements restant est suffisant et il a été choisi de privilégier les traces fiables et surtout faiblement inclinées.

3.6.4 Comparaison avec une méthode de reconstruction de traces interne

Il est possible d'estimer l'efficacité d'un plan de détection en reconstruisant une trace de référence à partir des trois autres plans. Cette méthode a été mentionnée dans la

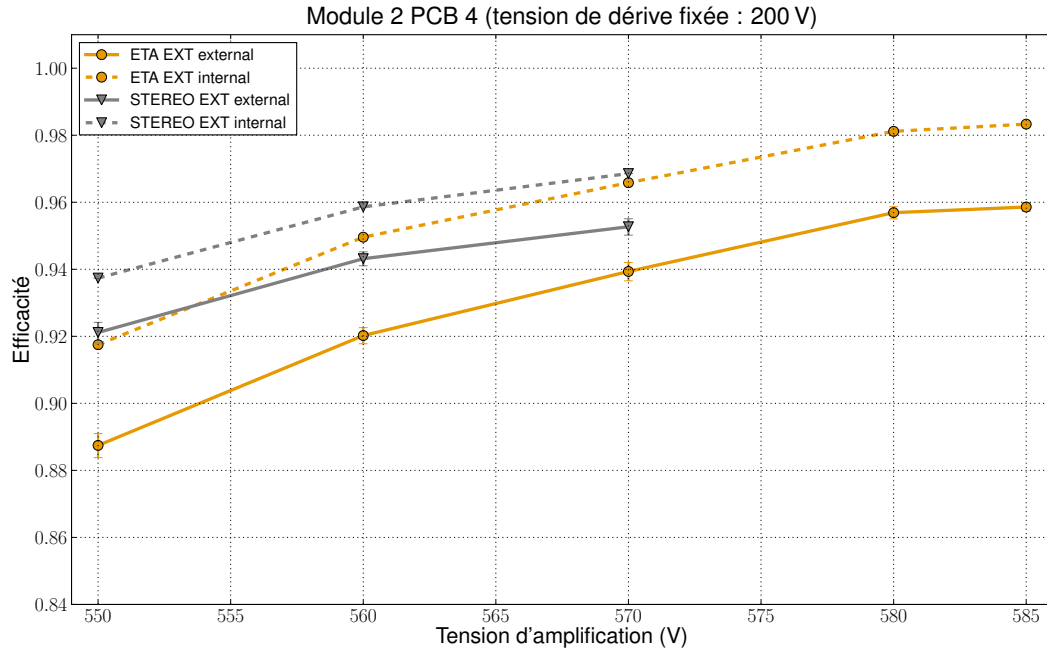


Figure 3.32 – Courbes d'efficacité externe (trait plein) et interne (trait pointillé) de différentes couches du Module 2, en fonction de la tension d'amplification. La tension de dérivation est fixée à 200 V. Seule l'incertitude statistique sur l'efficacité obtenue avec la méthode externe est présente.

section 3.4.3 et présente l'avantage de s'affranchir d'une référence externe telle que celle utilisée dans ce travail. Cependant il est nécessaire que les différents plans fonctionnent raisonnablement bien pour que la reconstruction de traces de référence soit fiable.

Les bonnes performances du Module 2 ont motivé l'utilisation de cette méthode qui pourrait être pérennisée dans l'optique de la production des modules suivants. Partant des amas reconstruits comme expliqué dans la section 3.4.2, une méthode de reconstruction de traces interne a été implémentée. Les efficacités précédemment montrées sont donc comparées aux résultats de cette nouvelle méthode sur les figures 3.32 et 3.33.

À ces fins, seules les couches « ETA EXT » et « STEREO EXT » sont considérées. La méthode ne permet pas d'obtenir une valeur d'efficacité fiable pour les autres couches au vu des conditions de prise de données. Certaines moitiés de circuit imprimé ne supportent en effet pas une tension suffisante, rendant impossible la reconstruction de traces de référence.

La figure 3.32 compare les deux méthodes d'estimation de l'efficacité, appelées « externe » et « interne », en fonction de la tension d'amplification similairement à la figure 3.26. De même, la figure 3.33 compare les deux méthodes en fonction de la tension de dérivation similairement à la figure 3.28. L'efficacité en fonction de la tension de dérivation est très

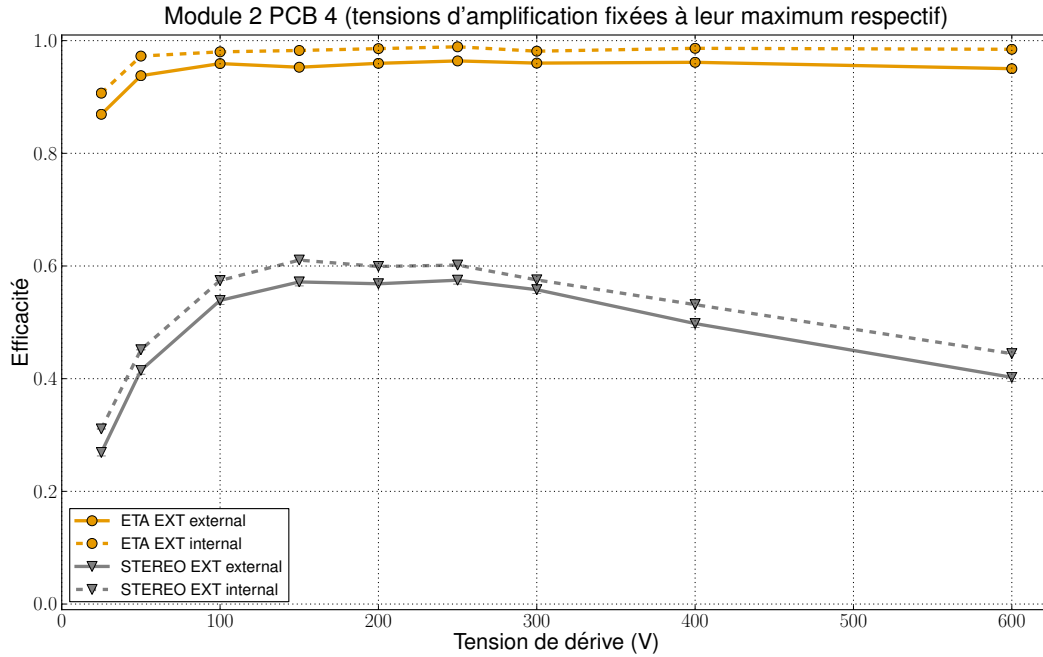


Figure 3.33 – Courbes d'efficacité externe (trait plein) et interne (trait pointillé) de différentes couches du Module 2, en fonction de la tension de dérive. La tension d'amplification est fixe et dépend de la couche considérée : 585 V pour « ETA EXT », et 510 V pour « STEREO EXT ». Seule l'incertitude statistique sur l'efficacité obtenue avec la méthode externe est présente.

différente pour les deux couches puisque la tension d'amplification est très différente : 585 V pour « ETA EXT », et 510 V pour « STEREO EXT ».

Ces figures montrent un bon accord entre les deux méthodes représentées pour chacun des points de mesure. Il est à noter toutefois un écart systématique d'environ 2 % entre les deux courbes : la méthode interne donne un résultat supérieur à la méthode externe. Ceci est tout à fait compréhensible puisque la méthode interne prend uniquement en compte les événements pour lesquels toutes les couches de référence ont détecté un amas, rendant la reconstruction de trace possible. Les événements pour lesquels au moins une couche ne possède pas d'amas reconstruit sont donc ignorés pour le calcul, et ils correspondent potentiellement à une inefficacité pour la méthode externe. La méthode interne sur-estime donc l'efficacité. Inversement la méthode externe sous-estime l'efficacité du fait de la coupure sur la valeur des résidus comme expliqué dans la section 3.4.3, ainsi que la sélection imparfaite des traces Mcube conformes à la réalité, générant une part d'inefficacité qui n'a pas lieu d'être. Ainsi l'efficacité réelle se situe entre les estimations des deux méthodes.

La méthode externe reste pour le moment la méthode de référence, mais des études

plus poussées sur la méthode interne pourraient permettre de s'affranchir de la référence externe pour l'analyse des modules futurs. Il serait notamment souhaitable de tester la méthode sur une portion de module possédant quatre plans pleinement fonctionnels, ce qui est impossible avec le Module 2. Cette nouvelle méthode permettrait d'être en mesure de caractériser un module entier avec une seule prise de données et de ne pas devoir le déplacer du fait de taille limitée du trajectographe externe. Ceci faciliterait et accélérerait grandement l'analyse d'un module, mais nécessite que ses quatre couches soient fonctionnelles.

3.7 Conclusion et perspectives

Une chaîne de caractérisation des modules LM1 pour le projet NSW a été mise en place à l'Irfu. Ce travail de thèse s'est concentré sur l'élaboration d'un code d'analyse de données du banc cosmique destiné à caractériser ces modules. Des coups puis des amas sont reconstruits à partir des amplitudes des signaux enregistrés sur les différentes pistes du module étudié, et leur position est comparée à une référence externe afin d'estimer notamment le gain et l'efficacité du module.

Les premiers modules de production « Module 1 » et « Module 2 » ont été caractérisés en détail. Les résultats obtenus sont cohérents mais souffrent d'une trop grande quantité de secteurs de haute tension ne supportant pas la tension d'amplification nominale. La caractérisation du Module 2 montre cependant une nette amélioration sur ce point par rapport au Module 1, ce qui est encourageant pour la suite de la production. Une partie de la solution à ce problème pourrait être de modifier la composition du gaz utilisé, comme il a été présenté dans ce travail.

Il reste encore environ 30 modules à produire dans les mois à venir, ce qui va nécessiter plusieurs améliorations ainsi qu'une plus grande automatisation de la chaîne de caractérisation. L'estimation de l'efficacité d'un module pourrait notamment être effectuée en interne sans référence externe. Une version préliminaire de cette méthode a été comparée à la méthode standard, mais cette étude est limitée du fait du manque d'efficacité de certaines couches. Il est donc nécessaire de répéter l'étude sur un module possédant quatre plans pleinement fonctionnels.

Le code d'analyse développé dans ce travail est en ce moment utilisé au Département de Physique des Particules pour caractériser les nouveaux modules produits. À titre d'illustration, la figure 3.34 représente l'efficacité du Module 4 en deux dimensions. Les extrémités du module sont coupées sur l'axe des ordonnées puisque la largeur du trajectographe externe est inférieure à celle du module. On peut remarquer une efficacité proche de 100 % sur la majeure partie du module, mais aussi plusieurs secteurs de haute tension avec une efficacité très faible. Ces secteurs ne supportent pas une haute tension suffisante pour être pleinement efficaces.

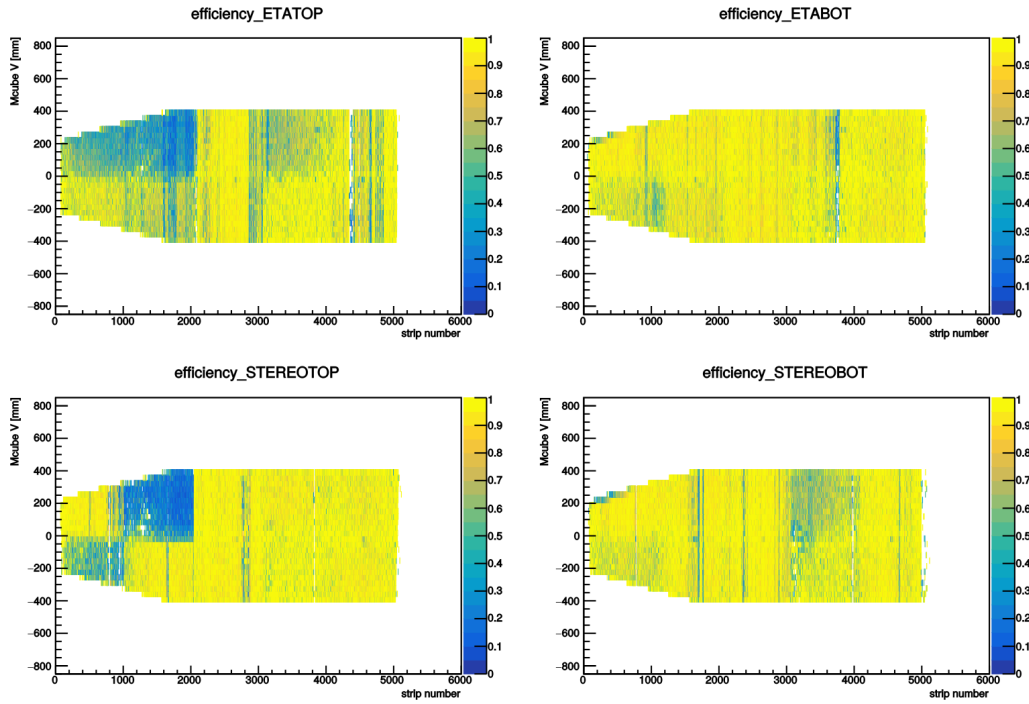


Figure 3.34 – Efficacité des quatre couches du Module 4 en deux dimensions. De gauche à droite puis de haut en bas, les couches sont d’altitude décroissante.

Le résultat montré sur la figure 3.34 a été obtenu en utilisant des cartes électroniques multiplexées. Ces cartes relient les pistes deux à deux afin de réduire la quantité d’électronique nécessaire à une prise de données par deux. L’analyse de données multiplexées est plus délicate que l’analyse standard [20].

Références

- [1] G APOLLINARI et al. *High-Luminosity Large Hadron Collider (HL-LHC) Preliminary Design Report*. 2015. DOI : 10.5170/cern-2015-005.
- [2] ATLAS COLLABORATION. *CERN-LHCC-2013-006 / ATLAS-TDR*. Rapp. tech. 2013. URL : <https://cds.cern.ch/record/1552862>.
- [3] S. MAJEWSKI et al. “A thin multiwire chamber operating in the high multiplication mode”. In : *Nuclear Instruments and Methods in Physics Research* 217.1-2 (nov. 1983), p. 265-271. DOI : 10.1016/0167-5087(83)90146-1.
- [4] G. MIKENBERG. “Thin-gap gas chambers for hadronic calorimetry”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 265.1-2 (mar. 1988), p. 223-227. DOI : 10.1016/0168-9002(88)91075-3.
- [5] Fabien JEANNEAU. “Design and Construction of Large Size Micromegas Chambers for the ATLAS Upgrade of the Muon Spectrometer”. In : *IEEE Transactions on Nuclear Science* 63.4 (août 2016), p. 2336-2342. DOI : 10.1109/tns.2016.2590601.
- [6] Y. GIOMATARIS et al. “MICROMEGAS : a high-granularity position-sensitive gaseous detector for high particle-flux environments”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 376.1 (juin 1996), p. 29-35. DOI : 10.1016/0168-9002(96)00175-1.
- [7] Fabio SAULI. *Gaseous Radiation Detectors : Fundamentals and Applications*. T. 36. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2014, pp.1-497. ISBN : 9781107043015.
- [8] M. DIXIT et al. “Micromegas TPC studies at high magnetic fields using the charge dispersion signal”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 581.1-2 (oct. 2007), p. 254-257. DOI : 10.1016/j.nima.2007.07.099.
- [9] T. ALEXOPOULOS et al. “A spark-resistant bulk-micromegas chamber for high-rate applications”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 640.1 (juin 2011), p. 110-118. DOI : 10.1016/j.nima.2011.03.025.
- [10] G. SEKHNIADZE. “Construction and performance of the sTGC and MicroMegas chambers for ATLAS NSW upgrade”. In : *Journal of Instrumentation* 12.03 (mar. 2017), p. C03020-C03020. DOI : 10.1088/1748-0221/12/03/c03020.

-
- [11] G. De GERONIMO et al. “VMM1—An ASIC for Micropattern Detectors”. In : *IEEE Transactions on Nuclear Science* 60.3 (juin 2013), p. 2314-2321. DOI : 10.1109/tns.2013.2258683.
 - [12] D. ATTIE et al. “The readout system for the Clas12 Micromegas vertex tracker”. In : *2014 19th IEEE-NPSS Real Time Conference*. IEEE, mai 2014. DOI : 10.1109/rtc.2014.7097517.
 - [13] C. FLOUZAT et al. “Dream : a 64-channel Front-end Chip with Analog Trigger Latency Buffer for the Micromegas Tracker of the CLAS12 Experiment”. In : *TWEPP 2014* (2014).
 - [14] ZEBRA[®] *Elastomeric Connectors data sheet*. URL : <https://www.es.co.th/Schemetic/PDF/ZEBRA-CONN.PDF>.
 - [15] Simon BOUTEILLE. “Development and applications of micro-pattern gaseous detectors for muon tomography”. Thèse de doct. Université Paris-Saclay, 2017.
 - [16] S. BOUTEILLE et al. “Large resistive 2D Micromegas with genetic multiplexing and some imaging applications”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 834 (oct. 2016), p. 187-191. DOI : 10.1016/j.nima.2016.08.006.
 - [17] T. ALEXOPOULOS et al. “Development of large size Micromegas detector for the upgrade of the ATLAS Muon system”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 617.1-3 (mai 2010), p. 161-165. DOI : 10.1016/j.nima.2009.06.113.
 - [18] Archana SHARMA. “Properties of some gas mixtures used in tracking detectors”. In : (1998). URL : <https://www.slac.stanford.edu/pubs/icfa/summer98/paper3/paper3a.html>.
 - [19] D. ATTIE. “TPC review”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 598.1 (jan. 2009), p. 89-93. DOI : 10.1016/j.nima.2008.08.114.
 - [20] S. PROCUREUR, R. DUPRÉ et S. AUNE. “Genetic multiplexing and first results with a $50 \times 50 \text{ cm}^2$ Micromegas”. In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 729 (nov. 2013), p. 888-894. DOI : 10.1016/j.nima.2013.08.071.

4

Le processus de production de quatre quarks top

Le modèle standard de la physique des particules détaillé au chapitre 1 permet de décrire les particules élémentaires constitutives de la matière ainsi que leurs interactions. Plusieurs limitations de ce modèle comme l'absence d'une description cohérente de l'interaction gravitationnelle poussent à développer de nouveaux modèles ou bien des extensions du modèle standard, comme expliqué dans le chapitre 1. Ils sont qualifiés de « Nouvelle Physique » ou encore de théories au-delà du modèle standard, ou *Beyond Standard Model* (BSM). À l'heure actuelle aucun de ces modèles de nouvelle physique n'a été suffisamment conforté par l'expérience et la physique au-delà du modèle standard reste une question ouverte.

Le quark top est la particule élémentaire la plus massive connue à ce jour : $m_t = 173,0 \pm 0,4 \text{ GeV}$ [1]. Son couplage de Yukawa au boson de Higgs est proche de 1, ce qui fait de lui un élément-clé de nombreux modèles de nouvelle physique. En particulier, la section efficace du processus de production de quatre quarks top $\sigma_{t\bar{t}t\bar{t}}$ est augmentée dans de nombreux modèles BSM. On s'intéresse notamment à des modèles faisant apparaître de nouvelles particules se désintégrant préférentiellement en quarks top : un modèle de dimensions supplémentaires (2UED) et un modèle à deux doublets de Higgs (2HDM), mais aussi à un modèle développé dans le cadre de la théorie des champs effectifs (CI). La mesure de $\sigma_{t\bar{t}t\bar{t}}$ est donc une sonde pour ces modèles, mais permet également de tester le modèle standard dans un environnement complexe et très énergétique. L'existence du diagramme de gauche de la figure 4.4 ainsi que d'autres diagrammes similaires permet notamment de contraindre le couplage de Yukawa du quark top via la mesure de $\sigma_{t\bar{t}t\bar{t}}$, et ainsi de placer des contraintes sur la largeur du boson de Higgs en combinant ce résultat avec la mesure de la section efficace du processus $pp \rightarrow t\bar{t}H$ [2]. Ce chapitre décrit le processus de production de quatre quarks top ainsi que quelques modèles de nouvelle physique pouvant conduire à l'augmentation de $\sigma_{t\bar{t}t\bar{t}}$.

Les meilleurs résultats actuels sur $\sigma_{t\bar{t}t\bar{t}}^{MS}$ au LHC à 13 TeV sont $12,6^{+5,8}_{-5,2} \text{ fb}$ pour l'expérience CMS [3], et $28,5^{+12}_{-11} \text{ fb}$ pour l'expérience ATLAS [4]. L'étude expérimentale de ce processus est l'objet des chapitres 5 et 6.

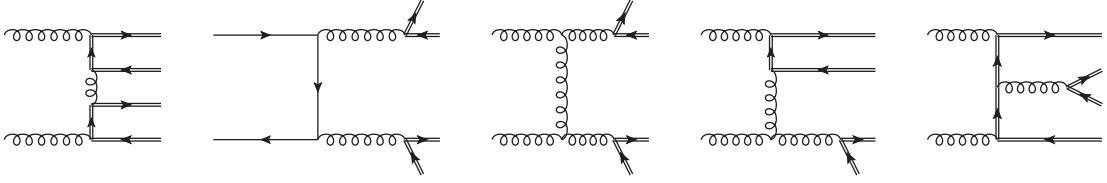


Figure 4.1 – Diagrammes de Feynman représentatifs pour l’amplitude de Born $pp \rightarrow t\bar{t}t\bar{t}$. Les diagrammes sont d’ordre $\mathcal{O}(\alpha_s^2)$. Les lignes doubles représentent les quarks top, les lignes simples correspondent aux quarks légers, et les boucles sont les gluons. [7]

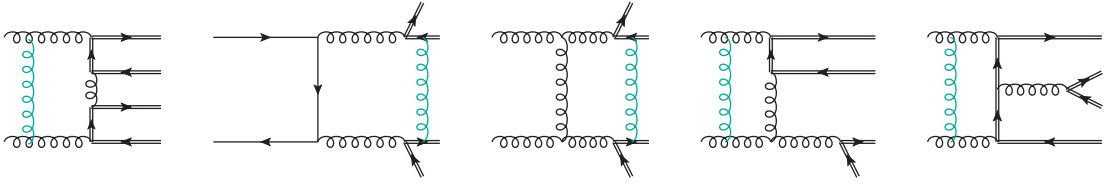


Figure 4.2 – Diagrammes de Feynman représentatifs pour l’amplitude $pp \rightarrow t\bar{t}t\bar{t}$ à une boucle. Les diagrammes sont d’ordre $\mathcal{O}(\alpha_s^3)$. Les lignes doubles représentent les quarks top, les lignes simples correspondent aux quarks légers, et les boucles sont les gluons. [7]

4.1 Production de quatre quarks top au LHC

4.1.1 Section efficace théorique

Au LHC, les quarks top sont le plus souvent produits par paire avec une section efficace $pp \rightarrow t\bar{t}$ d’environ 800 pb à 13 TeV [5], ou encore seuls avec une section efficace de l’ordre de 100 pb à 13 TeV [6]. La production de quatre quarks top $t\bar{t}t\bar{t}$ est beaucoup plus rare. Au LHC à 13 TeV, quatre quarks top peuvent être produits via une interaction entre deux gluons ou entre un quark et un anti-quark. Le premier cas est le plus probable, représentant environ 90 % de la section efficace totale [7].

Une série de diagrammes de Feynman représentatifs du processus $pp \rightarrow t\bar{t}t\bar{t}$ sont représentés sur les figures 4.1, 4.2, et 4.3. Les diagrammes de la figure 4.1 représentent l’amplitude de Born du processus, c’est-à-dire qu’il s’agit des diagrammes représentant le processus avec le nombre minimal de vertex, ici quatre. Les diagrammes de la figure 4.2 représentent l’amplitude à une boucle, et ceux de la figure 4.3 représentent l’amplitude avec une émission réelle de parton. Ces trois figures présentent des diagrammes qui contiennent exclusivement des vertex d’interaction forte puisqu’ils sont a priori dominants.

Toutefois les contributions électrofaibles au processus ne sont pas négligeables, et la figure 4.4 montre des diagrammes de Feynman incluant un ou plusieurs vertex électrofaibles. D’autres diagrammes doivent être considérés, par exemple ceux pour lesquels H est remplacé par γ ou Z .

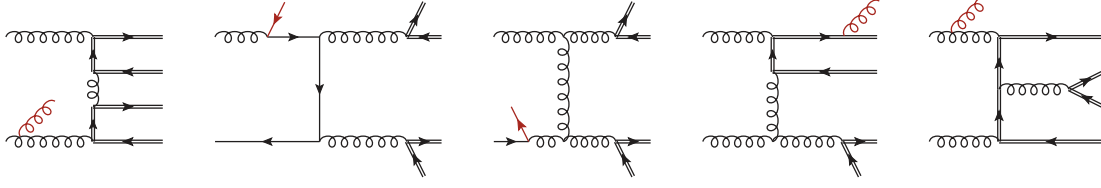


Figure 4.3 – Diagrammes de Feynman représentatifs pour l’amplitude $pp \rightarrow t\bar{t}t\bar{t}$ avec une émission réelle de parton (en rouge). Les diagrammes sont d’ordre $\mathcal{O}(\alpha_s^{5/2})$. Les lignes doubles représentent les quarks top, les lignes simples correspondent aux quarks légers, et les boucles sont les gluons. [7]

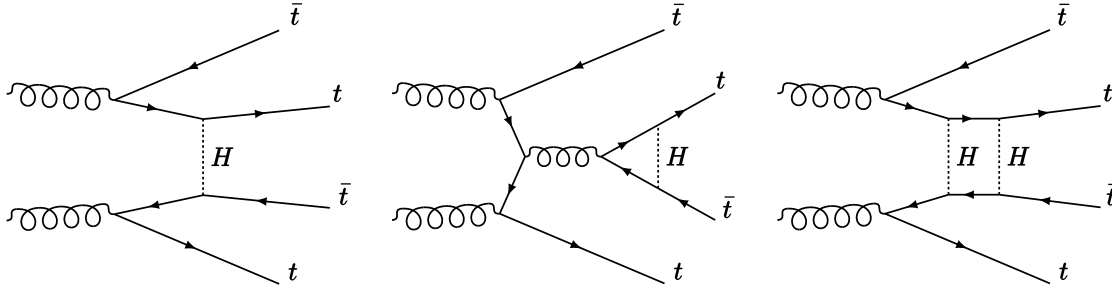


Figure 4.4 – Diagrammes de Feynman représentatifs pour l’amplitude $pp \rightarrow t\bar{t}t\bar{t}$ avec des vertex électrofaibles. Le diagramme de gauche est représentatif de l’amplitude de Born, tandis que les deux autres sont représentatifs de l’amplitude à une boucle [8]. Le diagramme de gauche est d’ordre $\mathcal{O}(\alpha_s\alpha)$, celui du milieu est d’ordre $\mathcal{O}(\alpha_s^2\alpha)$, et celui de droite est d’ordre $\mathcal{O}(\alpha_s\alpha^2)$.

Un calcul de section efficace fait intervenir la norme au carré d’une somme d’amplitudes, chaque amplitude pouvant être représentée par un diagramme de Feynman. L’ordre d’un diagramme est obtenu en multipliant $\alpha_s^{1/2}$ par vertex QCD et $\alpha^{1/2}$ par vertex électrofaible. Le nombre minimal de vertex dans un diagramme de Feynman $pp \rightarrow t\bar{t}t\bar{t}$ étant quatre, on en déduit que les diagrammes représentant l’amplitude de Born du processus sont d’ordre $\mathcal{O}(\alpha_s^i\alpha^j)$ avec $i + j = 2$. La section efficace calculée à l’ordre dominant (LO) est donc de l’ordre de la norme au carré d’un diagramme, soit $\mathcal{O}(\alpha_s^i\alpha^j)$ avec $i + j = 4$.

Un calcul complet doit prendre en compte tous les termes respectant $i + j = 4$. Puisque $\alpha/\alpha_s \sim 0,1$ à l’énergie du LHC, les diagrammes comportant le plus de vertex QCD dominant a priori le calcul. C’est pourquoi les termes LO_{QCD}, c’est-à-dire les termes pour lesquels $i = 4$ et $j = 0$ sont souvent les seuls pris en compte. Il s’agit des normes au carré des amplitudes représentées par les diagrammes similaires à ceux de la figure 4.1 (ordre $(\alpha_s^2)^2 = \alpha_s^4$), ainsi que les termes correspondant à leurs interférences (ordre $\alpha_s^2 \times \alpha_s^2 = \alpha_s^4$).

Un calcul au second ordre dominant (NLO) doit quant à lui prendre en compte tous les termes respectant $i + j = 5$. Souvent, seul le terme NLO_{QCD} est considéré puisqu'il est a priori dominant. Il s'agit du terme pour lequel $i = 5$ et $j = 0$. Il contient donc toutes les normes au carré des amplitudes des diagrammes avec une émission réelle de parton du type de ceux représentés sur la figure 4.3 (ordre $(\alpha_s^{5/2})^2 = \alpha_s^5$), ainsi que tous les termes d'interférence entre les diagrammes représentatifs de l'amplitude de Born comme ceux représentés sur la figure 4.1 et les diagrammes à une boucle similaires à ceux représentés sur la figure 4.2 (ordre $\alpha_s^2 \times \alpha_s^3 = \alpha_s^5$).

Un calcul NLO en QCD, c'est-à-dire $\text{LO}_{\text{QCD}} + \text{NLO}_{\text{QCD}}$, a été mené dans la référence [9]. La section efficace calculée est de $9,2_{-2,4}^{+2,9}$ fb à 13 TeV. Cette valeur est utilisée dans le chapitre 5. Un calcul plus récent fournit une valeur compatible mais un peu plus élevée : $11,1_{-2,6}^{+2,1}$ fb [8]. Les incertitudes sur les sections efficaces incluent les incertitudes dues à la dépendance du calcul aux échelles (renormalisation, factorisation), ainsi qu'à la dépendance aux fonctions de distribution de partons (PDF). Les incertitudes dues aux échelles dominent l'incertitude totale.

Les termes pour lesquels $j \neq 0$ ne sont cependant pas négligeables et ont été pris en compte dans la référence [8]. La section efficace obtenue ainsi est de $12,0_{-2,5}^{+2,2}$ fb à 13 TeV. Cette valeur est utilisée dans le chapitre 6. Ce calcul prend en compte des termes additionnels LO, soit tous les termes pour lesquels $i + j = 4$ et $j \neq 0$. Par exemple les termes pour lesquels $i = 2$ et $j = 2$ sont issus de la norme au carré de l'amplitude représentée par les diagrammes similaires au diagramme de gauche de la figure 4.4 (ordre $(\alpha_s \alpha)^2 = \alpha_s^2 \alpha^2$). Autre exemple, les termes pour lesquels $i = 3$ et $j = 1$ proviennent des interférences entre les diagrammes similaires au diagramme de gauche de la figure 4.4 et ceux de la figure 4.1 (ordre $\alpha_s \alpha \times \alpha_s^2 = \alpha_s^3 \alpha$). Le calcul prend également en compte des termes additionnels NLO pour lesquels $i + j = 5$ et $j \neq 0$. Par exemple, les interférences entre le diagramme central (respectivement de droite) de la figure 4.4 et les diagrammes de la figure 4.1 sont des termes d'ordre $\alpha_s^2 \alpha \times \alpha_s^2 = \alpha_s^4 \alpha$ (respectivement $\alpha_s \alpha^2 \times \alpha_s^2 = \alpha_s^3 \alpha^2$).

Le calcul NLO complet utilise une définition dynamique des échelles de factorisation μ_f et de renormalisation μ_r ayant pour valeur commune

$$\mu_c = \frac{H_T}{4}, \quad (4.1)$$

où H_T est la somme des masses transverses $m_T = \sqrt{m^2 + p_T^2}$ de toutes les particules dans l'événement. Le choix des échelles influence grandement le calcul, notamment l'importance relative des termes électrofaibles dans la section efficace totale. Comme le témoignent les incertitudes, la section efficace totale est augmentée (respectivement diminuée) d'environ 20 % suite à une multiplication (respectivement division) par 2 de μ_c [10].

4.1.2 Signatures expérimentales

Un quark top se désintègre quasi-exclusivement en quark bottom et en boson W : $t \rightarrow Wb$ [1]. Son temps de vie est d'environ 5×10^{-25} s, tandis que le temps caractéristique de l'interaction forte est de l'ordre de 10^{-23} s. Le quark top se désintègre donc trop vite pour s'hadroniser, et se déplace pendant un temps trop court pour que son déplacement soit observable dans un détecteur. Seuls ses produits de désintégration peuvent être observés. Les événements comportant quatre quarks top peuvent donc être considérés comme des événements avec quatre quarks b et quatre bosons W . Il est utile de remarquer à ce stade que l'état final du processus étudié est riche en quarks b , ce qui justifie un critère de sélection des événements fondé en partie sur le nombre de jets de quarks b dans l'événement.

Chaque boson W provenant d'un quark top se désintègre également très rapidement, avec un temps de vie d'environ 3×10^{-25} s. Ainsi seuls ses produits de désintégration peuvent être observés. Plusieurs canaux de désintégration existent, qu'il est possible de résumer ainsi :

$$W \rightarrow \begin{cases} \ell \nu_\ell & (\sim 30 \%) \\ qq' & (\sim 70 \%) \end{cases} \quad (4.2)$$

avec ℓ désignant un lepton chargé, ν_ℓ un neutrino (ou anti-neutrino) de même saveur que ce lepton, et qq' deux quarks légers par exemple $u\bar{d}$ ou $c\bar{s}$. Les leptons τ ne sont pas identifiés explicitement dans les travaux présentés aux chapitres 5 et 6, seuls leurs produits de désintégration le sont. Un τ est donc vu comme un électron, un muon, ou encore des quarks, selon son canal de désintégration. Par la suite on appelle « leptonique » une désintégration d'un boson W en électron ou en muon, et « hadronique » une désintégration en quarks.

Connaissant les rapports d'embranchement des désintégrations du boson W et du lepton τ [1], on peut en déduire les différents états finaux engendrés par la désintégration de quatre bosons W . Ces derniers sont catégorisés en fonction du nombre de désintégrations leptoniques et hadroniques, et sont représentés avec leurs rapports d'embranchement sur la figure 4.5. On obtient une probabilité d'environ 25 % (respectivement 75 %) pour une désintégration leptonique (respectivement hadronique) d'un boson W , en prenant en compte la désintégration des leptons τ .

Certaines analyses s'intéressent aux événements contenant un seul lepton et des jets, soit la catégorie « lhhh » sur la figure. Ce canal est intéressant de par son haut rapport d'embranchement d'environ 40 % du total des événements $pp \rightarrow t\bar{t}t\bar{t}$. De plus, ces analyses incluent souvent les événements du type « llhh OC », à savoir contenant deux leptons de charge opposée et des jets, qui peuvent être analysés avec des méthodes similaires. En considérant ces deux catégories à la fois, on atteint 56,5 % des événements $pp \rightarrow t\bar{t}t\bar{t}$, soit plus de la moitié. Cependant les deux canaux mentionnés souffrent d'un bruit de fond très important, provenant principalement du processus $pp \rightarrow t\bar{t}$ dont la section efficace

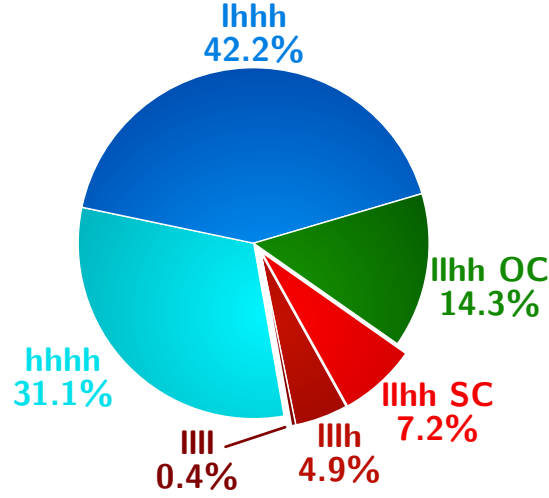


Figure 4.5 – États finaux et leurs rapports d'embranchement pour la désintégration de quatre bosons W . Le symboles « l » et « h » désignent respectivement une désintégration leptonique (e, μ , ou $\tau \rightarrow e, \mu$) et hadronique. Les abréviations « SC » et « OC » désignent respectivement des événements avec deux leptons de même charge électrique, ou de charge opposée.

est cinq ordres de grandeur supérieure à $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}}$. Par conséquent le rapport signal sur bruit de l'analyse est faible.

Les événements comportant au moins deux leptons de même charge électrique sont également étudiés : « llh SC », « llh », et « llhl ». C'est l'état final choisi dans ce travail de thèse et celui étudié par les analyses présentées aux chapitres 5 et 6. Cet état final ne représente que 12,5 % des événements $pp \rightarrow t\bar{t}t\bar{t}$, mais présente un bruit de fond très faible. En effet peu de processus du modèle standard produisent deux leptons de même charge, on s'attend donc à un bon rapport signal sur bruit pour ce canal.

4.2 Production de quatre quarks top dans les modèles de nouvelle physique

La section efficace du processus $pp \rightarrow t\bar{t}t\bar{t}$ peut être augmentée dans plusieurs modèles BSM, via de nouvelles particules ou encore une interaction effective à quatre quarks top. Ces modèles introduisent par exemple une ou plusieurs nouvelles particules pouvant se désintégrer en paire de quarks top $X \rightarrow t\bar{t}$. Une telle particule peut être produite par paire $XX \rightarrow t\bar{t}t\bar{t}$ ou associée à une paire de quarks top $t\bar{t}X \rightarrow t\bar{t}t\bar{t}$. Si cette particule peut être produite à l'énergie du LHC alors il peut être possible d'observer une résonance. Dans le cas contraire la particule est trop massive mais son existence peut engendrer des effets à plus basse énergie, par exemple à l'énergie du LHC. Il est donc possible de

sonder plusieurs classes de modèles résonants et non résonants via l'étude du processus de production de quatre quarks top.

Cette section décrit succinctement les modèles influençant la section efficace $\sigma_{t\bar{t}t\bar{t}}$ dans des proportions qui pourraient être observées au LHC à 13 TeV, et qui ont été étudiés expérimentalement, comme il sera discuté dans le chapitre 5. Il s'agit d'un modèle d'interaction de contact (CI), d'un modèle de dimensions supplémentaires (2UED), et d'un modèle à deux doublets de Higgs (2HDM), dont les diagrammes de Feynman sont représentés sur les figures 4.6, 4.7, et 4.8.

D'autres modèles ont un impact sur la section efficace de production de $\sigma_{t\bar{t}t\bar{t}}$ mais ne sont pas présentés ici, comme par exemple des modèles supersymétriques [11, 12], des modèles topo-philés [13], ou encore des modèles introduisant par exemple des sextuplets ou octuplets possédant une charge de couleur [14, 15, 16, 17, 18].

4.2.1 Modèle effectif d'interaction de contact entre quatre quarks top

De nombreuses recherches de résonances lourdes se désintégrant en paire de quarks top $t\bar{t}$ ont été menées, mais aucune déviation significative vis-à-vis du modèle standard n'a encore été observée [19, 20, 21]. L'étude du processus $pp \rightarrow t\bar{t}t\bar{t}$ peut ouvrir une nouvelle fenêtre vers des phénomènes n'affectant pas la production de paire de quarks top. De nombreuses analyses au LHC interprètent notamment leurs résultats dans le cadre de la théorie des champs effectifs, ou *Effective Field Theory* (EFT). L'idée d'un tel dispositif est de supposer que les processus de nouvelle physique se situent à des énergies trop élevées pour être directement accessibles au LHC, mais qu'ils ont des effets détectables à basse énergie (celle du LHC). Ceci est modélisé via l'ajout d'opérateurs effectifs au lagrangien du modèle standard. Les opérateurs de ce lagrangien sont d'ordre quatre dans le modèle standard, et l'approche EFT s'intéresse à des opérateurs d'ordres supérieurs. L'avantage de l'approche EFT et de n'être fondée sur aucun modèle de nouvelle physique particulier. Elle permet donc d'étudier une classe de modèles plutôt qu'un modèle précis.

La mesure de la section efficace $\sigma_{t\bar{t}t\bar{t}}$ peut contribuer à contraindre certains opérateurs de dimension 6. Elle est notamment très compétitive en ce qui concerne l'opérateur de contact à quatre fermions dont deux quarks top $qqtt$ [22]. Les limites actuelles sur $\sigma_{t\bar{t}t\bar{t}}$ permettent de contraindre l'opérateur $qqtt$ environ autant que les mesures effectuées sur le processus $pp \rightarrow t\bar{t}$ qui ont une précision de l'ordre du pourcent. En effet si on note C/Λ^2 le coefficient de l'opérateur $qqtt$, alors la section efficace du processus de production de quatre quarks top dépend de ce coefficient à la puissance quatre avec un terme de la forme $(CE^2/\Lambda^2)^4$, E étant l'énergie du processus $pp \rightarrow t\bar{t}t\bar{t}$. La grande énergie du processus de production de quatre quarks top, ainsi que la présence de ce terme à la puissance quatre augmente grandement sa sensibilité à l'opérateur $qqtt$.

Un modèle EFT est étudié dans l'analyse présentée dans le chapitre 5 [23]. Il s'agit d'un modèle d'interaction effective à quatre quarks top, appelé modèle d'interaction de

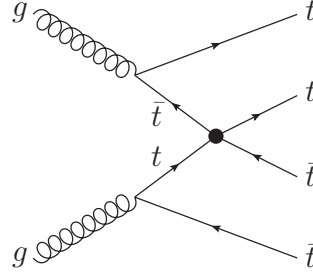


Figure 4.6 – Diagramme de Feynman de production de quatre quarks top via interaction de contact (CI).

contact, ou *Contact Interaction* (CI). Il suit l'approche décrite dans la référence [24]. Un diagramme de Feynman correspondant à cette interaction est montré sur la figure 4.6. L'interférence entre les processus MS et CI de production de quatre quarks top est négligeable et donc ignorée. Le lagrangien associé à cette interaction est :

$$\mathcal{L}_{4t} = \frac{C_{4t}}{\Lambda^2} (\bar{t}_R \gamma^\mu t_R) (\bar{t}_R \gamma_\mu t_R), \quad (4.3)$$

avec t_R le spineur droit du quark top, γ_μ les matrices de Dirac, C_{4t} une constante sans dimension, et Λ l'échelle d'énergie de nouvelle physique.

Le terme d'interaction de contact de l'équation 4.3 est un opérateur effectif de dimension 6, qui n'est pas renormalisable. Il peut être vu comme une approximation à basse énergie de plusieurs théories introduisant de nouvelles résonances lourdes et fortement couplées au quark top, comme les modèles composites [25, 26, 27, 28, 29], ou encore les modèles de dimensions supplémentaires [30, 31]. La nouvelle particule résonante échangée entre les quatre quarks top est trop massive pour être observée mais présente des effets à basse énergie, ce qui est modélisé par une interaction effective à quatre corps. L'établissement de limites sur la section efficace d'un tel processus permet alors indirectement de contraindre une classe de modèles comme expliqué auparavant. Ceci rappelle l'interaction de Fermi postulée en 1934 [32], à savoir une interaction effective à quatre corps représentant l'échange d'un boson W , trop massif pour être observé au moment de l'élaboration du modèle.

Les deux paramètres du modèle CI sont l'échelle de nouvelle physique qui peut être prise comme la masse de la résonance, ainsi que son couplage avec le quark top droit. L'analyse présentée au chapitre 5 utilise un modèle simplifié en fixant l'échelle de nouvelle physique à une valeur de 100 TeV afin d'être dans le régime d'interaction de contact. Il ne reste donc qu'un seul paramètre libre, à savoir le couplage C_{4t} qui contrôle directement la valeur de section efficace du processus de nouvelle physique. L'établissement d'une limite sur la section efficace du processus CI permet donc d'extraire une limite sur le paramètre C_{4t}/Λ^2 .

4.2.2 Modèle de dimensions supplémentaires

Les modèles introduisant deux dimensions en plus des quatre dimensions usuelles, appelés *Two Universal Extra Dimensions* (2UED), peuvent aussi augmenter la section efficace de production de quatre quarks top. Ces dimensions sont compactifiées dans la géométrie du plan projectif réel, comme décrit dans [33], ce qui introduit deux rayons R_5 et R_6 . Le plan projectif réel peut être construit à partir d'un rectangle, similairement au tore qui peut être obtenu en identifiant les côtés opposés du rectangle. Dans le cas du plan projectif réel, il s'agit également d'identifier les côtés opposés d'un rectangle mais en leur appliquant un demi-tour à la manière d'un ruban de Möbius. Cette construction n'est pas concrètement réalisable dans notre espace à trois dimensions spatiales. La compactification entraîne la discrétisation des impulsions le long de ces nouvelles dimensions, et chaque état d'impulsion peut être interprété comme une nouvelle particule, appelée excitation de Kaluza-Klein (KK). Les états sont indexés par deux entiers l et k représentant le niveau d'excitation dans chacune des nouvelles dimensions. Un couple d'indices (l, k) est appelé « niveau ». Les excitations KK $(0, 0)$ correspondent aux particules du modèle standard, on peut ainsi voir les nouvelles particules introduites comme des excitations des particules du modèle standard dans les deux nouvelles dimensions. La spécificité de ce modèle est que tous les champs du modèle standard peuvent se propager dans les nouvelles dimensions, d'où l'appellation *Universal*. La masse d'une excitation KK est

$$m^2 = m_0^2 + \frac{l^2}{R_5^2} + \frac{k^2}{R_6^2}, \quad (4.4)$$

avec m_0 la masse de la particule correspondante dans le modèle standard. Il existe des corrections au second ordre dominant différenciant les particules d'un même niveau (l, k) , mais elles sont faibles comparées aux masses des particules [31]. En introduisant les notations $m_{\text{KK}} = \frac{1}{R_5}$ et $\xi = \frac{R_5}{R_6}$, l'équation précédente s'écrit :

$$m^2 = m_0^2 + m_{\text{KK}}^2 (l^2 + k^2 \xi^2). \quad (4.5)$$

Selon ce modèle, il est possible de produire au LHC des paires de particules dans le niveau $(1, 1)$, qui se désintègrent en une cascade d'autres particules du niveau $(1, 1)$ étant donné la faible différence entre les masses des particules d'un même niveau. La particule en fin de cascade est la particule la plus légère du niveau $(1, 1)$, à savoir un photon lourd $A_\mu^{(1,1)}$. On suppose que ce dernier se désintègre exclusivement en paire de quarks top $t\bar{t}$, parce que le quark top est la particule la plus massive dans le modèle standard et pourrait donc se coupler le plus fortement avec les excitations KK. Il est donc possible d'obtenir quatre quarks top ainsi que d'autres quarks et leptons de basse énergie dans l'état final, comme illustré avec le diagramme de Feynman représenté sur la figure 4.7.

Certaines observations cosmologiques contraignent la valeur du paramètre m_{KK} , qui doit être au maximum de l'ordre de 1 TeV [31, 34]. L'analyse présentée dans le chapitre 5

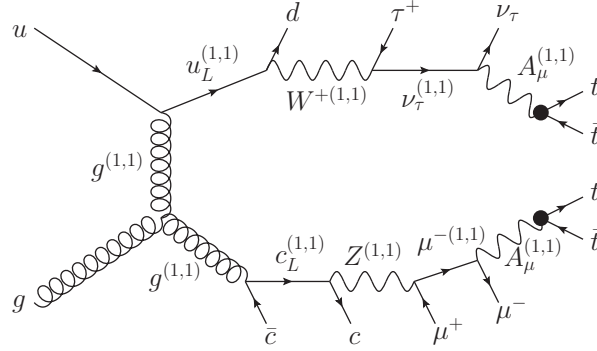


Figure 4.7 – Diagramme de Feynman de production de quatre quarks top dans le modèle à deux dimensions supplémentaires compactifiées (2UED).

sonde le modèle 2UED pour différentes valeurs du paramètre m_{KK} avec $\xi = 1$. Ces valeurs s'étendent de $m_{KK} = 600 \text{ GeV}$ à $m_{KK} = 2 \text{ TeV}$, pour une section efficace variant entre environ 300 fb et environ 1 fb.

4.2.3 Modèle avec deux doublets de Higgs

Un autre modèle de nouvelle physique étudié propose une extension du secteur du boson de Higgs. Le boson scalaire massif découvert autour de 125 GeV est décrit par un doublet de Higgs, comme expliqué dans la section 1.4, mais n'exclut pas l'existence d'autres doublets. On considère ici un modèle avec deux doublets de Higgs Φ_1 et Φ_2 qui brisent spontanément la symétrie électrofaible $SU(2)_L \otimes U(1)_Y$ [35]. On appelle ce modèle *Two-Higgs Doublet Model* (2HDM). Outre son intérêt dans le cadre de l'étude présentée ici, ce modèle est motivé par des modèles supersymétriques, des modèles d'axions, ou encore par le fait qu'il introduit de nouvelles sources de violation de CP et qu'il pourrait ainsi générer une asymétrie matière/antimatière d'amplitude suffisante pour se conformer aux observations.

Le lagrangien du modèle 2HDM est similaire à celui du secteur de Higgs du modèle standard :

$$\mathcal{L}_{2HDM} = y_1^{ij} \bar{\psi}_i \Phi_1 \psi_j + y_2^{ij} \bar{\psi}_i \Phi_2 \psi_j \quad (4.6)$$

où i et j sont les indices des générations de fermions. Parmi les différents modèles 2HDM, le modèle dit « de Type-II » est considéré. Il s'agit du modèle 2HDM le plus étudié car utilisé dans les modèles supersymétriques. Il introduit une différence entre les quarks de charge $+2/3$ d'un côté et les quarks de charge $-1/3$ ainsi que les leptons chargés de l'autre. En effet le champ Φ_1 interagit avec les quarks d , s , b , ainsi que les leptons chargés, tandis que Φ_2 interagit avec les quarks u , c , t . Deux valeurs moyennes dans le vide v_1 et v_2 sont présentes, et on a $v = \sqrt{v_1^2 + v_2^2} = 246 \text{ GeV}$, qui est la valeur moyenne dans le

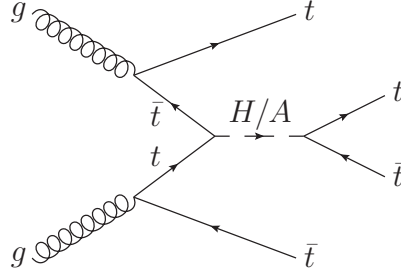


Figure 4.8 – Diagramme de Feynman de production de quatre quarks top dans le modèle à deux doublets de Higgs (2HDM).

vide du modèle standard. Ce modèle fait apparaître plusieurs bosons de Higgs : le boson scalaire usuel h , conforme à celui découvert au LHC, un boson scalaire neutre lourd H , un boson pseudo-scalaire A , ainsi que deux bosons chargés $H^\pm$. Les six paramètres libres du modèle sont les quatre masses des bosons de Higgs, ainsi que $\tan \beta = v_2/v_1$ et l'angle de mélange α entre les deux particules scalaires du modèle à l'origine des bosons h et H .

On s'intéresse à la production de bosons H ou A associée à une paire de quarks top, qui peut aboutir à l'état final $t\bar{t}t\bar{t}$ si ces bosons se désintègrent préférentiellement en paire $t\bar{t}$. Cette hypothèse est vérifiée dès lors que la masse des bosons est suffisante, et que la valeur de $\tan \beta$ est faible : $1 < \tan \beta \lesssim 5$. Pour des valeurs plus élevées de $\tan \beta$, la désintégration en paire $b\bar{b}$ devient dominante. L'étude porte donc sur la réaction suivante, illustrée sur la figure 4.8 :

$$gg \rightarrow t\bar{t}H/A \rightarrow t\bar{t}t\bar{t}. \quad (4.7)$$

La mesure des propriétés du boson de Higgs découvert au LHC contraint les modèles 2HDM, qui doivent être dans la « limite d'alignement » [36], à savoir $\alpha = \beta$. Ainsi, le boson h est bien similaire au boson de Higgs du modèle standard, et les autres bosons sont nettement plus lourds. Les bosons H et A ont également la même masse $m_{H/A}$ dans la limite d'alignement. La section efficace prédite par ce modèle pour le processus de production de quatre quarks top ne dépend donc que de $m_{H/A}$ ainsi que du paramètre $\tan \beta$. Pour des valeurs de $\tan \beta$ entre 1 et 5 et des valeurs de $m_{H/A}$ entre $2m_t$ et environ 1000 GeV, la section efficace de production de quatre quarks top prédite varie entre environ 10 fb et environ 0,1 fb.

L'analyse présentée dans le chapitre 5 étudie le modèle 2HDM dans le cas de la limite d'alignement $\alpha = \beta$, pour des valeurs de $m_{H/A}$ entre 500 GeV et 1000 GeV. Deux scénarios sont considérés, dans lesquels soit le boson lourd H seul contribue au processus de production de quatre quarks top, soit les deux bosons H et A contribuent. L'interférence du processus 2HDM avec le processus du modèle standard est négligeable.

Références

- [1] Particle Data GROUP. “Review of Particle Physics”. In : *Physical Review D* 98.3 (août 2018). DOI : 10.1103/physrevd.98.030001.
- [2] Qing-Hong CAO, Shao-Long CHEN et Yandong LIU. “Probing Higgs width and top quark Yukawa coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ productions”. In : *Physical Review D* 95.5 (mar. 2017). DOI : 10.1103/physrevd.95.053004.
- [3] CMS COLLABORATION. *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Rapp. tech. CMS-PAS-TOP-18-003. Geneva : CERN, 2019. URL : <https://cds.cern.ch/record/2668710>.
- [4] ATLAS COLLABORATION. “Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Physical Review D* 99.5 (mar. 2019). DOI : 10.1103/physrevd.99.052009.
- [5] Michał CZAKON, Paul FIEDLER et Alexander MITOV. “Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_s^4)$ ”. In : *Physical Review Letters* 110.25 (juin 2013). DOI : 10.1103/physrevlett.110.252004.
- [6] Mathias BRUCHERSEIFER, Fabrizio CAOLA et Kirill MELNIKOV. “On the NNLO QCD corrections to single-top production at the LHC”. In : *Physics Letters B* 736 (sept. 2014), p. 58-63. DOI : 10.1016/j.physletb.2014.06.075.
- [7] G. BEVILACQUA et M. WOREK. “Constraining BSM physics at the LHC : four top final states with NLO accuracy in perturbative QCD”. In : *Journal of High Energy Physics* 2012.7 (juil. 2012). DOI : 10.1007/jhep07(2012)111.
- [8] Rikkert FREDERIX, Davide PAGANI et Marco ZARO. “Large NLO corrections in $t\bar{t}W$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions”. In : *Journal of High Energy Physics* 2018.2 (fév. 2018). DOI : 10.1007/jhep02(2018)031.
- [9] J. ALWALL et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In : *Journal of High Energy Physics* 2014.7 (juil. 2014). DOI : 10.1007/jhep07(2014)079.
- [10] Fabio MALTONI, Davide PAGANI et Ioannis TSINIKOS. “Associated production of a top-quark pair with vector bosons at NLO in QCD : impact on $t\bar{t}H$ searches at the LHC”. In : *Journal of High Energy Physics* 2016.2 (fév. 2016). DOI : 10.1007/jhep02(2016)113.

- [11] Bobby S. ACHARYA et al. “Identifying Multi-Top Events from Gluino Decay at the LHC”. In : (22 jan. 2009). arXiv : 0901.3367v1 [hep-ph].
- [12] Gordon L. KANE et al. “Top channel for early supersymmetry discovery at the LHC”. In : *Physical Review D* 84.9 (nov. 2011). DOI : 10.1103/physrevd.84.095004.
- [13] Ezequiel ALVAREZ et al. “Four tops for LHC”. In : *Nuclear Physics B* 915 (fév. 2017), p. 19-43. DOI : 10.1016/j.nuclphysb.2016.11.024.
- [14] Michael GERBUSH et al. “Color-octet scalars at the CERN LHC”. In : *Physical Review D* 77.9 (mai 2008). DOI : 10.1103/physrevd.77.095003.
- [15] Chuan-Ren CHEN et al. “Color sextet scalars at the CERN Large Hadron Collider”. In : *Physical Review D* 79.5 (mar. 2009). DOI : 10.1103/physrevd.79.054002.
- [16] J. A. AGUILAR-SAAVEDRA et J. SANTIAGO. “Four top quarks and the $t\bar{t}$ forward-backward asymmetry”. In : *Physical Review D* 85.3 (fév. 2012). DOI : 10.1103/physrevd.85.034021.
- [17] Thomas GRÉGOIRE, Emanuel KATZ et Veronica SANZ. “Four top quarks in extensions of the standard model”. In : *Physical Review D* 85.5 (mar. 2012). DOI : 10.1103/physrevd.85.055024.
- [18] Luc DARMÉ, Benjamin FUKS et Mark GOODSSELL. “Cornering sgluons with four-top-quark events”. In : *Physics Letters B* 784 (sept. 2018), p. 223-228. DOI : 10.1016/j.physletb.2018.08.001.
- [19] ATLAS COLLABORATION. *Search for heavy particles decaying into a top-quark pair in the fully hadronic final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Rapp. tech. arXiv :1902.10077. * Temporary entry *. Geneva : CERN, fév. 2019. URL : <https://cds.cern.ch/record/2664828>.
- [20] ATLAS COLLABORATION. “Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *The European Physical Journal C* 78.7 (juil. 2018). DOI : 10.1140/epjc/s10052-018-5995-6.
- [21] CMS COLLABORATION. *Search for resonant $t\bar{t}$ production in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Rapp. tech. arXiv :1810.05905. CMS-B2G-17-017-003. * Temporary entry *. Geneva : CERN, oct. 2018. URL : <https://cds.cern.ch/record/2643605>.
- [22] Cen ZHANG. “Constraining $q\bar{q}t\bar{t}$ operators from four-top production : a case for enhanced EFT sensitivity”. In : *Chinese Physics C* 42.2 (fév. 2018), p. 023104. DOI : 10.1088/1674-1137/42/2/023104.

- [23] ATLAS COLLABORATION. “Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Journal of High Energy Physics* 2018.12 (déc. 2018). DOI : 10.1007/jhep12(2018)039.
- [24] Céline DEGRANDE et al. “Non-resonant new physics in top pair production at hadron colliders”. In : *Journal of High Energy Physics* 2011.3 (mar. 2011). DOI : 10.1007/jhep03(2011)125.
- [25] Howard GEORGI et al. “Effects of top quark compositeness”. In : *Physical Review D* 51.7 (avr. 1995), p. 3888-3894. DOI : 10.1103/physrevd.51.3888.
- [26] Ben LILLIE, Jing SHU et Tim M.P TAIT. “Top compositeness at the Tevatron and LHC”. In : *Journal of High Energy Physics* 2008.04 (avr. 2008), p. 087-087. DOI : 10.1088/1126-6708/2008/04/087.
- [27] Alex POMAROL et Javi SERRA. “Top quark compositeness : Feasibility and implications”. In : *Physical Review D* 78.7 (oct. 2008). DOI : 10.1103/physrevd.78.074026.
- [28] Kunal KUMAR, Tim M.P TAIT et Roberto VEGA-MORALES. “Manifestations of top compositeness at colliders”. In : *Journal of High Energy Physics* 2009.05 (mai 2009), p. 022-022. DOI : 10.1088/1126-6708/2009/05/022.
- [29] Da LIU et Rakhi MAHBUBANI. “Probing top-antitop resonances with $t\bar{t}$ scattering at LHC14”. In : *Journal of High Energy Physics* 2016.4 (avr. 2016), p. 1-26. DOI : 10.1007/jhep04(2016)116.
- [30] M. GUCHAIT, F. MAHMOUDI et K. SRIDHAR. “Associated production of a Kaluza-Klein excitation of a gluon with a $t\bar{t}$ pair at the LHC”. In : *Physics Letters B* 666.4 (sept. 2008), p. 347-351. DOI : 10.1016/j.physletb.2008.07.085.
- [31] Giacomo CACCIAPAGLIA et al. “Four tops on the real projective plane at LHC”. In : *Journal of High Energy Physics* 2011.10 (oct. 2011). DOI : 10.1007/jhep10(2011)042.
- [32] Fred L. WILSON. “Fermi’s Theory of Beta Decay”. In : *American Journal of Physics* 36.12 (déc. 1968), p. 1150-1160. DOI : 10.1119/1.1974382.
- [33] Giacomo CACCIAPAGLIA, Aldo DEANDREA et Jérémie LLODRA-PEREZ. “A dark matter candidate from Lorentz invariance in 6D”. In : *Journal of High Energy Physics* 2010.3 (mar. 2010). DOI : 10.1007/jhep03(2010)083.
- [34] Alexandre ARBEY et al. “Dark Matter in a twisted bottle”. In : *Journal of High Energy Physics* 2013.1 (jan. 2013). DOI : 10.1007/jhep01(2013)147.

- [35] G.C. BRANCO et al. “Theory and phenomenology of two-Higgs-doublet models”. In : *Physics Reports* 516.1-2 (juil. 2012), p. 1-102. DOI : 10.1016/j.physrep.2012.02.002.
- [36] P. S. Bhupal DEV et Apostolos PILAFTSIS. “Maximally symmetric two Higgs doublet model with natural standard model alignment”. In : *Journal of High Energy Physics* 2014.12 (déc. 2014). DOI : 10.1007/jhep12(2014)024.

5

Recherche de production de quatre quarks tops avec le détecteur ATLAS au LHC avec $36,1 \text{ fb}^{-1}$

5.1 Introduction

La production de quatre quarks top au LHC $pp \rightarrow t\bar{t}t\bar{t}$ est un processus physique extrêmement rare. Son étude permet d'une part de mettre à l'épreuve une prédiction du modèle standard dans un environnement complexe et très énergétique, et d'autre part de sonder différents modèles BSM prédisant une augmentation de sa section efficace. La discussion théorique de ces aspects est l'objet du chapitre 4.

Ce travail de thèse décrit dans ce chapitre s'inscrit dans l'analyse de ce processus [1] portant sur les données du Run 2 prises en 2015 et 2016 par le détecteur ATLAS, correspondant à une luminosité intégrée d'environ $36,1 \text{ fb}^{-1}$. Seuls les événements comportant au moins deux leptons de même charge électrique ainsi que des jets de quarks b sont considérés ici. Il est important de rappeler que les leptons τ ne sont pas reconstruits dans l'analyse présentée ici, seuls leurs produits de désintégration le sont. C'est pourquoi dans ce chapitre le terme « lepton » ou encore le symbole ℓ désigne un électron e ou un muon μ , mais pas un lepton τ .

Cette analyse s'intéresse à plusieurs processus physiques, décrits dans le modèle standard ou bien via des modèles BSM :

- production de quatre quarks top dans le cadre du modèle standard ;
- production de quatre quarks top via un modèle d'interaction de contact (CI) ;
- production de quatre quarks top via un modèle de dimensions supplémentaires (2UED) ;
- production de quatre quarks top via un modèle de doublet de Higgs de Type-II (2HDM) ;
- production de *Vector-Like Quarks* (VLQ) ;
- production de deux quarks top de même charge via un modèle de matière noire.

Mon travail personnel est principalement lié aux modèles de production de quatre quarks top, en particulier la production conforme au modèle standard et dans une moindre mesure les modèles CI, 2UED, et 2HDM. On ne va donc s'intéresser qu'à ces modèles dans le cadre de cette thèse.

Une analyse portant sur les données du Run 1 prises par le détecteur ATLAS à

8 TeV, correspondant à une luminosité intégrée d'environ 20,3 fb⁻¹, place une limite supérieure observée (attendue) de 70 fb (27 fb) sur la section efficace de production de quatre quarks top prédite par le modèle standard $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}}$ [2]. Cette même analyse a été adaptée aux premières données du Run 2 prises en 2015, correspondant à une luminosité intégrée d'environ 3,2 fb⁻¹, et place une limite supérieure observée (attendue) de 95 fb (107 fb) sur $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}}$ [3]. Deux analyses des données prises en 2016 par le détecteur CMS, correspondant à une luminosité intégrée d'environ 35,9 fb⁻¹, ont également étudié ce processus. Une analyse BSM [4] place une limite supérieure observée (attendue) de 42 fb (27 fb) sur $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}}$. L'autre analyse [5] étudie spécifiquement le processus $pp \rightarrow t\bar{t}t\bar{t}$ dans le cadre du MS et place une limite supérieure observée (attendue) de 42 fb (21 fb) sur $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}}$. Cette dernière analyse fournit également une mesure de $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}} = 16,9^{+13,8}_{-11,4}$ fb, avec une signification statistique de 1,6 déviations standard. Finalement deux analyses préliminaires [6, 7] mesurent respectivement $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}} = 13^{+11}_{-9}$ fb et $\sigma_{t\bar{t}t\bar{t}}^{\text{MS}} = 12,6^{+5,6}_{-5,2}$ fb.

On s'intéresse dans ce chapitre à l'analyse publiée dans *Journal of High Energy Physics* [1]. Cette analyse est combinée avec l'analyse utilisant les canaux à un seul lepton et à deux leptons de même charge électrique [8].

Dans le cadre de cette analyse mes contributions personnelles portent sur l'optimisation des régions de signal de l'analyse présentée dans la section 5.3.2, sur le calcul des incertitudes systématiques portant sur les bruits de fond irréductibles détaillé dans la section 5.5.3, sur la mise en place du dispositif d'analyse statistique présenté dans la section 5.6 ainsi que sur la production des résultats décrits dans la section 5.7 et les études de recouvrement pour les résultats combinés.

5.2 Sélection des objets utilisés dans l'analyse

L'analyse de données présentée dans ce travail [1] utilise différents objets tels que les électrons, les muons, les jets, ou encore l'énergie transverse manquante, reconstruits à partir des informations des différents sous-détecteurs d'ATLAS. Les méthodes de reconstruction des objets sont présentées dans le chapitre 2. La sélection appliquée aux objets est décrite dans cette section et résumée dans le tableau 5.1.

Les électrons utilisés doivent avoir une impulsion transverse $p_T > 28$ GeV. Ils doivent se trouver dans la région de l'espace définie par $|\eta| < 2,47$, en excluant la région de transition entre tonneau et bouchons $1,37 < |\eta| < 1,52$. Dans le cas d'événements avec deux électrons ou encore un électron et un muon, la coupure est plus stricte : $|\eta| < 1,37$. Cette coupure permet de rejeter une large proportion du bruit de fond dû à la mauvaise identification de la charge des électrons décrit dans la section 5.4.2. Les muons quant à eux doivent respecter $p_T > 28$ GeV et $|\eta| < 2,5$.

Les électrons et les muons doivent provenir du vertex primaire. Pour s'en assurer, des conditions sont définies à partir de leur paramètre d'impact d_0 dans le plan $x - y$

	Électrons		Muons		Jets	Jets de quarks b
	lâche	strict	lâche	strict		
p_T [GeV]	> 28		> 28		> 25	> 25
$ \eta $	< 1.37 ou $1.52 - 2.47$ (< 1.37 pour ee et $e\mu$)		< 2.5		< 2.5	< 2.5
Qualité de l'identification	mediumLH	tightLH	medium		cleaning + JVT	MV2c10 77 %
Isolation	aucune	trace et calorimètre	aucune	trace		
Vertex de la trace :						
– $ d_0/\sigma_{d_0} $	< 5		< 3			
– $ z_0 \sin \theta $ [mm]	< 0.5		< 0.5			

Table 5.1 – Résumé de la sélection des objets utilisés dans l'analyse.

et de son incertitude σ_{d_0} . Les électrons doivent satisfaire $|d_0/\sigma_{d_0}| < 5$ et les muons $|d_0/\sigma_{d_0}| < 3$. Une coupure est également appliquée sur $|z_0 \sin \theta|$, où z_0 est la différence entre la coordonnée z du vertex primaire et celle du point de la trace du lepton considéré le plus proche de l'axe du faisceau. On requiert $|z_0 \sin \theta| < 0,5$ mm pour les électrons et les muons.

Les leptons doivent satisfaire un critère d'identification strict ou lâche. L'échantillon strict est plus restrictif et est inclus dans l'échantillon lâche. Seuls les événements comportant des leptons stricts sont analysés, sauf pour l'estimation du bruit de fond dû aux faux leptons et aux leptons non-prompts décrit dans la section 5.4.2, pour lequel tous les leptons lâches sont pris en compte. La qualité de l'identification requise pour les électrons stricts (lâches) est le critère « tight » (« medium ») défini en [9]. Ces critères correspondent à des coupures différentes sur une valeur de vraisemblance calculée en combinant plusieurs paramètres de la trace reconstruite par le trajectographe avec des paramètres de gerbe mesurés par les différentes couches du calorimètre électromagnétique. La qualité d'identification requise pour les muons stricts et lâche est la même, il s'agit du critère « medium » défini en [10]. Ce dernier est calculé à partir de paramètres de traces du muon mesurés par le trajectographe ainsi que par le spectromètre à muons, incluant notamment un ajustement global de la trace dans les deux sous-détecteurs.

Les leptons stricts doivent de plus satisfaire un critère d'isolation calculé à partir de la somme des p_T des traces situées dans un cône de taille variable autour du lepton, excluant la trace du lepton. Cette somme doit être inférieure à 6 % du p_T du lepton. La taille de cône utilisée est $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = \min(10 \text{ GeV}/p_T, 0,2(0,3))$ pour les électrons (muons). Les électrons doivent de plus respecter un critère calculé à partir de

la somme des p_T des amas d'énergie dans le calorimètre située dans un cône de taille $\Delta R = 0,2$ autour de l'électron, excluant l'énergie de l'électron. Cette somme doit être inférieure à 6 % du p_T de l'électron.

Les jets sont reconstruits à partir d'amas d'énergie dans le calorimètre en utilisant l'algorithme anti- k_t [11] avec un paramètre de rayon de 0,4. Les jets doivent vérifier $p_T > 25 \text{ GeV}$ et $|\eta| < 2,5$. Des critères de qualité sont appliqués dans le but de réduire la contamination de faux jets comme indiqué dans le tableau 5.1 sous le terme « cleaning » [12]. Afin de rejeter les jets issus du *pile-up*, une observable « JVT » est définie à partir de la fraction de l'impulsion totale du jet portée par ses traces associées au vertex primaire [13]. Les jets vérifiant $p_T < 60 \text{ GeV}$ et $|\eta| < 2,4$ subissent une coupure sur cette observable.

Les jets issus d'un hadron b sont identifiés grâce à une technique multivariée [14]. Le critère MV2c10 est appliqué avec une efficacité de sélection de jets de quarks b de 77 % dans des événements $t\bar{t}$ simulés. Les facteurs de réjection des jets de quarks c et des jets légers sont alors de 6 et 134 respectivement.

L'énergie transverse manquante est calculée à partir des impulsions transverses des objets reconstruits dans l'événement : électrons, photons, leptons τ se désintégrant en hadrons, jets, muons, ainsi que la somme p_T des traces compatibles avec le vertex primaire mais n'étant pas associées à l'un de ces objets, appelé *soft term* [15]. Il s'agit de l'opposé de la somme de toutes ces impulsions transverses. On note sa norme E_T^{miss} .

Enfin, des critères de chevauchement sont appliqués aux leptons lâches. D'abord, les électrons partageant une trace avec un muon sont éliminés. Ensuite le jet le plus proche d'un électron dans un intervalle $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta \phi)^2} = 0,2$ est supprimé pour éviter un double comptage. Puis les électrons dans un intervalle $\Delta R_y = 0,4$ d'un jet sont enlevés. Les jets possédant moins de trois traces et se trouvant dans un intervalle $\Delta R_y = 0,4$ d'un muon sont ensuite retirés. Enfin, les muons dans un intervalle $\Delta R_y = 0,04 + 10 \text{ GeV}/p_{T,\mu}$ d'un jet sont supprimés.

5.3 Sélection des événements

Les données utilisées par cette analyse ont été enregistrées par le détecteur ATLAS en 2015 et 2016. Elles correspondent à des collisions proton-proton pp du Run 2 du LHC à une énergie dans le centre de masse de 13 TeV. La luminosité intégrée de l'ensemble des données considérées représente $36,1 \pm 0,8 \text{ fb}^{-1}$. Le nombre moyen d'interactions par croisement de faisceaux est appelé empilement, ou *pile-up*. Il vaut environ 24 dans ce jeu de données.

Dans cette analyse, la section efficace théorique du processus $pp \rightarrow t\bar{t}t\bar{t}$ prédite par le modèle standard est calculée au second ordre dominant (NLO) en QCD, comme décrit

dans la section 4.1.1. Elle vaut environ 9,2 fb [16]. Par conséquent environ 300 événements $pp \rightarrow t\bar{t}t\bar{t}$ sont attendus pour une luminosité intégrée de $36,1 \text{ fb}^{-1}$, ce qui est très faible.

Plusieurs critères de sélection des événements sont définis afin de maximiser la sensibilité de l'analyse au processus $t\bar{t}t\bar{t}$ ainsi qu'aux modèles BSM étudiés. Une pré-sélection est appliquée à tous les événements, puis plusieurs jeux de coupures définissent différentes sélections d'événements appelées régions de signal. Chacune de ces régions présente un rapport signal sur bruit différent et peut donc être utile pour contraindre la modélisation d'un certain type de bruit de fond ou encore peut être sensible à un signal en particulier. Huit régions de signal sont ainsi définies et sont combinées dans l'analyse statistique finale, selon la méthode décrite dans la section 5.6.

5.3.1 Pré-sélection

Les périodes de prises de données considérées sont celles pour lesquelles tous les sous-détecteurs ont fonctionné de manière nominale, et pour lesquelles le LHC était en mode de collision stable. Les événements utilisés par l'analyse doivent comporter un vertex primaire associé à au moins deux particules chargées vérifiant $p_T > 0,4 \text{ GeV}$. Deux catégories de déclenchements sont considérées :

- des déclenchements par deux leptons, intéressants car sensibles à des leptons à bas p_T ;
- ainsi que des déclenchements par un lepton à haut p_T .

Les événements doivent avoir déclenché au moins l'un de ces deux critères. Ces derniers comportent des critères d'identification et d'isolation lâches, puisque des critères stricts compliqueraient l'estimation du bruit de fond dû aux faux leptons décrit dans la section 5.4.2. La valeur de p_T minimale pour les déclenchements par deux leptons varie selon le type de lepton et l'année de prise de données entre 7 et 24 GeV. Dans le cas du déclenchement par un muon seul il s'agit de 50 GeV, et dans le cas d'un électron seul 24 GeV pour les données prises en 2015 et 60 GeV pour les données prises en 2016. L'efficacité de sélection associée au déclenchement est supérieure à 95 % pour toutes les combinaisons de leptons possibles dans les événements étudiés.

Les événements doivent contenir au moins un jet et au moins deux leptons répondant au critère de sélection strict. Si deux leptons parmi les trois leptons ayant le p_T le plus élevé répondent au critère strict mais pas le troisième, alors ils doivent avoir la même charge électrique. De plus, s'il s'agit de deux électrons, ils doivent passer une coupure supplémentaire de rejet de quarkonia/ Z sur leur masse invariante : $m_{ee} > 15 \text{ GeV}$ et $|m_{ee} - 91 \text{ GeV}| > 10 \text{ GeV}$. Une telle coupure n'est pas nécessaire dans le cas d'une paire de muons puisque la probabilité qu'un boson Z produise deux muons de même charge électrique, donc avec une mauvaise identification de la charge d'un muon, est négligeable. Les événements satisfaisant ces critères sont classés en tant qu'événements avec deux leptons de même charge : ee , $e\mu$, ou $\mu\mu$. Si les trois leptons ayant le p_T le plus élevé

répondent au critère strict, l'événement est classé en tant qu'événement avec trois leptons : eee , $ee\mu$, $e\mu\mu$, ou $\mu\mu\mu$. Les deux catégories d'événements sont traitées séparément dans cette analyse.

5.3.2 Méthode d'optimisation des régions de signal

L'analyse présentée dans ce chapitre cible plusieurs modèles de physique BSM ainsi que le processus de production de quatre quarks tops prédit par le modèle standard. Ces processus physiques ont pour point commun de pouvoir produire deux leptons de même charge électrique, c'est pourquoi ils sont recherchés dans cette analyse. Malgré ce point commun les modèles sont différents donc une sélection des événements optimale pour l'un ne l'est pas forcément pour l'autre. Cependant la définition de régions de signal optimales pour chacun des modèles étudiés résulterait en des centaines de régions. Le choix du groupe d'analyse a été de définir des régions communes à tous les modèles afin de simplifier la phase d'analyse statistique.

Les variables offrant la meilleure séparation entre signal et bruit de fond sont le nombre de jets de quarks b N_b , le nombre de jets N_j , l'énergie totale dans l'événement H_T , et dans une moindre mesure l'énergie transverse manquante E_T^{miss} . La grandeur H_T est calculée comme la somme des normes des impulsions transverses de tous les leptons et les jets reconstruits dans l'événement. La figure 5.1 représente les distributions de ces quatre grandeurs pour le bruit de fond et plusieurs signaux représentatifs. Il est clair que les signaux recherchés par l'analyse présentent des valeurs de ces grandeurs nettement plus élevées que le bruit de fond. Finalement, les événements avec deux leptons de même charge d'une part et les événements avec trois leptons d'autre part sont sélectionnés dans des régions de signal distinctes puisque les bruits de fond associés à ces deux catégories sont différents.

Six modèles représentatifs ont été choisis pour l'optimisation des régions de signal, dont la production de quatre quarks top dans le cadre du modèle standard, ainsi que le modèle 2UED pour une masse de 1400 GeV (voir section 4.2.2). Les autres sont des modèles de VLQ étudiés par l'analyse mais non décrits dans ce manuscrit, deux modèles à basse masse (800 GeV) et deux autres à haute masse (1200 GeV). La variable d'optimisation est la moyenne des significations statistiques attendues de chacun des six modèles.

Au moment de l'optimisation, l'estimation des bruits de fond était encore préliminaire et entièrement basée sur des simulations Monte-Carlo. Les bruits de fond réductibles notamment étaient estimés à partir d'échantillons Monte-Carlo, contrairement aux méthodes finales d'estimation présentées dans la section 5.4.2. Les échantillons choisis pour modéliser ces bruits de fond irréductibles sont ceux des processus supposés majoritairement responsables de ce type de bruit de fond : production d'une paire de quarks top $t\bar{t}$, d'un seul quark top associé à d'autres particules, ou d'un boson vecteur associé à des jets.

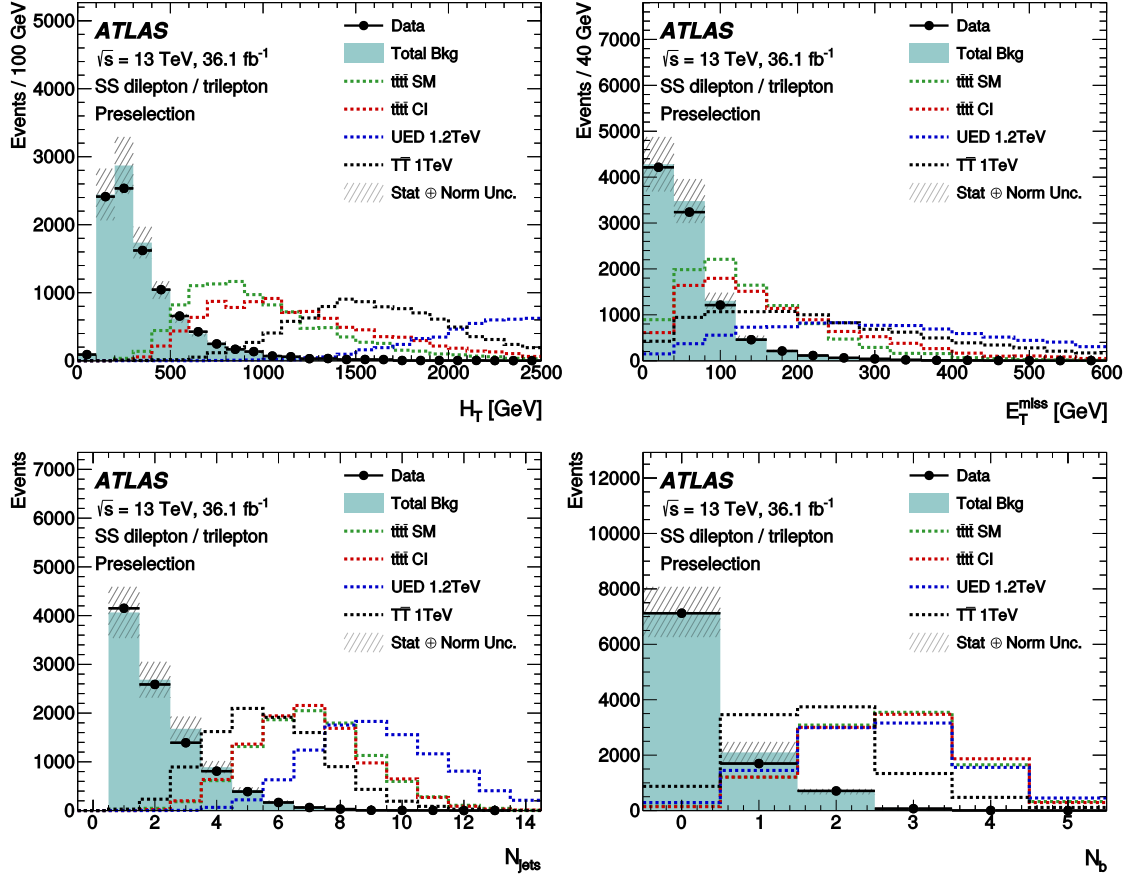


Figure 5.1 – Distributions de grandeurs cinématiques pour le bruit de fond et plusieurs signaux représentatifs. De gauche à droite et de haut en bas, il s’agit des grandeurs H_T , E_T^{miss} , N_j , et N_b . Les distributions des signaux (histogrammes vides en pointillés) sont normalisées suivant l’intégrale de la distribution attendue de bruit de fond (histogramme plein).

La signification statistique d’un modèle dans une région de signal donnée a été approximée avec la formule suivante :

$$Z_A = \frac{s}{\sqrt{s + b + \sigma_b^2}}, \quad (5.1)$$

où s est le nombre d’événements de signal dans la région considérée, b le nombre d’événements de bruit de fond, et σ_b l’incertitude sur b . Cette étude ayant été effectuée de manière préliminaire, l’incertitude sur b n’était pas connue. Une fourchette de valeurs allant de $\sigma_b = 0,10b$ à $\sigma_b = 1,0b$ a par conséquent été utilisée. La combinaison de significations statistiques de différentes régions pour un même signal a été effectuée grâce

à la méthode de Fisher, en définissant une grandeur χ^2 :

$$\chi_F^2 = \sum_{i=1}^{n_{SR}} \ln p_i, \quad (5.2)$$

où i parcourt les numéros des régions combinées, n_{SR} est le nombre de régions de signal, et p_i est la valeur- p correspondant à la signification Z_A dans la région i qui est supposée gaussienne. La valeur- p globale est alors la probabilité d'obtenir une valeur de χ_F^2 supérieure à la valeur calculée, pour une distribution du χ^2 à $2n_{SR}$ degrés de liberté. Cette valeur- p globale est alors convertie en un chiffre de signification gaussienne pour le signal considéré.

Ce travail d'optimisation a fourni plusieurs résultats, notamment les coupures sur les grandeurs H_T et E_T^{miss} des régions de signal décrites dans le tableau 5.2. Il a permis de confirmer l'intérêt de traiter séparément les événements comportant deux leptons de même charge et les événements à trois leptons. L'utilisation de régions de signal avec quatre jets de quarks b a été exclue, puisqu'elle n'apporte qu'un gain modéré de sensibilité et implique une estimation très délicate du bruit de fond de faux leptons décrit dans la section 5.4.2. Enfin, ce travail a surtout permis l'introduction de régions de signal utilisant des valeurs modérées de H_T (régions avec le suffixe « _L » dans le tableau 5.2) pour améliorer la sensibilité de l'analyse au signal de production de quatre quarks top selon le modèle standard.

5.3.3 Définition des régions de signal

Utilisant la méthode présentée dans la section 5.3.2, les régions de signal sont définies en fonction du nombre de leptons dans l'événement N_ℓ : deux leptons de même charge ou trois leptons, ainsi que du nombre de jets de quarks b N_b . Dans chacune de ces régions des coupures sur les variables H_T et E_T^{miss} ont été optimisées. La définition des huit régions de signal de l'analyse est donnée dans le tableau 5.2.

Les coupures sur la grandeur H_T déterminées via la méthode d'optimisation sont toutes très élevées, au-delà de 1 TeV. Il a été démontré que ces coupures réduisaient la sensibilité de l'analyse à la production de quatre quarks top prédite par le modèle standard. Elles sont toutefois utiles car elles augmentent la sensibilité de l'analyse à la plupart des signaux BSM. Ainsi des régions additionnelles acceptant des valeurs de H_T plus faibles ont été définies afin d'améliorer la signification statistique obtenue pour le signal conforme au modèle standard. Il s'agit des régions avec le suffixe « _L ». Ces régions supplémentaires ne concernent que les événements avec au moins trois jets de quarks b , étant donné que les régions de signal comportant moins de jets de quarks b ne sont pas sensibles au signal $t\bar{t}t\bar{t}$ dans le cadre du modèle standard. Une coupure additionnelle sur le nombre de jets N_j est définie dans ces régions car elle augmente la

sensibilité. Une telle coupure serait dommageable aux modèles BSM et est par conséquent absente des autres régions.

Nom	N_j	N_b	N_ℓ	H_T [GeV]	E_T^{miss} [GeV]
SR1b2 ℓ	≥ 1	1	2	> 1000	> 180
SR2b2 ℓ	≥ 2	2	2	> 1200	> 40
SR3b2 ℓ _L	≥ 7	≥ 3	2	$500 - -1200$	> 40
SR3b2 ℓ	≥ 3	≥ 3	2	> 1200	> 100
SR1b3 ℓ	≥ 1	1	3	> 1000	> 140
SR2b3 ℓ	≥ 2	2	3	> 1200	> 100
SR3b3 ℓ _L	≥ 5	≥ 3	3	$500 - -1000$	> 40
SR3b3 ℓ	≥ 3	≥ 3	3	> 1000	> 40

Table 5.2 – Définition des régions de signal de l’analyse.

Les efficacités de sélection pour différents modèles de production de quatre quarks top dans les régions de signal de l’analyse sont présentées dans le tableau 5.3. Si on considère la production dans le cadre du modèle standard, l’efficacité de pré-sélection est d’environ 3 %, tandis que l’efficacité de sélection dans les régions de signal varie entre 0,02 % et 0,4 % environ. Le rapport d’embranchement des événements présentant au moins deux leptons de même charge parmi tous les événements à quatre tops, valant approximativement 12,5 % comme expliqué dans la section 4.1.2, est inclus dans ces efficacités. Le nombre théorique d’événements à quatre tops conformément au modèle standard est d’environ 300 pour une luminosité de $36,1 \text{ fb}^{-1}$. Étant donné les chiffres d’efficacité de sélection, le nombre d’événements attendus dans les régions de signal après sélection varie ainsi entre 0,1 et 1,2 environ.

Les bruits de fond principaux dans les régions de signal proviennent des processus $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$, ainsi que des bruits de fond réductibles (voir section 5.4). La production de deux bosons vecteurs est également un bruit de fond important, mais uniquement dans les régions à un jet de quark b qui ne sont pas sensibles au signal de production de quatre quarks top conforme au modèle standard. Les nombres d’événements observés ainsi que les estimations de bruit de fond dans ces régions de signal sont données dans la section 5.7.1.

5.3.4 Définition des régions de validation

Des régions dites de validation sont également définies afin de vérifier la modélisation du bruit de fond avant de révéler les données dans les régions de signal. Elles sont donc proches des régions de signal, mais orthogonales. Le bruit de fond dans les régions de validation doit être similaire au bruit de fond dans les régions de signal afin de valider

Modèle	Efficacité de pré-sélection [%]	Efficacités de sélection dans les régions de signal [%]			
		SR1b2 ℓ / 3 ℓ	SR2b2 ℓ / 3 ℓ	SR3b2 ℓ _L / 3 ℓ _L	SR3b2 ℓ / 3 ℓ
MS	2,7	0,02 / 0,02	0,11 / 0,04	0,37 / 0,17	0,20 / 0,18
CI	3,0	0,06 / 0,05	0,23 / 0,08	0,30 / 0,16	0,33 / 0,27
2HDM (700 GeV)	3,1	0,02 / 0,03	0,11 / 0,03	0,62 / 0,24	0,19 / 0,17
2UED (1,4 TeV)	3,3	0,27 / 0,16	0,62 / 0,31	8×10^{-4} / 0,0	0,89 / 0,51

Table 5.3 – Efficacités de sélection dans les régions de signal pour différents modèles de production de quatre quarks top. Le rapport d’embranchement des événements présentant au moins deux leptons de même charge parmi tous les événements à quatre tops est inclus dans ces efficacités.

sa modélisation, mais la proportion de signal doit être faible. Ces régions sont ignorées lors de l’analyse statistique décrite dans la section 5.6. Elles sont au nombre de six, et sont construites sur le même modèle que les régions de signal. Leur définition est donnée dans le tableau 5.4. La colonne intitulée « veto SR » exprime le fait que tout événement entrant dans une région de signal n’est pas accepté dans les régions de validation, ce qui garantit l’orthogonalité entre les régions.

Nom	N _j	N _b	N _ℓ	critère cinématique	veto SR
VR1b2 ℓ	≥ 1	1	2	$400 \text{ GeV} < H_T < 2400 \text{ GeV}$ ou $E_T^{\text{miss}} < 40 \text{ GeV}$	✓
VR2b2 ℓ	≥ 2	2	2	$H_T > 400 \text{ GeV}$	✓
VR3b2 ℓ	≥ 3	≥ 3	2	$400 \text{ GeV} < H_T < 1400 \text{ GeV}$ ou $E_T^{\text{miss}} < 40 \text{ GeV}$	✓
VR1b3 ℓ	≥ 1	1	3	$400 \text{ GeV} < H_T < 2000 \text{ GeV}$ ou $E_T^{\text{miss}} < 40 \text{ GeV}$	✓
VR2b3 ℓ	≥ 2	2	3	$400 \text{ GeV} < H_T < 2400 \text{ GeV}$ ou $E_T^{\text{miss}} < 40 \text{ GeV}$	✓
VR3b3 ℓ	≥ 3	≥ 3	3	$H_T > 400 \text{ GeV}$	✓

Table 5.4 – Définition des régions de validation de l’analyse. La colonne intitulée « veto SR » exprime le fait que tout événement entrant dans une région de signal n’est pas accepté dans les régions de validation.

L’accord entre l’estimation du bruit de fond et les données observées dans les régions de validation est décrit dans la section 5.4.3.

5.4 Estimation des bruits de fond et des signaux

Plusieurs processus physiques autres que la production de quatre quarks top sont susceptibles de générer deux leptons de même charge électrique et des jets de quarks b . Il s'agit des différents bruits de fond de l'analyse, qui peuvent être séparés en deux catégories :

- les bruits de fond irréductibles font référence aux processus produisant précisément les particules recherchées dans l'analyse, notamment deux leptons ayant la même charge électrique ;
- les bruits de fond réductibles quant à eux sont issus d'erreurs de reconstruction des objets dans le détecteur, entraînant la détection de deux leptons de même charge issus de processus qui ne devraient pas pouvoir les produire.

Les bruits de fond réductibles sont estimés à partir des données en utilisant des méthodes dédiées présentées dans la section 5.4.2, tandis que les bruits de fond irréductibles sont modélisés grâce à des simulations Monte-Carlo comme expliqué dans la section 5.4.1.

Les différents signaux étudiés dans cette analyse sont également modélisés via des simulations Monte-Carlo. En ce qui concerne les signaux de production de quatre quarks top, selon le modèle standard ou selon les modèles BSM considérés, le générateur utilisé est MG5_AMC@NLO [16] et l'hadronisation est réalisée avec PYTHIA [17]. La modélisation des processus physiques par des simulations Monte-Carlo est décrite dans la section 2.4.

5.4.1 Bruits de fond irréductibles

De nombreux processus physiques sont considérés en tant que bruits de fond irréductibles de cette analyse :

- les bruits de fond principaux sont les processus $t\bar{t}V$ et VV où V désigne un boson W ou un boson Z , ainsi que $t\bar{t}H$;
- d'autres contributions moins importantes sont catégorisées en tant que « other » : $t\bar{t}WW$, tWZ , tZ , $t\bar{t}t$, VH , VVV ;
- la production de quatre quarks top conformément au modèle standard est également un bruit de fond lorsque les modèles BSM sont étudiés.

Chacun des processus listés peut produire un état final comportant au moins deux leptons de même charge, dans le cas de désintégrations $t \rightarrow Wb$ puis $W \rightarrow l\nu_l$, ainsi que $Z \rightarrow l^+l^-$. Par exemple, le processus $t\bar{t}W$ peut produire trois leptons suite à la désintégration leptonique des deux quarks t ainsi que du boson W , ou encore deux leptons de même charge dans le cas d'une désintégration leptonique du boson W et du quark t ayant la même charge.

Ces processus sont simulés avec MG5_AMC@NLO et PYTHIA, sauf les processus à deux ou trois bosons vecteurs VV et VVV qui sont simulés avec SHERPA [18] et les processus VH qui sont simulés avec PYTHIA seul.

Le calcul des éléments de matrice est effectué au second ordre dominant (NLO) pour les processus $t\bar{t}V$, $t\bar{t}H$ et tWZ , et à l'ordre dominant (LO) pour $t\bar{t}t$, $t\bar{t}t\bar{t}$, $t\bar{t}WW$ et tZ . Les sections efficaces quant à elles sont calculées au NLO en QCD pour tous les processus sauf tZ pour lequel le calcul LO est utilisé. Lors de l'étude d'un échantillon Monte-Carlo, le nombre d'événements sélectionnés est normalisé par sa section efficace théorique, ainsi que par la luminosité intégrée.

Les informations au niveau partonique des échantillons Monte-Carlo sont utilisées afin de s'assurer que les leptons sélectionnés possèdent une charge adéquate. Ceci permet d'éviter un double comptage des événements dans lesquels la charge d'un électron a été mal reconstruite. En effet ce type de bruit de fond est estimé grâce à des méthodes spécifiques, présentées dans la section 5.4.2.

5.4.2 Bruits de fond réductibles

Les bruits de fond réductibles proviennent de processus ne produisant pas deux leptons de même charge électrique. Une erreur de reconstruction dans un tel événement peut toutefois le faire entrer dans une région de signal. Ce type d'événements rentrant à tort dans les régions de signal est séparé en deux catégories :

- les événements qui contiennent un jet reconstruit en tant que lepton ou encore un lepton ne provenant pas de l'interaction initiale, respectivement appelés « faux lepton » et « lepton non-prompt » ;
- les événements associés à une mauvaise identification de la charge électrique d'un électron, entraînant la reconstruction d'électrons de même charge au lieu d'électrons de charge opposée.

Ces événements peuvent provenir de nombreux processus physiques différents, le processus majoritaire étant ici la production d'une paire de quarks top $pp \rightarrow t\bar{t}$. Les probabilités de tels événements étant faibles, il est difficile de les modéliser de manière satisfaisante et avec une faible incertitude statistique. Cependant les sections efficaces des processus sous-jacents se situent plusieurs ordres de grandeur au-dessus de la section efficace de production de quatre quarks top dans le cadre du modèle standard. Ainsi, ce type de bruit de fond est important mais souvent mal modélisé via des simulations Monte-Carlo, en particulier dans une analyse avec un espace des phases très restreint. C'est pourquoi des techniques d'estimation à partir des données, ou *data-driven*, sont mises en place afin d'estimer ce type de bruit de fond.

Faux leptons et leptons non-prompts

Certains événements comportent des leptons qui ne proviennent pas réellement d'une désintégration d'un boson W issu de la désintégration d'un quark t . On a distingué précédemment les leptons « faux » et les leptons « non-prompts » aussi qualifiés de

« secondaires ». Dans la suite de ce texte ces deux catégories sont rassemblées sous la dénomination de « faux » leptons, et les autres leptons sont appelés « vrais » leptons. Plusieurs sources sont envisageables pour de tels leptons :

- une désintégration semi-leptonique dans un jet de quark lourd ;
- la conversion d'un photon en paire e^+e^- ;
- un électron qui émet un photon se convertissant un paire e^+e^- , processus appelé « trident » (cas particulier de conversion) ;
- la reconstruction d'un jet en tant qu'électron, par exemple à cause de la proximité entre un photon ou pion et un hadron chargé.

Les événements avec des faux leptons peuvent provenir d'un grande variété de processus physiques différents, les plus importants étant la production d'une paire de quarks top, d'un unique quark top accompagné d'autres particules ou encore d'un boson vecteur accompagné de jets. Une méthode d'estimation à partir des données appelée méthode de la matrice, ou *Matrix Method*, est utilisée.

Les leptons reconstruits sont séparés en deux catégories, comme expliqué dans la section 5.2 : les leptons stricts et les leptons lâches. L'échantillon strict subit une sélection plus restrictive et est donc inclus dans l'échantillon lâche. La méthode de la matrice est fondée sur l'estimation du taux de leptons lâches vrais qui passent les coupures strictes, ainsi que du taux de leptons lâches faux qui passent ces mêmes coupures. Dans le cas d'un événements avec deux leptons, on a la relation matricielle suivante :

$$\begin{pmatrix} N^{tt} \\ N^{t\bar{t}} \\ N^{\bar{t}t} \\ N^{\bar{t}\bar{t}} \end{pmatrix} = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1 \bar{r}_2 & r_1 \bar{f}_2 & f_1 \bar{r}_2 & f_1 \bar{f}_2 \\ \bar{r}_1 r_2 & \bar{r}_1 f_2 & \bar{f}_1 r_2 & \bar{f}_1 f_2 \\ \bar{r}_1 \bar{r}_2 & \bar{r}_1 \bar{f}_2 & \bar{f}_1 \bar{r}_2 & \bar{f}_1 \bar{f}_2 \end{pmatrix} \begin{pmatrix} N_{rr}^{ll} \\ N_{rf}^{ll} \\ N_{fr}^{ll} \\ N_{ff}^{ll} \end{pmatrix} \quad (5.3)$$

où t désigne un lepton strict, $\bar{t}$ un lepton lâche et non-strict, et l un lepton lâche strict ou non-strict. Ainsi par exemple N^{tt} désigne le nombre d'événements avec deux leptons stricts et $N^{t\bar{t}}$ le nombre d'événements avec un lepton strict et un lepton lâche et non-strict ordonnés par p_T décroissant. Les lettres r et f font référence aux vrais et faux leptons respectivement. Les termes N_{rr}^{ll} et N_{rf}^{ll} désignent donc respectivement le nombre d'événements avec deux vrais leptons lâches, et le nombre d'événements avec un vrai lepton et un faux lepton lâches, ordonnés par p_T décroissant. Les éléments de la matrice sont les taux, r_1, r_2 (f_1, f_2) désignant la probabilité des vrais (faux) leptons lâches rangés par ordre décroissant de p_T de passer les coupures strictes. $\bar{r}$ et $\bar{f}$ valent respectivement $1 - r$ et $1 - f$. La relation matricielle se généralise dans le cas d'un événement contenant trois leptons.

L'équation 5.3 montre que si les taux r_1, r_2, f_1, f_2 , ainsi que les nombres d'événements observés $N^{tt}, N^{t\bar{t}}$, etc. sont connus, alors il est possible d'inverser la relation matricielle et d'en déduire le nombre d'événements avec des leptons vrais et faux N_{rr}^{ll}, N_{rf}^{ll} , etc. Ces

informations permettent d'estimer le bruit de fond dû aux faux leptons. Il est également possible d'appliquer cette méthode à une distribution en l'utilisant dans chaque intervalle de l'histogramme correspondant.

La relation matricielle est résolue pour chaque événement. Par exemple dans le cas d'un événement avec deux leptons stricts, N^{tt} vaut 1 et les autres composantes du vecteur de gauche de l'équation 5.3 valent 0. La relation matricielle est inversée afin d'obtenir le membre de droite de l'équation. Les termes $N_{\text{rf}}^{\text{ll}}$, $N_{\text{ff}}^{\text{ll}}$, et $N_{\text{ff}}^{\text{ll}}$ obtenus sont vus comme des poids pour l'événement en question. La somme de tous ces poids pour tous les événements donne l'estimation du bruit de fond dû aux faux leptons. Cependant ces poids peuvent être négatifs, ce qui est problématique quand le nombre d'événements disponibles est faible. De plus il peut y avoir des instabilités dues à de grands poids notamment si les taux des vrais et faux leptons sont trop similaires. C'est pourquoi la méthode de la matrice utilisée dans cette analyse a été modifiée pour y inclure un terme de vraisemblance Poissonnien afin d'obtenir le nombre d'événements avec des faux leptons le plus probable. Cette méthode permet d'obtenir une estimation positive du nombre d'événements dus aux faux leptons de par l'utilisation de la vraisemblance Poissonnienne. Elle est également plus précise et plus stable que la méthode de la matrice usuelle en particulier quand les taux de faux leptons lâches passant les coupures strictes sont élevés ou quand les taux des leptons faux et vrais varient fortement d'un lepton à l'autre [19].

L'extraction des taux est un élément central de la méthode de la matrice. Dans ce but, des événements de données avec au moins un lepton sont analysés dans des régions de contrôle enrichies en vrais ou faux leptons. La définition de ces régions est présentée dans le tableau 5.5. Les taux sont estimés séparément pour les événements sans jet de quark b ($0b$) et ceux avec au moins un jet de quark b ($\geq 1b$), et séparément pour les événements contenant un seul jet ($1j$) et ceux avec au moins deux jets ($\geq 2j$). On considère les quatre combinaisons possibles $\{0b, \geq 1b\} \times \{1j, \geq 2j\}$ pour chaque ligne du tableau, donc quatre régions de contrôle par quantité mesurée. Seuls les taux $\geq 1b$ sont utilisés dans les régions de signal de l'analyse, c'est-à-dire les taux estimés dans la région $\{\geq 1b\} \times \{1j\}$ pour les événements avec un seul jet, et ceux estimés dans la région $\{\geq 1b\} \times \{\geq 2j\}$ pour les autres événements. Les autres taux sont quant à eux utilisés dans certaines régions de validation.

Les régions de contrôle pour les électrons sont choisies sachant que la plupart des vrais électrons sélectionnés proviennent d'une désintégration d'un boson W tandis que la plupart des faux électrons proviennent d'événements avec de multiples jets. Ces derniers ne contiennent pas de neutrinos énergétiques, contrairement aux événements contenant des vrais électrons. C'est pourquoi les coupures choisies sont $E_{\text{T}}^{\text{miss}} > 150 \text{ GeV}$ pour les vrais électrons, et $m_{\text{T}}(W) < 20 \text{ GeV}$ et $E_{\text{T}}^{\text{miss}} + m_{\text{T}}(W) < 60 \text{ GeV}$ pour les faux électrons, où $m_{\text{T}}(W)$ est la masse transverse du lepton. Dans le cas des muons une coupure $m_{\text{T}}(W) > 100 \text{ GeV}$ permet de sélectionner une majorité de vrais leptons provenant

Quantité mesurée	N_b et N_j	Critères additionnels
électron r	$\{0b, \geq 1b\} \times \{1j, \geq 2j\}$	$E_T^{\text{miss}} > 150 \text{ GeV}$
électron f	$\{0b, \geq 1b\} \times \{1j, \geq 2j\}$	$m_T(W) < 20 \text{ GeV}$ $E_T^{\text{miss}} + m_T(W) < 60 \text{ GeV}$
muon r	$\{0b, \geq 1b\} \times \{1j, \geq 2j\}$	$m_T(W) > 100 \text{ GeV}$
muon f	$\{0b, \geq 1b\} \times \{1j, \geq 2j\}$	$ d_0 /\sigma_{d_0} > 5$

Table 5.5 – Définition des régions de contrôle utilisées pour mesurer les taux des vrais et faux leptons. Tous les événements considérés doivent contenir au moins un lepton lâche. La notation $\{0b, \geq 1b\} \times \{1j, \geq 2j\}$ signifie que les taux sont estimés séparément pour les quatre combinaisons de coupures sur N_b et sur N_j .

d’une désintégration d’un boson W . Les faux muons proviennent principalement de désintégrations de jets de quarks c ou b , et sont sélectionnés par leur incompatibilité avec le vertex primaire : $|d_0|/\sigma_{d_0} > 5$.

Les taux sont estimés en fonction des variables cinématiques les plus influentes : p_T , $|\eta|$, et $\Delta R(\ell, \text{jet})$ qui est la distance entre le lepton ℓ et le jet le plus proche. On rappelle la définition de cette grandeur : $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Les taux des électrons sont exprimés comme le produit de deux distributions : taux en fonction de $(p_T, |\eta|)$, et taux en fonction de $\Delta R(e, \text{jet})$. Le résultat est normalisé en divisant par le taux global moyen. Les taux des muons sont exprimés comme le produit de trois distributions : taux en fonction de p_T , taux en fonction de $|\eta|$, et taux en fonction de $\Delta R(\mu, \text{jet})$. Le résultat est normalisé en divisant par le taux global moyen au carré. La distribution bi-dimensionnelle $(p_T, |\eta|)$ est utilisée pour les électrons car les variables p_T et $|\eta|$ sont corrélées. Cette corrélation n’est pas observée dans le cas des muons. Les valeurs des taux utilisés dans la matrice de l’équation 5.3 sont donc choisies en accord avec les valeurs de p_T , $|\eta|$, et $\Delta R(\ell, \text{jet})$ des leptons dans l’événement.

Malgré les coupures définies pour les régions enrichies en faux leptons, il reste certains événements contenant des vrais leptons dans ces régions de contrôle. Cette contribution est estimée à partir d’échantillons de simulation Monte-Carlo, puis soustraite. Dans les régions avec au moins un jet de quark b , la proportion de ces événements est de l’ordre de 5 % maximum pour les muons, et se situe entre 30 % et 60 % environ pour les électrons, selon le déclenchement considéré.

La figure 5.2 présente les taux pour les électrons, obtenus dans le cas $N_b \geq 1$ et $N_j \geq 2$. La mesure est effectuée dans les données prises en 2016. Les distributions des taux en fonction de $(p_T, |\eta|)$, et en fonction de $\Delta R(e, \text{jet})$ sont montrées. La figure 5.3 présente les taux pour les muons, obtenus dans les mêmes conditions et avec les mêmes données. Les distributions des taux en fonction de p_T , $|\eta|$, et $\Delta R(\mu, \text{jet})$ sont présentées.

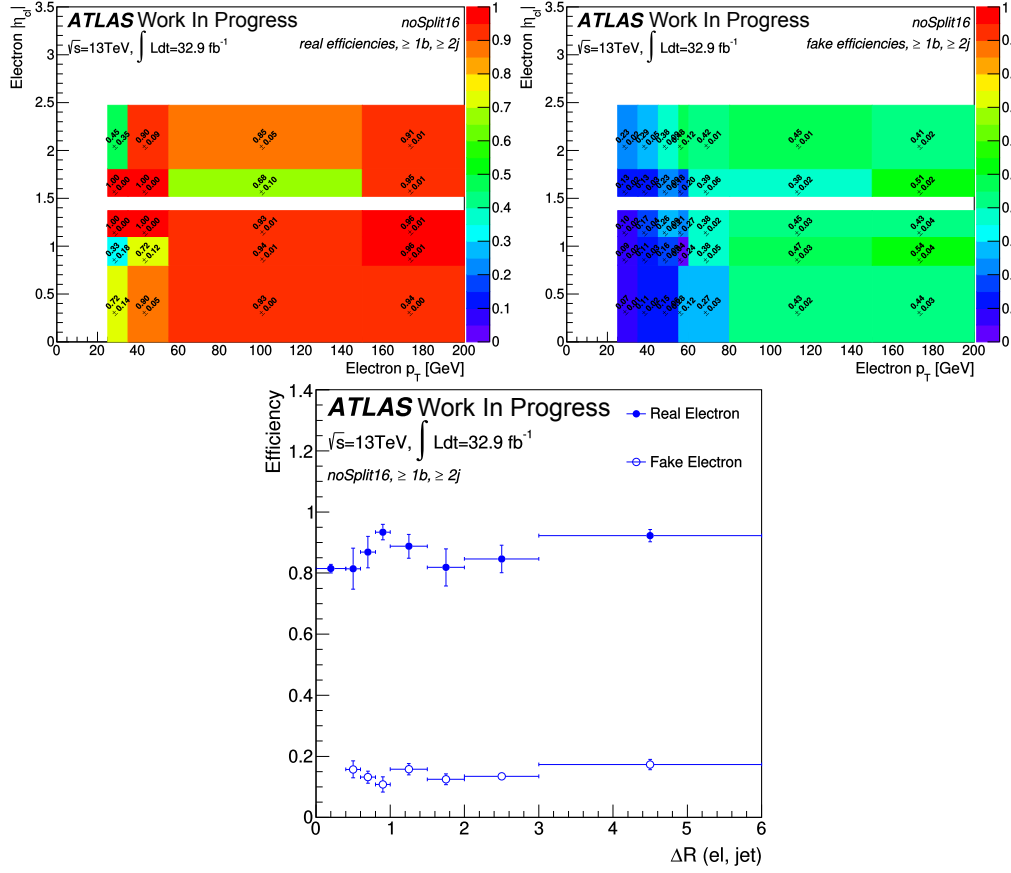


Figure 5.2 – Taux d’électrons lâches passant les coupures strictes, en fonction des valeurs de $(p_T, |\eta|)$, et $\Delta R(e, \text{jet})$. La mesure est effectuée dans les données prises en 2016. Les taux sont estimés pour les événements avec au moins un jet de quark b et au moins deux jets. Les taux des vrais électrons sont représentés sur la figure en haut à gauche, ainsi que sur la figure du bas à l’aide de points bleus. Ceux des faux électrons sont représentés sur la figure en haut à droite, ainsi que sur la figure du bas à l’aide de cercles bleus. Les figures du haut représentent les taux en fonction de $(p_T, |\eta|)$, et la figure du bas les représente en fonction de $\Delta R(e, \text{jet})$.

Les taux r des vrais électrons varient typiquement de 70 % à 95 %, et les taux f des faux électrons de 8 % à 50 %, selon les valeurs des paramètres p_T , $|\eta|$, et $\Delta R(e, \text{jet})$. Quant aux muons, ils présentent des valeurs de r entre 55 % et 97 %, et des valeurs de f entre 7 % et 30 %. De manière générale les valeurs de r et f diminuent avec la distance au jet le plus proche, et augmentent avec le p_T du lepton.

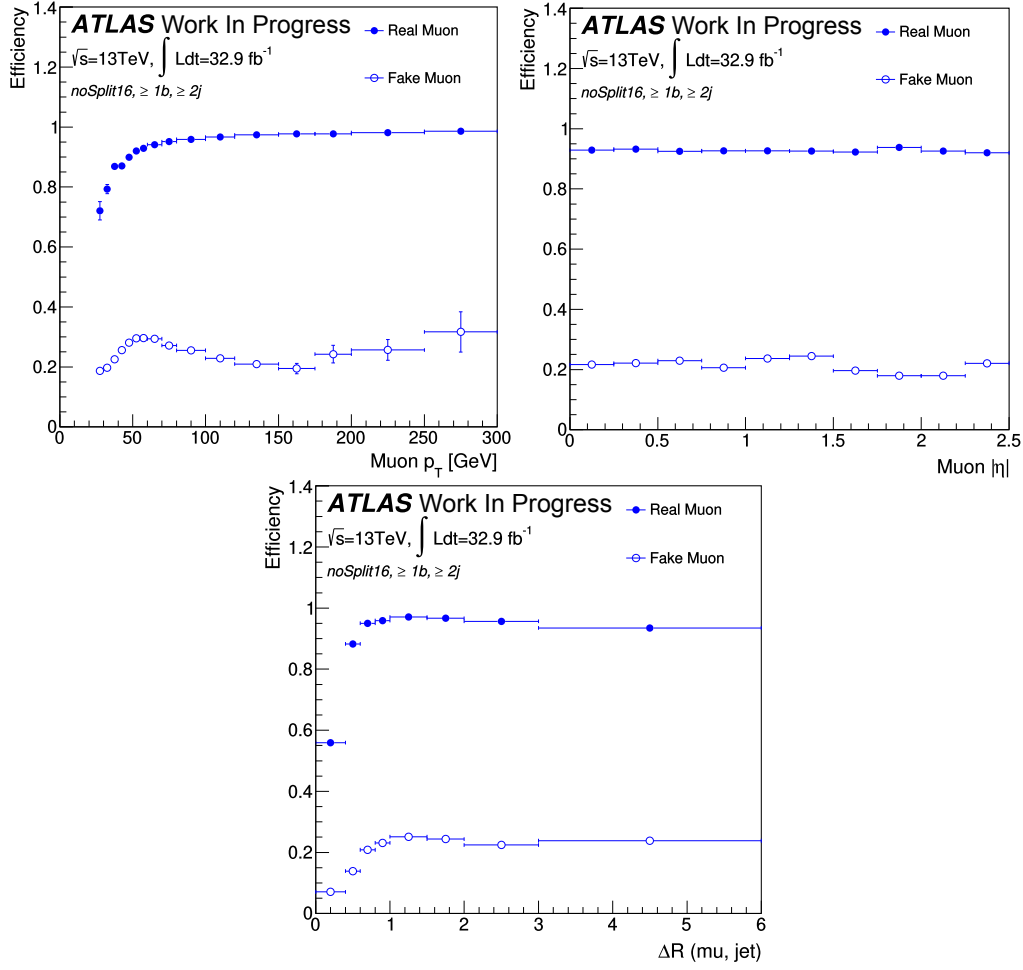


Figure 5.3 – Taux de muons lâches passant les coupures strictes, en fonction des valeurs de p_T , $|\eta|$, et $\Delta R(\mu, \text{jet})$. La mesure est effectuée dans les données prises en 2016. Les taux sont estimés pour les événements avec au moins un jet de quark b et au moins deux jets. Les taux des vrais électrons sont représentés à l’aide de points bleus, et ceux des faux électrons à l’aide de cercles bleus. Les figures représentent les taux en fonction des valeurs de p_T (haut, gauche), des valeurs de $|\eta|$ (haut, droite), et en fonction de $\Delta R(\mu, \text{jet})$ (bas).

Mauvaise identification de la charge d’un électron

De nombreux processus physiques produisent deux leptons de charge électrique opposée. La mauvaise identification de la charge de l’un d’entre eux entraîne la reconstruction d’une paire de leptons de même charge, pouvant potentiellement entrer dans les régions de signal de cette analyse. Ce type d’erreur peut provenir de deux sources :

- il peut être difficile pour le trajectographe de reconstruire la charge d’une particule, par exemple dans le cas d’une particule très énergétique ayant une trace très peu courbée ;
- un électron peut émettre un photon de rayonnement de freinage ou *bremsstrahlung* se désintégrant en paire e^+e^- , soit $e^\pm \rightarrow e^\pm \gamma \rightarrow e^\pm e^+e^-$. Lors d’un tel phénomène qualifié de « trident », l’amas d’énergie dans le calorimètre peut être associé à la mauvaise trace d’électron, entraînant une erreur sur la charge.

Ces phénomènes sont négligeables pour les muons grâce au grand bras de levier du spectromètre et le fait que la charge des muons est mesurée à la fois dans le détecteur interne et dans le spectromètre. De plus, les muons émettent beaucoup moins de rayonnement de freinage que les électrons de par leur plus grande masse. On n’impose aucune condition sur la charge des leptons dans les régions de signal avec trois leptons, c’est pourquoi ce bruit de fond n’est défini que dans les régions avec deux leptons ayant la même charge électrique $e^\pm e^\pm$ et $e^\pm \mu^\pm$.

Le taux de mauvaise identification de la charge d’un électron est noté ε et dépend des valeurs de $|\eta|$ et p_T de l’électron. Si ce taux ε est connu en fonction de $|\eta|$ et p_T , alors il est possible de l’appliquer sur des événements avec deux leptons de charge opposée au lieu de deux leptons de même charge, afin d’obtenir une estimation du bruit de fond. Pour estimer ce bruit de fond, la sélection des régions de signal est reproduite en modifiant le critère de charge afin de sélectionner deux leptons de charge opposée et non de charge identique. Parmi les N événements avec deux leptons de charge opposée, on peut distinguer N_{OC} événements avec des leptons à la charge bien identifiée, et N_{SC} événements avec une erreur sur la charge d’un des deux leptons. Si la charge des deux leptons est mal identifiée, l’événement se retrouve dans la catégorie N_{OC} . Si le taux de mauvaise identification de la charge d’un seul des deux leptons est appelé p , alors le nombre total d’événements avec deux leptons de charge opposée est

$$N = N_{OC} + N_{SC} = (1 - p)N + pN \quad (5.4)$$

avec $p = \varepsilon_1(1 - \varepsilon_2) + \varepsilon_2(1 - \varepsilon_1) = \varepsilon_1 + \varepsilon_2 - 2\varepsilon_1\varepsilon_2$, où ε_1 et ε_2 sont respectivement les taux de mauvaise identification du premier et du second lepton. Le taux pour le premier lepton est en effet supposé indépendant de celui pour le second lepton. Si l’un des deux leptons est un muon, ce taux est supposé être nul. Les indices $_1$ et $_2$ désignent chacun un intervalle bi-dimensionnel $(|\eta|, p_T)$ spécifique, au sein duquel les valeurs de ε sont considérées comme constantes.

Le rapport w est alors défini pour chaque intervalle $(|\eta|, p_T)$:

$$w = \frac{N_{SC}}{N_{OC}} = \frac{p}{1 - p} = \frac{\varepsilon_1 + \varepsilon_2 - 2\varepsilon_1\varepsilon_2}{1 - \varepsilon_1 - \varepsilon_2 + 2\varepsilon_1\varepsilon_2}. \quad (5.5)$$

L’estimation du bruit de fond de mauvaise identification de la charge d’un électron est effectuée par région de signal, en calculant puis en appliquant le poids w à chaque

événement observé avec deux leptons de charge opposée dans les données. D'après l'équation 5.5 le poids appliqué à chaque événement avec deux leptons de charge opposée fournit le nombre d'événements avec deux leptons de même charge pour chaque intervalle $(|\eta|, p_T)$: $N_{SC} = w \times N_{OC}$. En sommant sur tous les intervalles, on obtient bien une estimation du nombre d'événements de données entrant dans les régions de signal via une mauvaise identification.

Les taux ε sont mesurés dans une région de contrôle enrichie en événements $Z \rightarrow e^+e^-$, appelé « région du pic du Z ». Elle est définie par une condition sur la masse invariante de la paire d'électrons : $81 \text{ GeV} < m_{ee} < 101 \text{ GeV}$. Seuls les leptons stricts sont pris en compte dans cette étude. La majorité des événements observés comportent deux leptons de charge opposée, conformément à ce qu'on pourrait attendre du processus $Z \rightarrow e^+e^-$. Toutefois certains événements présentent deux leptons ayant la même charge. En gardant les mêmes notations que précédemment, on a donc :

$$N_{SC}^{ij} = pN^{ij} = N^{ij} (\varepsilon_1 + \varepsilon_2 - 2\varepsilon_1\varepsilon_2) \quad (5.6)$$

où i et j sont les indices représentant les intervalles $(|\eta|, p_T)$ des deux électrons. Les valeurs ε dans les différents intervalles sont alors extraites en maximisant la vraisemblance Poissonnienne que le nombre d'événements de données observé dans chacun des intervalles $(|\eta|, p_T)$ avec des paires de même charge corresponde à l'estimation de l'équation 5.6.

Afin d'éliminer les bruits de fond provenant d'autres processus que $Z \rightarrow e^+e^-$ dans la région du pic du Z , des bandes latérales situées à droite et à gauche et d'une largeur de 10 GeV sont utilisées. Les bruits de fond sont estimés à l'aide d'échantillons Monte-Carlo, puis soustraits aux grandeurs N et N_{SC} .

La méthode d'estimation du bruit de fond dû aux faux leptons inclut les événements de type trident $e^\pm \rightarrow e^\pm \gamma \rightarrow e^\pm e^+ e^-$. Afin d'éviter un double comptage de certains événements en tant que bruit de fond de mauvaise identification de la charge et de bruit de fond dû aux faux leptons, il est nécessaire de soustraire la contribution due aux faux leptons des grandeurs N et N_{SC} . Cette contribution est estimée via la méthode de la matrice. Par souci de cohérence avec cette méthode, une coupure $N_j \geq 1$ est appliquée aux événements de la région du pic du Z .

Les taux ε par intervalle $(|\eta|, p_T)$ sont représentés sur la figure 5.4. Ils augmentent globalement avec les valeurs de $|\eta|$ et p_T . Leurs valeurs s'étendent de $2 \cdot 10^{-5}$ à $1 \cdot 10^{-2}$ dans le tonneau, et peuvent aller jusqu'à 10 % dans le bouchon.

5.4.3 Accord entre simulation et données dans les régions de validation

La modélisation du bruit de fond de l'analyse est vérifiée dans les régions de validation définies dans la section 5.3.4. Dans chacune de ces régions, le nombre d'événements de données est comparé au nombre d'événements estimé de bruit de fond. Cette comparaison est caractérisée par une valeur- p , à savoir la probabilité que le bruit de fond attendu

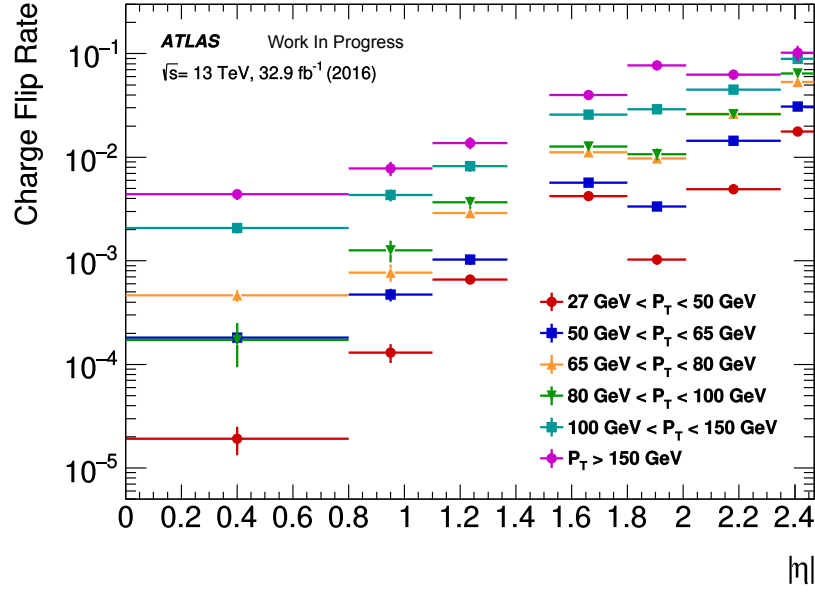


Figure 5.4 – Taux ε de mauvaise identification de la charge d’un électron par intervalle $(|\eta|, p_T)$. L’estimation de ces taux a été réalisée à partir des données prises en 2016.

fluctue vers des valeurs supérieures ou égales à celle des données observées. La valeur- p la plus basse obtenue ainsi est de 0,10 dans la région VR2b2 ℓ .

On s’intéresse aussi tout particulièrement à la modélisation des grandeurs E_T^{miss} et H_T qui sont utilisées pour définir les régions de signal et de validation. Les distributions de ces grandeurs dans les régions de validation sont montrées sur la figure 5.5 pour E_T^{miss} et sur la figure 5.6 pour H_T . Les valeurs de χ^2 représentant l’accord entre données et bruit de fond sont calculées en prenant en compte les incertitudes systématiques ainsi que les corrélations entre les intervalles des histogrammes. La probabilité qu’une fluctuation du bruit de fond produise une distribution ayant une compatibilité identique ou moins bonne avec les données peut ainsi être calculée. La plus petite valeur obtenue parmi les distributions considérées est de 0,02, pour la distribution de la grandeur E_T^{miss} dans la région VR1b2 ℓ . L’accord entre estimations de bruit de fond et données dans les régions de validation est considéré satisfaisant.

5.5 Incertitudes systématiques

L’ensemble des incertitudes prises en compte dans l’analyse est décrit dans cette section. En ce qui concerne les signaux, aucune incertitude de modélisation n’est considérée. La plupart des incertitudes sont constituées d’un couple de variations positive et négative du nombre d’événements, d’autres sont définies comme une valeur unique qu’il faut appliquer symétriquement comme erreur positive et négative.

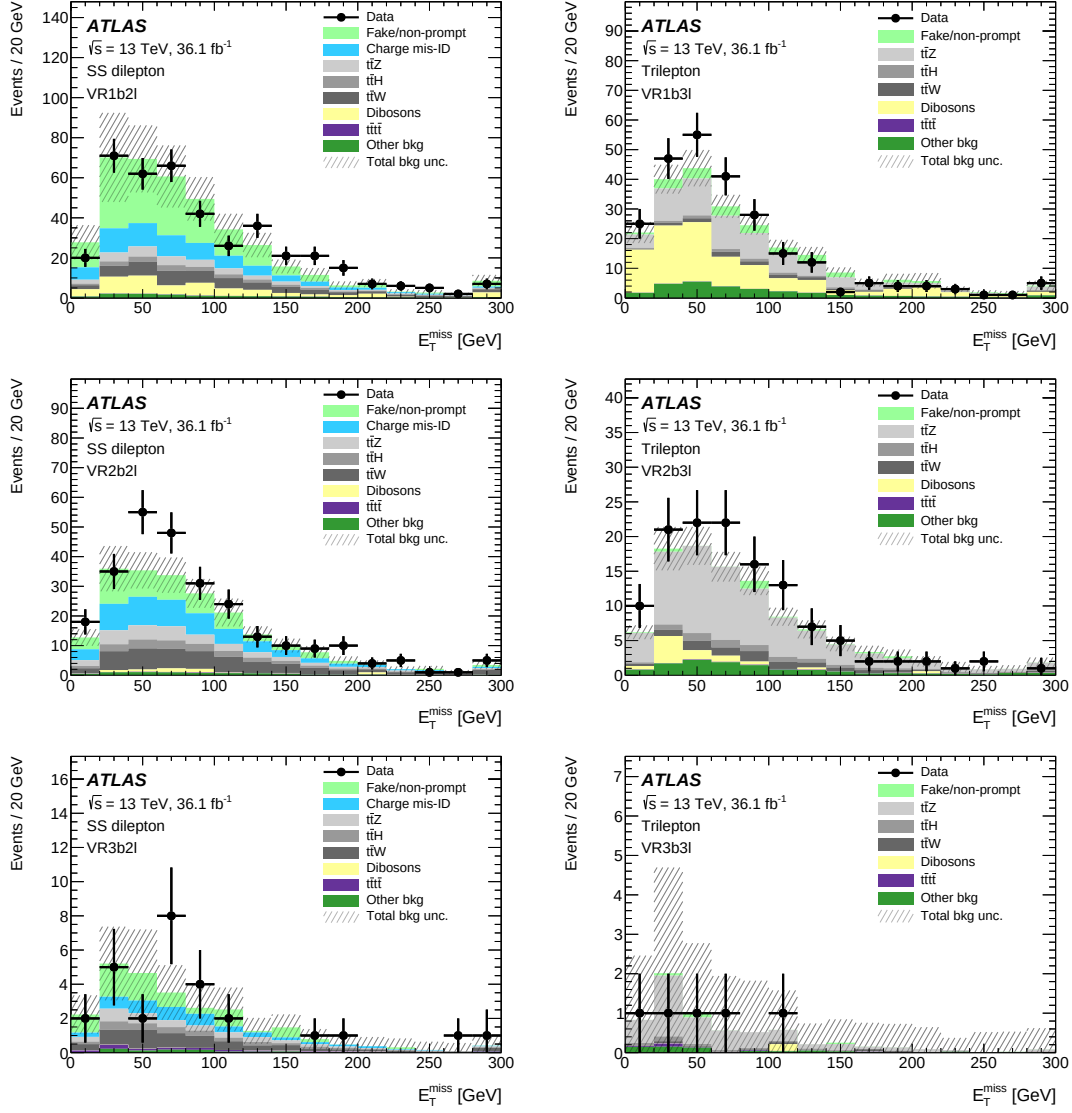


Figure 5.5 – Distributions de la grandeur E_T^{miss} dans les régions de validation. De haut en bas il s'agit des régions 1b, 2b, puis 3b. Les régions 2ℓ sont à gauche, et les régions 3ℓ à droite. Les histogrammes colorés empilés représentent l'estimation des bruits de fond, et les points noirs sont les données. Les bandes hachurées représentent les incertitudes statistiques et systématiques de l'ensemble des bruits de fond.

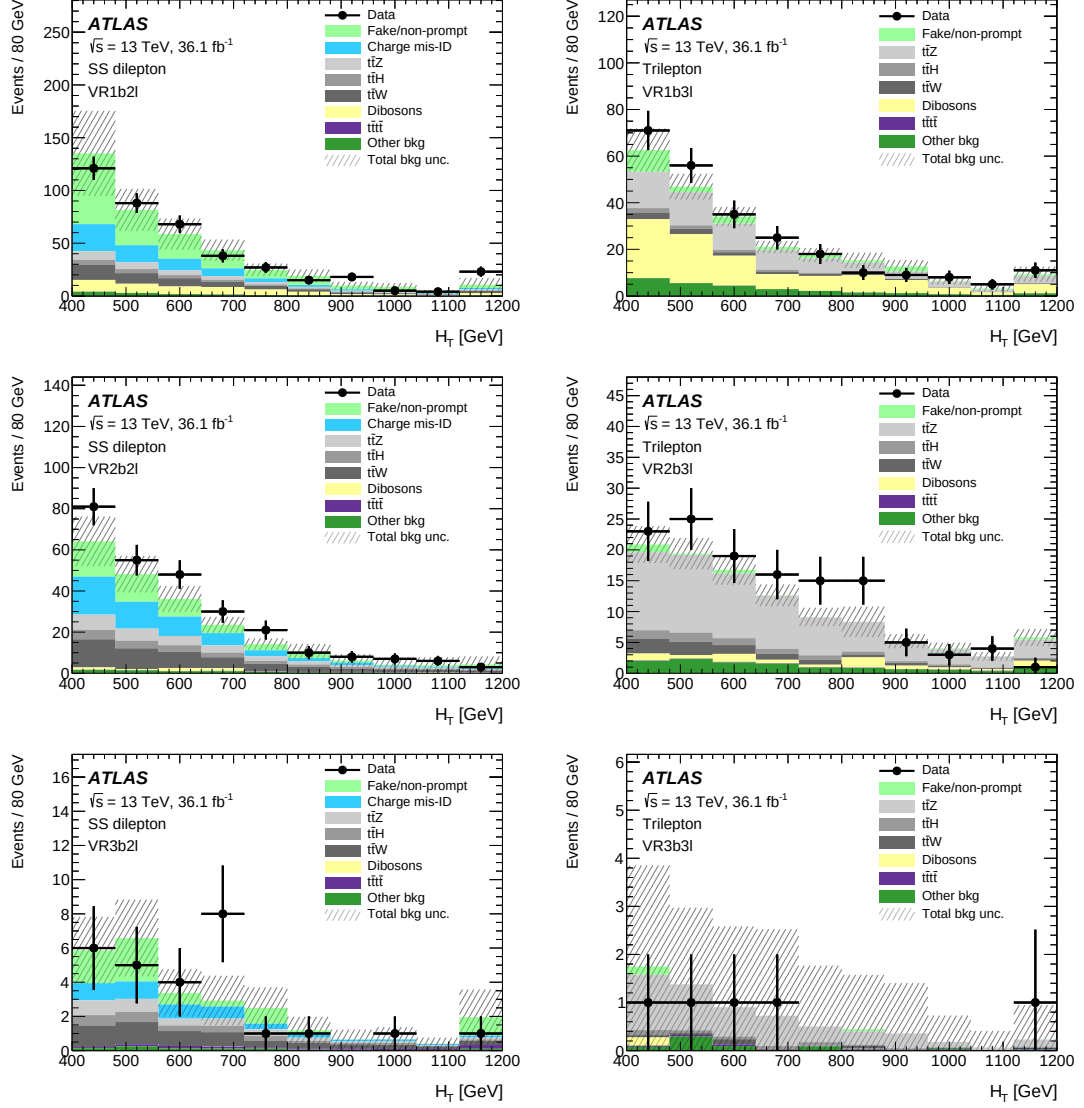


Figure 5.6 – Distributions de la grandeur H_T dans les régions de validation. De haut en bas il s’agit des régions 1b, 2b, puis 3b. Les régions 2 ℓ sont à gauche, et les régions 3 ℓ à droite. Les histogrammes colorés empilés représentent l’estimation des bruits de fond, et les points noirs sont les données. Les bandes hachurées représentent les incertitudes statistiques et systématiques de l’ensemble des bruits de fond.

5.5.1 Reconstruction des objets

La reconstruction des objets utilisés dans cette analyse (leptons, jets, etc.) comporte plusieurs incertitudes. Ces objets sont calibrés via des études dédiées qui aboutissent à des facteurs de normalisation pour corriger certains désaccords entre simulation et données. Ces facteurs ainsi que leurs incertitudes sont introduits en tant que paramètres de nuisance dans le modèle statistique de l'analyse qui est présenté dans la section 5.6.

Des incertitudes liées à la reconstruction, l'identification, l'isolation et le déclenchement sont prises en compte pour les leptons. Des incertitudes sur l'échelle d'énergie et la résolution en énergie sont considérées pour les leptons et les jets. L'étiquetage des jets de quarks b présente des incertitudes sur l'identification de jets de quarks b , ainsi que la contamination issue des jets de quarks c et des jets légers. Toutes ces incertitudes impactent aussi le calcul de l'énergie transverse manquante, qui possède donc ses propres incertitudes systématiques.

5.5.2 Luminosité

L'incertitude sur la luminosité intégrée pour les données prises en 2015 et en 2016 est de $\pm 2.1\%$. Elle est estimée via la calibration de l'échelle de luminosité utilisant des scans Van de Meer, similairement à la méthode présentée dans la référence [20], ainsi que grâce au détecteur LUCID-2 [21]. Elle est appliquée de manière corrélée à toutes les régions de signal.

5.5.3 Bruits de fond irréductibles

Les incertitudes systématiques appliquées aux bruits de fond irréductibles sont liées à la modélisation des processus physiques associés. Ces incertitudes sont de deux types : une incertitude sur la section efficace du processus utilisée pour normaliser les échantillons Monte-Carlo, et une incertitude associée aux coupures des régions de signal. En effet les critères de sélection d'une région de signal sont définis par des coupures sur certaines grandeurs physiques (H_T , N_b , etc.) pouvant être impactées par les paramètres des simulations Monte-Carlo utilisées. La forme des distributions de ces variables peut notamment varier, ce qui entraîne une incertitude dite de forme sur le nombre d'événements sélectionnés dans les régions considérées.

Ces deux incertitudes sont regroupées dans une même catégorie et impactent le nombre d'événements total d'un type de bruit de fond dans chaque région de signal. La valeur de ces incertitudes peut varier d'une région de signal à l'autre. Elles sont discutées séparément ci-dessous.

Incertitudes sur la section efficace de normalisation

L'incertitude sur la section efficace du processus $t\bar{t}W$ (respectivement $t\bar{t}Z$) est de 13 % (respectivement 12 %). Elle est de 6 % sur les processus VV où V désigne W ou Z , et de $^{+6}_{-9}\%$ sur le processus $t\bar{t}H$ [16].

Une incertitude sur la section efficace globale de 50 % est appliquée sur tous les autres processus. Ces derniers étant rassemblés dans une même catégorie « other », leurs incertitudes sont traitées comme étant non corrélées et sont donc additionnées quadratiquement dans chaque région en considérant leur nombre d'événements relatif.

Lors de l'étude de signaux BSM, le processus de production de quatre quarks top conforme au modèle standard fait partie des bruits de fond et est traité séparément avec une incertitude de 50 % sur sa section efficace de production.

Incertitudes sur la forme des distributions

Les incertitudes sur la forme sont estimées séparément des incertitudes sur la section efficace détaillées auparavant. Plusieurs échantillons Monte-Carlo sont produits pour un même processus, en faisant varier différents paramètres. La section efficace calculée pour ces échantillons n'est pas exactement la même, il est donc nécessaire de les normaliser à la même section efficace, à savoir celle de l'échantillon standard. Ceci permet d'éviter un double comptage de l'incertitude sur la section efficace.

Les incertitudes sur la forme ne sont estimées que pour les bruits de fond irréductibles principaux, à savoir $t\bar{t}V$ et VV où V désigne W ou Z . Pour les autres bruits de fond, elles sont supposées négligeables ou prises en compte dans l'incertitude sur la section efficace.

Deux sources d'incertitude de forme sont étudiées au niveau reconstruit pour les processus $t\bar{t}W$ et $t\bar{t}Z$.

- D'abord on étudie les variations liées à la dépendance de la section efficace aux échelles de factorisation « μ_F » (voir section 2.4) et de renormalisation « μ_R » (voir section 1.5). Chaque échelle est tantôt fixée à sa valeur nominale, tantôt multipliée par deux (variation « $_\text{up}$ ») ou divisée par deux (variation « $_\text{down}$ »). Neuf combinaisons de valeurs sont donc considérées :
 - le cas nominal, pour lequel aucune échelle n'est modifiée ;
 - quatre cas pour lesquels une seule échelle est modifiée : μ_F_up , μ_F_down , μ_R_up , μ_R_down ;
 - quatre cas pour lesquels les deux échelles sont modifiées : $\mu_R_\text{up}_\mu_F_\text{up}$, $\mu_R_\text{up}_\mu_F_\text{down}$, $\mu_R_\text{down}_\mu_F_\text{up}$, $\mu_R_\text{down}_\mu_F_\text{down}$.

Dans chaque région de signal, les différentes variations sont comparées : la plus grande variation positive ainsi que la plus grande variation négative sont prises comme étant l'incertitude.

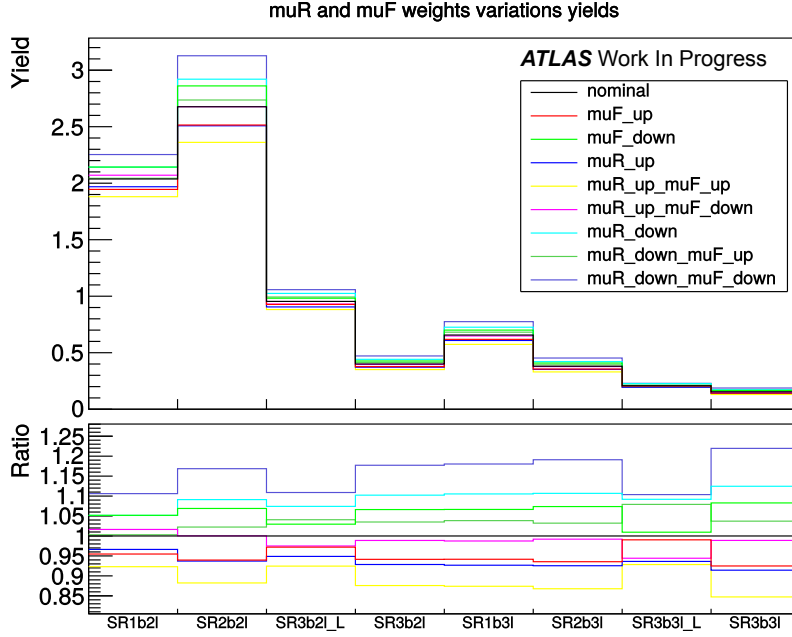


Figure 5.7 – Nombres d’événements $t\bar{t}W$ dans les régions de signal estimés avec différentes valeurs des échelles de factorisation et de renormalisation. Chaque couleur correspond à l’une des variations d’échelle (référence en noir). Le rectangle en bas de la figure présente le rapport du nombre d’événements observés pour la variation considérée par rapport à la valeur nominale.

- Une seconde incertitude étudiée provient du choix du générateur Monte-Carlo. La prédiction du générateur de référence MG5_AMC@NLO v2.2.2 est comparée à celle d’un générateur Monte-Carlo alternatif, à savoir SHERPA v2.2.1.

La figure 5.7 (respectivement figure 5.8) montre l’effet de la variation des échelles de factorisation et de renormalisation sur le processus $t\bar{t}W$ (respectivement $t\bar{t}Z$). Les rectangles en bas des figures montrent le rapport du nombre d’événements pour la variation considérée par rapport à la valeur nominale. On observe des variations maximales d’environ +10 % à +20 % pour le processus $t\bar{t}W$ (autour de +5 % pour le processus $t\bar{t}Z$); et des variations minimales d’environ –10 % à –15 % pour le processus $t\bar{t}W$ (environ –5 % à –10 % pour le processus $t\bar{t}Z$).

La comparaison des deux générateurs Monte-Carlo est illustrée sur la figure 5.9 (respectivement figure 5.10) pour le processus $t\bar{t}W$ (respectivement $t\bar{t}Z$). Les variations obtenues sont très différentes d’une région à l’autre avec un maximum d’environ +100 % pour le processus $t\bar{t}Z$ dans la région de signal SR3b2l. Ceci s’explique par la faible statistique de ce processus dans cette région, comme l’indiquent les barres d’erreur sur les figures.

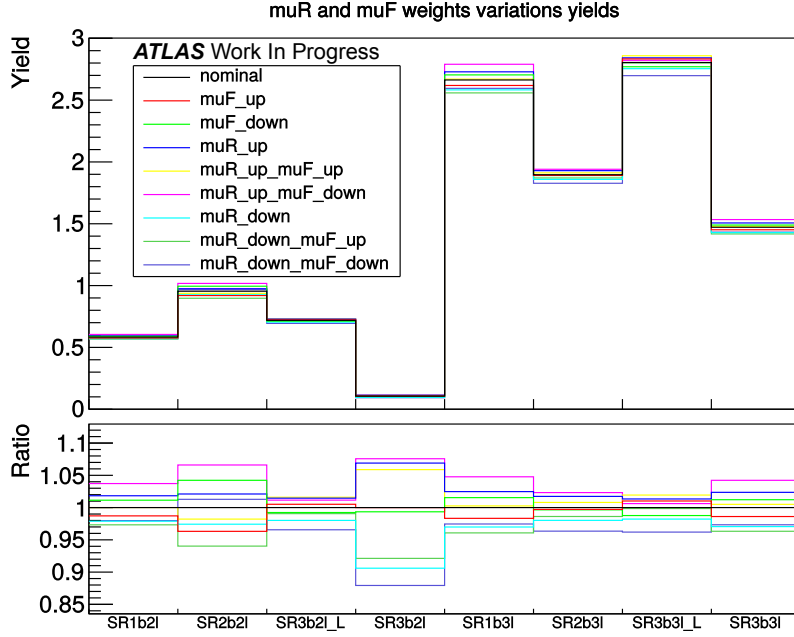


Figure 5.8 – Nombres d’événements $t\bar{t}Z$ dans les régions de signal estimés avec différentes valeurs des échelles de factorisation et de renormalisation. Chaque couleur correspond à l’une des variations d’échelle (référence en noir). Le rectangle en bas de la figure présente le rapport du nombre d’événements observés pour la variation considérée par rapport à la valeur nominale.

Les deux types d’incertitudes sont ajoutées quadratiquement pour chacun des deux processus et traitées comme un unique paramètre de nuisance par région de signal dans le modèle statistique. Les incertitudes ainsi obtenues sont présentées dans le tableau 5.6. La plupart de ces incertitudes sont de l’ordre de 10 ou 20 %, sauf dans le cas de régions à très faible statistique comme $SR3b2l$ pour le processus $t\bar{t}Z$. Dans ce cas de figure les incertitudes calculées dépassent 100 %, on choisit donc de fixer la variation négative à -99% afin d’éviter d’avoir un nombre d’événements négatif dans le modèle statistique.

Une méthode similaire est utilisée dans le but d’estimer les incertitudes de forme pour les processus VV . Toutefois une étude au niveau reconstruit comme dans le cas des processus $t\bar{t}V$ n’est pas possible par manque d’échantillons Monte-Carlo adéquats. C’est pourquoi l’étude est effectuée au niveau partonique. Différents types d’échelles dont la valeur théorique est incertaine sont variées de manière indépendante. L’impact de leur variation sur le nombre d’événements sélectionné est alors étudié. Les variations sont calculées de la manière suivante :

- Les échelles suivantes sont multipliées et divisées par deux :
- l’échelle de factorisation ;

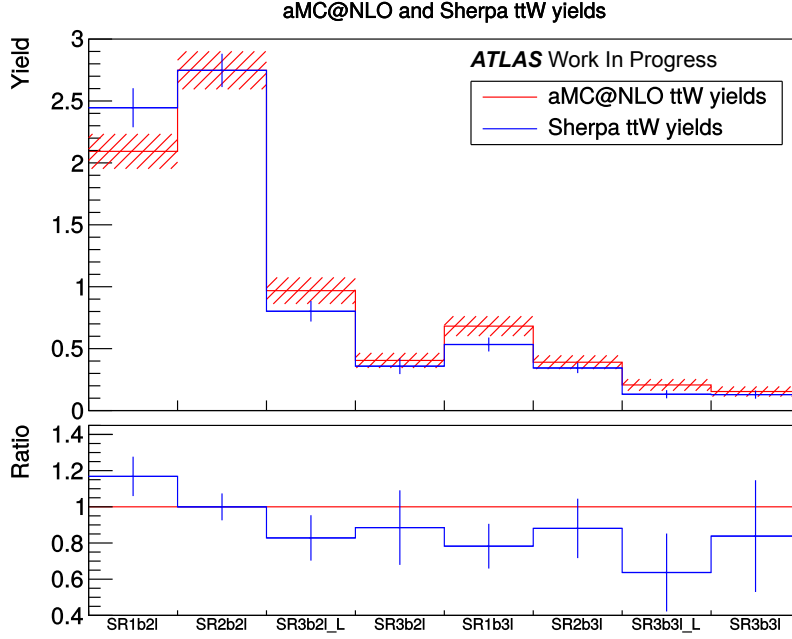


Figure 5.9 – Nombres d’événements $t\bar{t}W$ dans les régions de signal estimés via deux générateurs Monte-Carlo alternatifs. L’histogramme rouge correspond au générateur de référence et l’histogramme bleu à l’autre générateur. Le rectangle en bas de la figure présente le rapport du nombre d’événements observés avec chaque générateur par rapport à la référence.

- l’échelle de renormalisation ;
- l’échelle de resommation.
- L’échelle CKKW est placée à des valeurs de 15 GeV et 30 GeV, la valeur nominale étant 20 GeV.

Le nombre d’événements de certains échantillons utilisés pour cette étude étant assez faible (parfois seulement 10 000 événements), les incertitudes statistiques peuvent être très grandes. L’étude se restreint aux régions de signal avec un seul jet de quark b car le bruit de fond VV est négligeable dans les autres régions. Les variations observées ainsi que leurs incertitudes statistiques sont présentées dans le tableau 5.7.

Le tableau montre des valeurs très différentes selon la variation d’échelle considérée. De plus les incertitudes statistiques sont très élevées, souvent supérieures aux variations elles-mêmes. Les valeurs les plus significatives sont $-72 \pm 22\%$ pour la région $SR1b2\ell$ et $-60 \pm 20\%$ pour la région $SR1b3\ell$. Une incertitude conservatrice globale de 70 % est finalement appliquée dans toutes les régions de signal. Cette valeur a peu d’influence sur l’ajustement du modèle statistique de l’analyse sur les données puisque les régions avec un seul jet de quark b contiennent peu de signal et que le bruit de fond VV est

Région de signal	$t\bar{t}W$	$t\bar{t}Z$
SR1b2 ℓ	19,9 % −18,5 %	6,0 % −5,4 %
SR2b2 ℓ	16,9 % −11,8 %	8,7 % −8,2 %
SR3b2 ℓ _L	20,3 % −18,8 %	3,2 % −4,4 %
SR3b2 ℓ	21,1 % −16,9 %	130,4 % −99,0 %
SR1b3 ℓ	28,3 % −25,1 %	7,6 % −7,1 %
SR2b3 ℓ	22,5 % −17,8 %	16,6 % −16,8 %
SR3b3 ℓ _L	37,8 % −37,0 %	9,8 % −10,3 %
SR3b3 ℓ	27,3 % −22,3 %	5,9 % −5,5 %

Table 5.6 – Incertitudes systématiques globales appliquées aux processus $t\bar{t}W$ et $t\bar{t}Z$ dans les différentes régions de signal. Elles incluent les composantes de variation des échelles de factorisation et de renormalisation, ainsi que la modélisation du générateur Monte-Carlo.

Variation	SR1b2 ℓ	SR1b3 ℓ
Cas nominal	0 %	0 %
Factorisation $\times 2$	-49 ± 31 %	-28 ± 31 %
Factorisation $/2$	-46 ± 21 %	$+21 \pm 45$ %
Renormalisation $\times 2$	-8 ± 46 %	-38 ± 28 %
Renormalisation $/2$	-72 ± 22 %	$+40 \pm 51$ %
Resommation $\times 2$	-46 ± 30 %	-30 ± 30 %
Resommation $/2$	-48 ± 31 %	-6 ± 38 %
CKKW 15 GeV	-10 ± 44 %	-18 ± 32 %
CKKW 30 GeV	-25 ± 40 %	-60 ± 20 %

Table 5.7 – Variations des nombres d’événements VV par rapport à la valeur nominale, pour les différentes variations d’échelles étudiées. Les incertitudes sont statistiques.

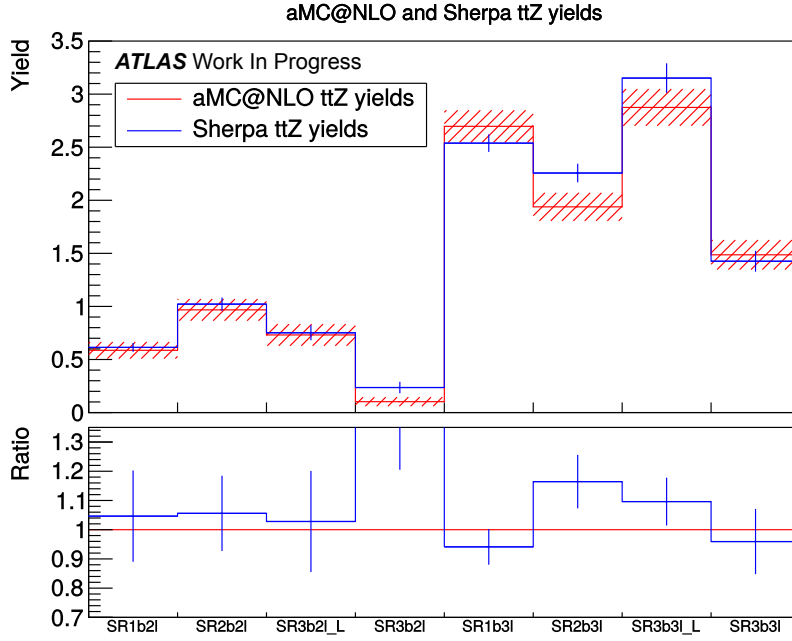


Figure 5.10 – Nombres d’événements $t\bar{t}Z$ dans les régions de signal estimés via deux générateurs Monte-Carlo alternatifs. L’histogramme rouge correspond au générateur de référence et l’histogramme bleu à l’autre générateur. Le rectangle en bas de la figure présente le rapport du nombre d’événements observés avec chaque générateur par rapport à la référence. La valeur du rapport pour $SR3b2l$ sort du cadre, elle vaut environ 2,3.

négligeable dans les autres régions.

Validité des incertitudes des bruits de fond sous-dominants

Comme expliqué auparavant, seuls les bruits de fond $t\bar{t}V$ et VV , où V désigne W ou Z , ont bénéficié d’études poussées en ce qui concerne les incertitudes systématiques. Des vérifications de la validité des incertitudes ont toutefois été effectuées sur les autres bruits de fond.

L’impact de l’incertitude de normalisation sur la section efficace du processus $t\bar{t}H$ a été étudiée puisque sa valeur par défaut est faible : $^{+6}_{-9}\%$, soit environ $\pm 10\%$. Ainsi les valeurs 10 %, 20 % et 30 % ont été testées. L’incertitude reste négligeable dans tous ces cas de figure, la valeur injectée pour cette incertitude n’a donc pas d’influence sur le résultat.

Les incertitudes systématiques appliquées sur bruits de fond de la catégorie « other » sont définies par région de signal. Une incertitude de 50 % est calculée pour chaque processus de cette catégorie dans chacune des régions de signal, puis ces valeurs sont

ajoutées quadratiquement par région de signal en respectant la composition du bruit de fond. Les incertitudes résultantes dans chaque région sont négligeables.

Lors de l'étude de signaux BSM l'incertitude systématique de 50 % sur la section efficace de production de quatre quarks top dans le cadre du modèle standard est présente. Elle n'est pas complètement négligeable, mais est nettement inférieure aux incertitudes dominantes : identification des jets de quarks b et bruit de fond de faux leptons.

5.5.4 Bruits de fond réductibles

Faux leptons et leptons non-prompts

Les incertitudes systématiques associées au bruit de fond dû aux faux leptons proviennent des taux de leptons lâches (vrais et faux) passant les coupures strictes, définis dans la section 5.4.2. Il existe trois sources d'incertitude :

- la définition des régions d'extraction des taux comme expliqué ci-dessous ;
- l'incertitude statistique dans les régions d'extraction des taux ;
- l'incertitude sur la normalisation des échantillons Monte-Carlo utilisés afin de soustraire les événements contenant des vrais leptons dans les régions d'extraction des taux des faux leptons f .

Les proportions de vrais et faux leptons dans les régions d'extraction des taux, ainsi que les importances relatives des sources de faux leptons (désintégration de jet de quark lourd, trident, conversion, etc.) dépendent des définitions de ces régions. Les incertitudes systématiques associées à ces effets sont estimées en modifiant les critères de sélection de ces régions et en comparant les taux qui en résultent. Les critères $E_T^{\text{miss}} > 150 \text{ GeV}$ de la région mesurant les valeurs de r des électrons est modifié en $E_T^{\text{miss}} > 175 \text{ GeV}$. Les critères $m_T(W) < 20 \text{ GeV}$ ainsi que $E_T^{\text{miss}} + m_T(W) < 60 \text{ GeV}$ de la région mesurant les valeurs de f des électrons sont remplacés par $E_T^{\text{miss}} < 20 \text{ GeV}$. Le critère $m_T(W) > 100 \text{ GeV}$ de la région mesurant les valeurs de r des muons devient $m_T(W) > 110 \text{ GeV}$. Enfin le critère $|d_0|/\sigma_{d_0} > 5$ de la région mesurant les valeurs de f des muons est remplacé par les critères $m_T(W) < 20 \text{ GeV}$ ainsi que $E_T^{\text{miss}} + m_T(W) < 60 \text{ GeV}$.

L'incertitude statistique est estimée en séparant les événements des régions d'extraction aléatoirement en quatre groupes, puis en estimant le bruit de fond dû aux faux leptons à partir des événements de chaque groupe. La moitié de la plus grande variation des nombres d'événements obtenus est prise comme incertitude. Enfin, l'incertitude sur la normalisation des échantillons Monte-Carlo est estimée en faisant varier cette normalisation de 10 % vers le haut et vers le bas.

Les incertitudes statistique et de définition des régions d'extraction des taux sont estimées indépendamment pour chaque région de signal dans les quatre régions d'extraction utilisées : électron r , électron f , muon r , et muon f . L'incertitude de normalisation Monte-Carlo est quant à elle unique pour chaque région de signal puisqu'elle affecte les

régions d'extraction des taux de manière corrélée. On compte ainsi neuf incertitudes systématiques par région de signal, traitées comme des paramètres de nuisance indépendants dans le modèle statistique.

L'incertitude totale sur le bruit de fond dû aux faux leptons se situe autour de 40 % ou 50 % selon la région de signal considérée. Les incertitudes provenant de la définition des régions d'extraction des taux dominent l'incertitude totale.

Mauvaise identification de la charge des électrons

Les taux de mauvaise identification de la charge des électrons présentent des incertitudes entraînant une incertitude systématique sur l'estimation du bruit de fond correspondant. Quatre sources sont considérées, et leurs valeurs typiques sont données ci-dessous :

- l'incertitude statistique provenant de la maximisation de la vraisemblance (15 %) ;
- les variations du résultat en considérant différentes définitions de la région du pic du Z comme indiqué ci-dessous (6 %) ;
- la différence entre le résultat de la méthode de maximisation de la vraisemblance et le nombre réel de mauvaises identifications dans un échantillon d'événements $Z \rightarrow e^+e^-$ simulés (5 %) ;
- les incertitudes sur le bruit de fond dû aux faux leptons (10 %).

Quatre définitions alternatives de la région du pic du Z sont considérées. Deux d'entre elles sont obtenues en agrandissant la région : la borne gauche ou droite de cette région est déplacée de 10 GeV, tout comme la bande latérale du côté affecté afin de conserver sa largeur. La troisième configuration est obtenue en augmentant la taille des bandes latérales à 20 GeV sans modifier la région sous le pic du Z , et la quatrième en supprimant les bandes latérales et donc en ignorant la soustraction de bruit de fond associée.

Ces incertitudes sont ajoutées quadratiquement pour une incertitude globale d'environ 20 %. On peut noter que la composante d'incertitude associée aux faux leptons est anti-corrélée avec l'incertitude du bruit de fond dû aux faux leptons.

5.5.5 Bilan

De par le faible nombre d'événements dans les régions de signal, l'incertitude statistique domine cette analyse. Les incertitudes systématiques dominantes concernent le bruit de fond dû aux faux leptons et la modélisation des bruits de fond irréductibles (section efficace et forme des distributions). Le tableau 5.8 donne la valeur des différentes incertitudes systématiques dans les régions de signal, exprimées en pourcentage du nombre total d'événements de bruit de fond. Il s'agit des incertitudes systématiques avant ajustement du modèle statistique.

Source d'incertitude	SR1b2 ℓ [%]	SR2b2 ℓ [%]	SR3b2 ℓ _L [%]	SR3b2 ℓ [%]	SR1b3 ℓ [%]	SR2b3 ℓ [%]	SR3b3 ℓ _L [%]	SR3b3 ℓ [%]
Résolution en énergie des jets	3	1	5	6	3	5	3	4
Échelle d'énergie des jets	3	3	9	6	3	5	11	6
b -tagging	5	3	6	7	3	4	9	9
Luminosité	1	1	2	2	2	2	2	2
Faux leptons	20	12	13	8	7	2	3	1
Charge mis-ID	2	3	1	2	-	-	-	-
Section efficace $\times$ forme	25	13	22	32	32	26	21	24

Table 5.8 – Incertitudes systématiques exprimées en pourcentage du nombre total d'événements de bruit de fond dans les régions de signal, par type d'incertitude. L'identification des jets de quarks b correspond à la ligne « b -tagging », et la mauvaise identification de la charge des électrons correspond à la ligne « Charge mis-ID ». Il s'agit des incertitudes systématiques avant ajustement du modèle statistique.

5.6 Analyse statistique

L'analyse présentée dans ce travail vise à évaluer la plausibilité des différents modèles BSM attendus à la lumière des données récoltées et de l'estimation des bruits de fond. Pour cela, une analyse statistique fondée sur un test d'hypothèses peut être menée de deux manières. Premièrement lorsque l'objectif est la découverte d'un nouveau processus physique, l'analyse doit déterminer si l'hypothèse de l'existence du processus est préférée à l'hypothèse de son inexistence. Ainsi l'hypothèse nulle H_0 correspond au cas où il n'y a que du bruit de fond tandis que l'alternative H_1 est l'hypothèse de signal plus bruit de fond. La configuration H_1 est notamment utilisée lorsqu'une mesure de section efficace est effectuée. Toutefois, en l'absence d'un excès significatif de données observées par rapport à l'estimation du bruit de fond, il est préférable de calculer une limite supérieure sur la section efficace du signal considéré plutôt que d'effectuer une mesure compatible avec une absence de signal. Dans ce cas de figure l'hypothèse nulle H_0 est l'hypothèse de signal plus bruit de fond et H_1 est l'hypothèse de bruit de fond seul. Au-delà de la limite obtenue, l'hypothèse de signal BSM est donc rejetée en faveur de l'hypothèse de bruit de fond seul, à un niveau de confiance choisi à 95 %. Les valeurs de section efficaces supérieures à cette limite sont dites exclues à un niveau de confiance de 95 %. Afin de quantifier le niveau d'accord entre une hypothèse H et les observations, une valeur- p peut être calculée. Il s'agit de la probabilité dans l'hypothèse H d'obtenir des données ayant une incompatibilité supérieure ou égale avec les prédictions de H . Une faible valeur- p

signifie alors que les données observées sont peu compatibles avec l'hypothèse H , puisqu'il n'existe qu'une faible probabilité d'obtenir des données encore moins compatibles. Le critère de niveau de confiance de 95 % peut alors se traduire en un critère sur la valeur- p : le critère de rejet est que la valeur- p est inférieure à 5 % : $p < 0,05$.

Dans le cas de l'analyse du signal de production de quatre quarks top conformément au modèle standard $t\bar{t}t\bar{t}$, ce processus est simplement supprimé de la liste de bruits de fond. Ainsi lorsque l'on veut mesurer $\sigma_{t\bar{t}t\bar{t}}$, on mène une analyse statistique de découverte du processus $pp \rightarrow t\bar{t}t\bar{t}$, donc H_0 est alors l'hypothèse de bruit de fond sans $t\bar{t}t\bar{t}$, et H_1 l'hypothèse de bruit de fond incluant $t\bar{t}t\bar{t}$.

Le modèle statistique utilisé est construit à partir de plusieurs termes de vraisemblance : les probabilités du nombre d'événements de données observées dans chaque région de signal vis-à-vis de leur estimation, ainsi qu'un terme modélisant les incertitudes systématiques introduites en tant que paramètres de nuisance. La vraisemblance est donc simplement la probabilité du nombre d'événements de données observé, vue comme une fonction de l'estimation du bruit de fond et des incertitudes. Seuls les nombres bruts d'événements après les coupures des régions de signal sont considérés, et aucune distribution cinématique n'est utilisée. Le nombre d'événements attendus de bruit de fond et de signal dans les régions de signal est estimé conformément aux méthodes présentées dans la section 5.4. La fonction de vraisemblance est définie ainsi :

$$\mathcal{L}(\mu, \boldsymbol{\theta} | \{n_i\}) = \prod_{i=1}^{n_{SR}} \mathcal{L}(n_i | \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})) \times \mathcal{L}(\boldsymbol{\theta}), \quad (5.7)$$

ou encore en explicitant le terme Poissonnien $\mathcal{L}$:

$$\mathcal{L}(\mu, \boldsymbol{\theta} | \{n_i\}) = \prod_{i=1}^{n_{SR}} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)} \times \mathcal{L}(\boldsymbol{\theta}), \quad (5.8)$$

où

- μ est appelé « force du signal », il s'agit du facteur de normalisation de la section efficace théorique du signal et donc du nombre d'événements attendus pour le signal, le cas de figure $\mu = 0$ (respectivement $\mu = 1$) désigne par exemple une absence de signal (respectivement un signal conforme à la théorie) ;
- $\boldsymbol{\theta}$ est le vecteur contenant les paramètres de nuisance ;
- i désigne le numéro de la région de signal ;
- n_{SR} est le nombre de régions de signal ;
- $s_i(\boldsymbol{\theta})$ et $b_i(\boldsymbol{\theta})$ sont le nombre d'événements attendus de signal et de bruit de fond dans la région de signal i , dépendant de $\boldsymbol{\theta}$;
- n_i est le nombre d'événements de données observé dans la région i , et $\{n_i\}$ désigne l'ensemble des n_i ;
- $\mathcal{L}(\boldsymbol{\theta})$ est le produit des termes de vraisemblance des différents paramètres de nuisance.

Les incertitudes systématiques sont encodées dans le terme $\boldsymbol{\theta}$. Elles représentent la méconnaissance de grandeurs physiques influençant le résultat de l'analyse. Le nombre d'événements de bruit de fond et de signal dépend notamment de ces paramètres de nuisance. Par exemple l'échelle d'énergie des jets a un impact sur l'impulsion transverse des jets reconstruits et donc sur des variables comme H_T ou E_T^{miss} qui servent de critère d'entrée dans les régions de signal. Une incertitude systématique est définie a priori comme une distribution gaussienne ou log-normale de la grandeur impactée, de valeur centrale ξ_m et de largeur σ_ξ . Une distribution log-normale est préférée à une distribution gaussienne pour un paramètre de nuisance positif. Ces informations a priori proviennent souvent d'une analyse complémentaire effectuée dans une ou plusieurs régions de contrôle, elles ne peuvent donc pas représenter parfaitement la réalité des régions de signal de l'analyse. Lors de l'ajustement du modèle statistique, la grandeur ξ a la possibilité de s'écarter de la valeur a priori ξ_m , mais avec une pénalité gaussienne ou log-normale, tandis que la largeur σ_ξ peut augmenter ou diminuer selon le niveau de contrainte apporté par l'ajustement. La composante associée dans le vecteur $\boldsymbol{\theta}$ est alors $(\xi - \xi_m)/\sigma_\xi$, soit la distance entre la valeur considérée et la valeur de référence en unités de déviation standard de la mesure.

Afin de tester une valeur donnée du paramètre μ , on définit un rapport de vraisemblance

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu) | \{n_i\})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}} | \{n_i\})}. \quad (5.9)$$

Les termes $\hat{\mu}$ et $\hat{\boldsymbol{\theta}}$ représentent le couple de valeurs $(\mu, \boldsymbol{\theta})$ maximisant la vraisemblance vis-à-vis des données $\{n_i\}$. Le terme $\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}} | \{n_i\})$ représente donc la vraisemblance maximale. La notation $\hat{\boldsymbol{\theta}}(\mu)$ représente la valeur de $\boldsymbol{\theta}$ maximisant $\mathcal{L}$ pour la valeur de μ considérée. Cette grandeur est une fonction de μ , c'est-à-dire que les valeurs des paramètres de nuisance maximisant la vraisemblance dépendent de la valeur de μ choisie. Le terme $\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu) | \{n_i\})$ représente donc la vraisemblance maximale pour une valeur de μ donnée. Par construction, $\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu) | \{n_i\}) \leq \mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}} | \{n_i\})$, avec égalité si $\mu = \hat{\mu}$. On a donc $\lambda(\mu) \in [0, 1]$, et une certaine valeur de μ est d'autant plus compatible avec les données observées que λ est proche de 1. Le rapport de vraisemblance λ est une fonction de μ qui ne dépend plus de $\boldsymbol{\theta}$. En effet, la sélection du maximum de vraisemblance de $\boldsymbol{\theta}$ pour une valeur de μ donnée fournit une fonction de μ : $\hat{\boldsymbol{\theta}}(\mu)$. Une statistique de test est alors définie à partir de λ afin de réaliser le test d'hypothèses décrit précédemment :

$$t_\mu = -2 \ln \lambda(\mu). \quad (5.10)$$

On peut remarquer que $t_\mu \in [0, +\infty[$, et que l'accord d'une valeur μ avec les observations est d'autant plus grand que t_μ est proche de 0. Inversement plus la valeur de t_μ est élevée, plus le désaccord est grand. Il est possible de quantifier ce désaccord pour une valeur

donnée de μ via le calcul d'une valeur- p

$$p_\mu = \int_{t_{\mu,\text{obs}}}^{+\infty} f(t_\mu|\mu) dt_\mu, \quad (5.11)$$

avec $t_{\mu,\text{obs}}$ la valeur de la statistique de test pour la valeur de μ considérée et les données $\{n_i\}$. Si les données réelles $\{n_i\}$ sont remplacées par d'autres valeurs, alors les paramètres $\hat{\mu}$ et $\hat{\theta}$ sont modifiés, et par conséquent la valeur de t_μ change. Le terme $f(t_\mu|\mu)$ représente la distribution des valeurs de t_μ obtenues pour des nombres d'événements distribués selon l'hypothèse d'une force de signal égale à μ . La valeur- p p_μ représente donc la probabilité, étant donné une force de signal égale à μ , d'obtenir des nombres d'événements donnant une valeur de t_μ supérieure à la valeur $t_{\mu,\text{obs}}$ calculée à partir des données réelles $\{n_i\}$. Le rejet des valeurs de μ produisant une valeur- p inférieure à un certain seuil permet d'obtenir un intervalle de confiance.

L'analyse présentée dans ce chapitre n'utilise cependant pas directement un critère sur cette valeur- p , mais construit une grandeur R_{CLs} comme discriminant pour calculer les limites supérieures, similairement à la méthode décrite dans la référence [22] :

$$R_{\text{CLs}}(\mu) = \frac{p_\mu}{p_0}, \quad (5.12)$$

où p_0 est la valeur- p dans le cas du test de l'hypothèse $\mu = 0$. Ce rapport de valeurs- p permet de comparer l'hypothèse d'un signal de force μ à celle d'une absence de signal ($\mu = 0$). Le critère pour rejeter l'hypothèse de signal de force μ est alors

$$R_{\text{CLs}}(\mu) < 0.05 \quad (5.13)$$

pour un niveau de confiance de 95 %. Si les données sont compatibles avec le bruit de fond seul, la grandeur $R_{\text{CLs}}(\mu)$ tend à diminuer avec l'augmentation de μ . Toutes les valeurs de μ pour lesquelles $R_{\text{CLs}}(\mu) < 0,05$ sont considérées comme exclues à un niveau de confiance de 95 %. Si dans le cas d'un modèle particulier $R_{\text{CLs}}(\mu = 1) < 0,05$, alors ce modèle est exclus.

Afin de calculer les valeurs- p nécessaires au critère de rejet d'hypothèse, il est essentiel de connaître la densité de probabilité $f(t_\mu|\mu)$. Cette densité de probabilité est calculée dans l'approximation asymptotique comme décrit dans la référence [23].

5.7 Résultats

Cette section présente le nombre d'événements dans les régions de signal de l'analyse, ainsi que l'ajustement du modèle statistique sur ces nombres. L'ajustement est réalisé dans le cadre de différents modèles, et est interprété soit sous forme de limite, soit comme une mesure de la section efficace du processus étudié. Enfin la combinaison des résultats de l'analyse présentée dans ce travail avec ceux de l'analyse étudiant les événements à un seul lepton est décrite.

5.7.1 Nombre d'événements dans les régions de signal

Le tableau 5.9 et la figure 5.11 présentent le nombre d'événements estimé pour les différents bruits de fond dans les huit régions de signal, ainsi que les nombres d'événements observés dans les données. Seuls quelques événements, une dizaine au maximum, sont observés dans les régions de signal étant donné les coupures strictes définissant les régions.

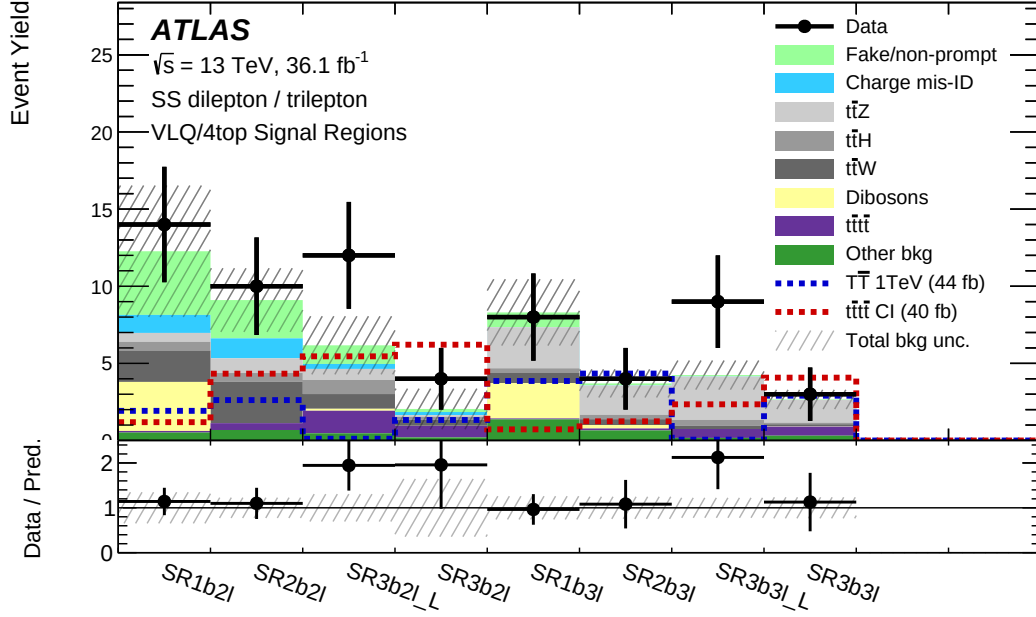


Figure 5.11 – Nombre d'événements de bruit de fond et de données dans les huit régions de signal. L'histogramme des bruits de fond est divisé en catégories représentées par des couleurs différentes. L'incertitude donnée par la bande hachurée inclut les incertitudes statistique et systématique sur l'ensemble des bruits de fond. Les données sont représentées par les points noirs avec leur incertitude. Le rectangle inférieur représente le rapport du nombre d'événements observé dans les données sur le nombre d'événements prédits de bruit de fond pour chaque région de signal.

Les valeurs de signification statistique indiquées pour chaque région de signal dans le tableau caractérisent l'excès du nombre d'événements observés vis-à-vis de l'estimation du bruit de fond. Elles sont exprimées en nombre de déviations standard gaussiennes. Les valeurs négatives pour la région SR1b3 ℓ proviennent d'un déficit de données. La valeur appelée « BSM signif. » correspond à cet excès de données comparé à l'ensemble des sources de bruit de fond. La valeur « $t\bar{t}t\bar{t}$ MS signif. » quant à elle est l'excès de données comparé aux bruits de fond privés du processus $t\bar{t}t\bar{t}$ conforme au modèle standard. C'est pourquoi la seconde valeur est toujours plus élevée que la première. Les nombres de

Source	SR1b2 ℓ	SR2b2 ℓ	SR3b2 ℓ _L	SR3b2 ℓ
$t\bar{t}W$	$2,04 \pm 0,14 \pm 0,49$	$2,68 \pm 0,15 \pm 0,55$	$0,95 \pm 0,11 \pm 0,31$	$0,40 \pm 0,06 \pm 0,10$
$t\bar{t}Z$	$0,58 \pm 0,08 \pm 0,10$	$0,95 \pm 0,11 \pm 0,17$	$0,72 \pm 0,11 \pm 0,19$	$0,11 \pm 0,05 \pm 0,13$
Dibosons	$3,2 \pm 1,5 \pm 2,4$	$< 0,5$	$0,13 \pm 0,13 \pm 0,27$	$< 0,5$
$t\bar{t}H$	$0,56 \pm 0,07 \pm 0,07$	$0,57 \pm 0,10 \pm 0,09$	$0,91 \pm 0,11 \pm 0,22$	$0,19 \pm 0,05 \pm 0,07$
$t\bar{t}t\bar{t}$	$0,10 \pm 0,01 \pm 0,05$	$0,44 \pm 0,03 \pm 0,23$	$1,46 \pm 0,05 \pm 0,74$	$0,75 \pm 0,04 \pm 0,38$
Other bkg	$0,52 \pm 0,07 \pm 0,14$	$0,68 \pm 0,09 \pm 0,24$	$0,47 \pm 0,08 \pm 0,18$	$0,20 \pm 0,04 \pm 0,06$
Faux leptons	$4,1 \pm 1,6 \pm 2,4$	$2,5 \pm 1,0 \pm 1,1$	$1,2 \pm 0,9 \pm 0,6$	$0,20 \pm 0,46 \pm 0,16$
Charge mis-ID	$1,17 \pm 0,10 \pm 0,27$	$1,29 \pm 0,10 \pm 0,28$	$0,32 \pm 0,04 \pm 0,09$	$0,21 \pm 0,04 \pm 0,04$
Total bkg	$12,3 \pm 2,2 \pm 3,4$	$9,1 \pm 1,2 \pm 1,2$	$6,2 \pm 1,0 \pm 1,2$	$2,0 \pm 0,5 \pm 0,3$
Data yield	14	10	12	4
BSM signif.	0,31	0,25	1,7	1,1
$t\bar{t}t\bar{t}$ MS signif.	0,33	0,38	2,1	1,6

Source	SR1b3 ℓ	SR2b3 ℓ	SR3b3 ℓ _L	SR3b3 ℓ
$t\bar{t}W$	$0,66 \pm 0,08 \pm 0,20$	$0,38 \pm 0,05 \pm 0,11$	$0,21 \pm 0,05 \pm 0,09$	$0,15 \pm 0,04 \pm 0,05$
$t\bar{t}Z$	$2,66 \pm 0,15 \pm 0,43$	$1,90 \pm 0,14 \pm 0,42$	$2,80 \pm 0,17 \pm 0,58$	$1,47 \pm 0,14 \pm 0,28$
Dibosons	$2,3 \pm 0,7 \pm 1,7$	$0,22 \pm 0,16 \pm 0,27$	$< 0,5$	$< 0,5$
$t\bar{t}H$	$0,30 \pm 0,04 \pm 0,04$	$0,28 \pm 0,05 \pm 0,05$	$0,38 \pm 0,06 \pm 0,07$	$0,10 \pm 0,03 \pm 0,02$
$t\bar{t}t\bar{t}$	$0,06 \pm 0,01 \pm 0,03$	$0,13 \pm 0,02 \pm 0,06$	$0,58 \pm 0,04 \pm 0,29$	$0,59 \pm 0,03 \pm 0,30$
Other bkg.	$1,37 \pm 0,13 \pm 0,45$	$0,65 \pm 0,10 \pm 0,27$	$0,17 \pm 0,09 \pm 0,10$	$0,31 \pm 0,07 \pm 0,11$
Faux leptons	$1,0 \pm 0,6 \pm 0,6$	$0,14 \pm 0,31 \pm 0,09$	$0,00 \pm 0,38 \pm 0,09$	$0,03 \pm 0,15 \pm 0,00$
Total bkg	$8,3 \pm 0,9 \pm 1,8$	$3,7 \pm 0,6 \pm 0,4$	$4,2 \pm 0,4 \pm 0,7$	$2,7 \pm 0,2 \pm 0,5$
Data yield	8	4	9	3
BSM signif.	-0,09	0,14	1,8	0,19
$t\bar{t}t\bar{t}$ MS signif.	-0,07	0,21	2,1	0,6

Table 5.9 – Nombre d'événements de bruit de fond et de données dans les huit régions de signal. La première incertitude est statistique et la seconde systématique. Les deux dernières lignes montrent la signification statistique de l'excès de données comparé à la somme des bruits de fond. Dans le cas « BSM », tous les bruits de fond sont pris en compte tandis que dans le cas « $t\bar{t}t\bar{t}$ MS », le processus $t\bar{t}t\bar{t}$ n'est pas considéré comme un bruit de fond. Les valeurs de signification statistique sont exprimées en nombre de déviations standard gaussiennes, et les valeurs négatives dans la région SR1b3 ℓ proviennent d'un déficit de données.

signification « BSM signif. » ne dépendent que du nombre estimé d'événements de bruit de fond et du nombre observé d'événements de données, et sont donc indépendants du modèle BSM considéré, tandis que les nombres « $t\bar{t}t\bar{t}$ MS signif. » correspondent aux significations statistiques du signal $t\bar{t}t\bar{t}$.

Les régions SR3b2 ℓ _L et SR3b3 ℓ _L, définies aux valeurs intermédiaires de H_T , présentent un excès d'événements de données représentant respectivement 1,7 et 1,8 déviations standard de la prédiction de bruit de fond. Ces valeurs de signification s'élèvent à 2,1 pour le signal $t\bar{t}t\bar{t}$, en supprimant ce même processus de la liste de bruits de fond. L'excès de données atteint 3,0 déviations standard pour le signal $t\bar{t}t\bar{t}$ en combinant toutes les régions de signal. Pour comparaison, la signification attendue calculée à partir de la simulation du processus $t\bar{t}t\bar{t}$ et non des données observées est de seulement 0,9 déviations standard. Les propriétés cinématiques des événements entrant dans la région SR3b2 ℓ _L ou dans la région SR3b3 ℓ _L (respectivement n'importe quelle région de signal) sont représentées sur la figure 5.12 (respectivement figure 5.13). La cinématique de l'excès de données a été comparée aux prédictions des différents modèles étudiés mais les observations ne sont en accord avec aucun de ces modèles, en particulier en ce qui concerne le nombre de jets et le nombre de jets de quarks b .

5.7.2 Ajustement du modèle

Le modèle statistique décrit dans la section 5.6 est ajusté sur les données observées ou bien sur des pseudo-données générées à partir des estimations de bruit de fond et de signal. Cet ajustement est réalisé séparément pour chaque hypothèse de signal prise en considération. Une limite est également calculée pour chacun de ces signaux. Lors d'une mesure l'hypothèse de bruit de fond plus signal est testée face à l'hypothèse nulle de bruit de fond seul, tandis que lors de l'établissement d'une limite c'est l'hypothèse de bruit de fond seul qui est testée face à l'hypothèse nulle de bruit de fond plus signal. Ainsi, les pseudo-données sont générées dans l'hypothèse $\mu = 1$ pour une mesure et dans l'hypothèse $\mu = 0$ pour l'établissement de limites. Un résultat obtenu via l'ajustement du modèle statistique sur les données réelles (respectivement pseudo-données) est dit « observé » (respectivement « attendu »).

Comme il a été décrit plus haut, les paramètres de nuisance du modèle statistique sont ajustés afin de maximiser la vraisemblance vis-à-vis des données observées (ou le cas échéant des pseudo-données utilisées). Il est donc important de vérifier l'ajustement final des paramètres de nuisance afin de détecter toute irrégularité. De plus, puisque l'analyse est limitée par la statistique disponible, aucune contrainte n'est attendue sur les paramètres de nuisance. La figure 5.14 présente les paramètres de nuisance les plus importants dans le cadre de l'analyse du signal de production de quatre quarks top conforme au modèle standard. À chacun de ces paramètres est associé un point et une barre noirs, représentant respectivement le déplacement après ajustement par rapport

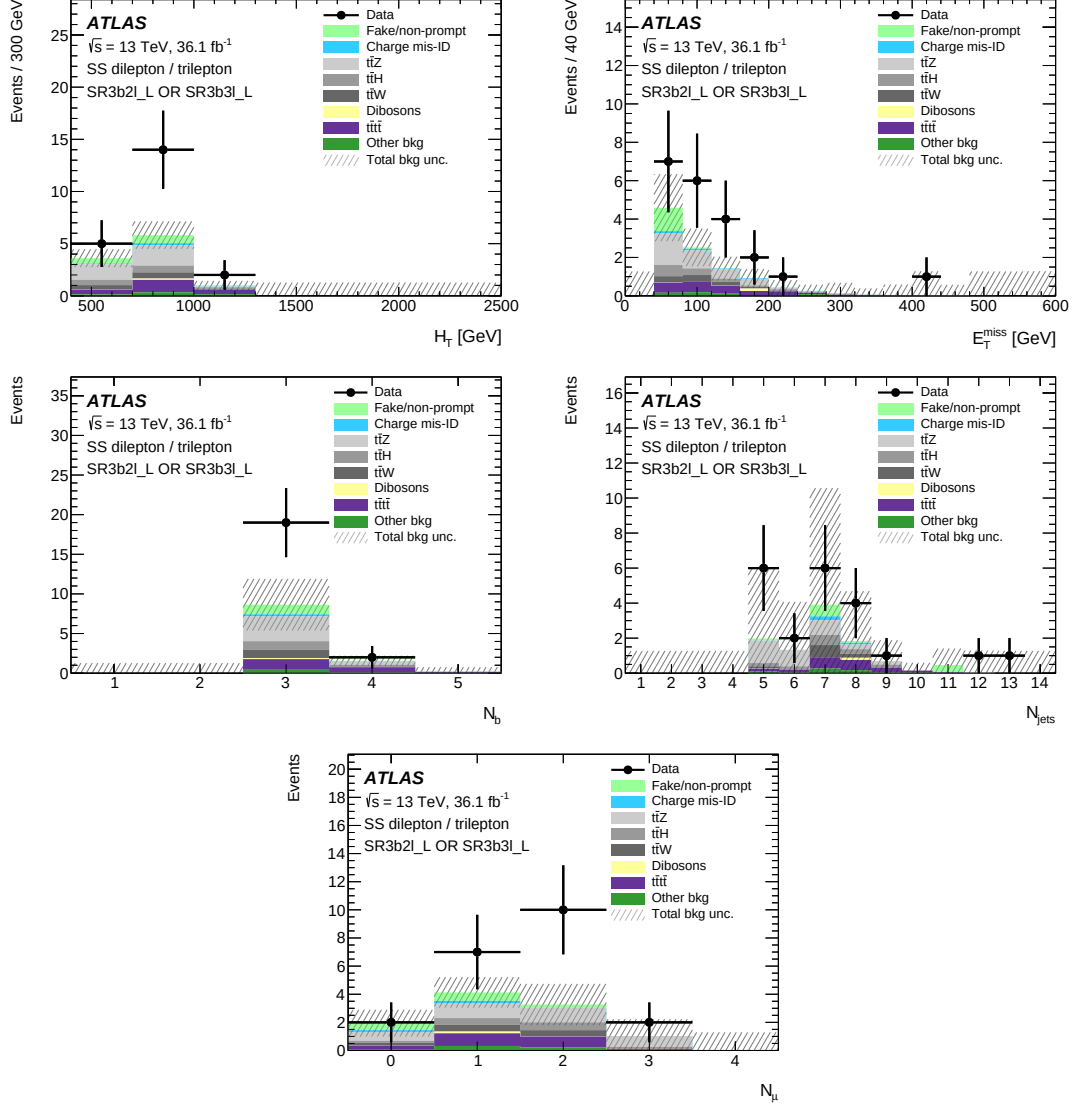


Figure 5.12 – Distributions de variables cinématiques des événements entrant dans la région de signal SR3b2 ℓ _L ou dans la région de signal SR3b3 ℓ _L. De gauche à droite puis de haut en bas les variables sont H_T , E_T^{miss} , le nombre de jets de quarks b N_b , le nombre de jets N_{jets} , et enfin le nombre de muons N_{μ} .

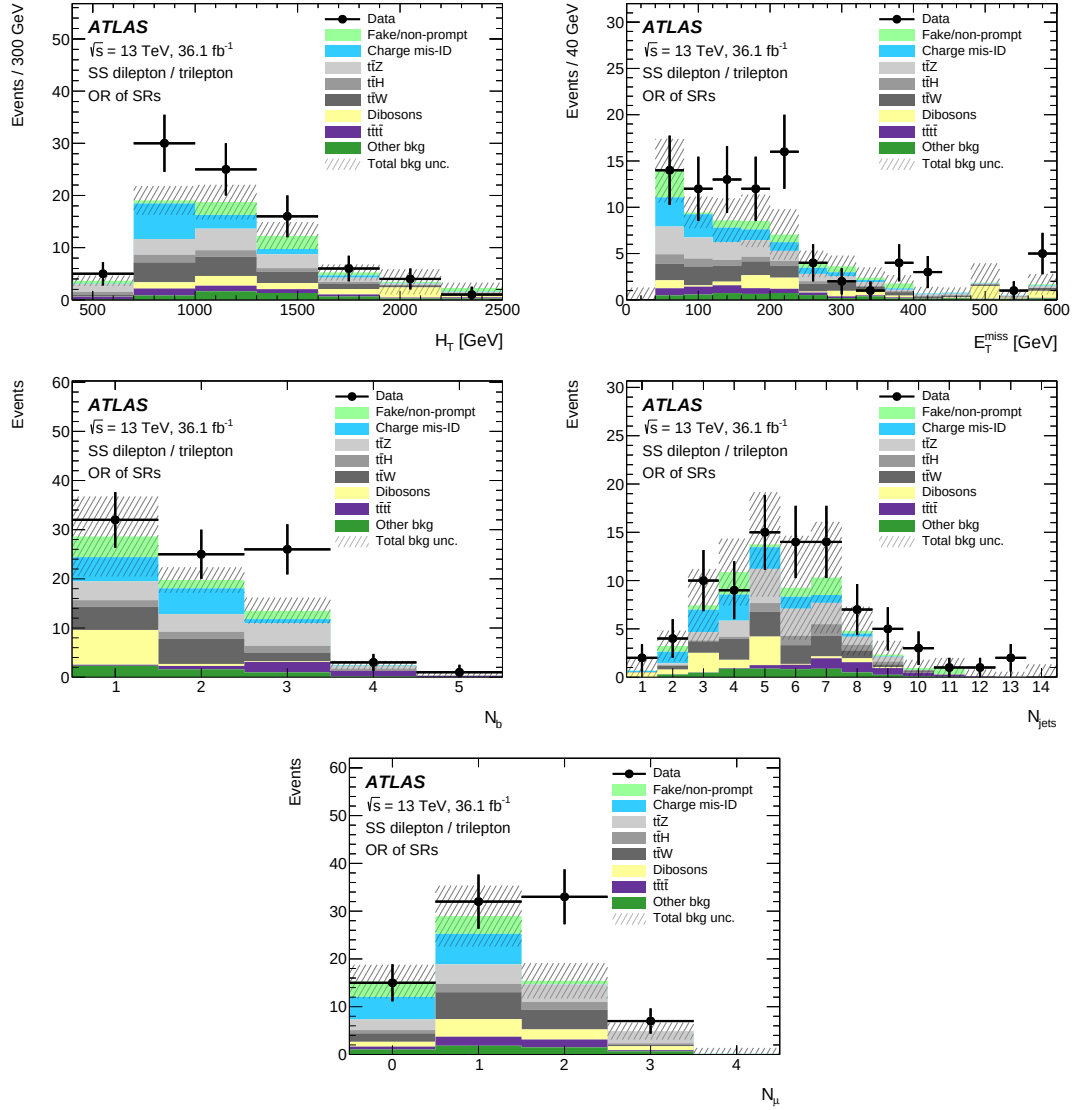


Figure 5.13 – Distributions de variables cinématiques des événements entrant dans n'importe quelle région de signal. De gauche à droite puis de haut en bas les variables sont H_T , E_T^{miss} , le nombre de jets de quarks b N_b , le nombre de jets N_{jets} , et enfin le nombre de muons N_{μ} .

Modèle	$\hat{\mu}$
MS	$4,4 \pm 1,7$
CI	$0,019 \pm 0,011$
2HDM (700 GeV)	$0,28 \pm 0,13$
2UED (1400 GeV)	$0,17 \pm 0,25$

Table 5.10 – Valeurs ajustées de force du signal pour plusieurs modèles représentatifs.

à la valeur a priori avant ajustement $(\xi - \xi_m)/\sigma_\xi$ (noté $(\hat{\theta} - \theta_0)/\Delta\theta$ sur la figure) et la largeur σ_ξ après ajustement. Les points noirs sont tous proches de la valeur 0 donc les valeurs après ajustement des différents paramètres de nuisance sont proches des valeurs avant ajustement. De plus, les barres noires s'étendent toutes environ entre les valeurs -1 et $+1$ ce qui signifie que l'ajustement ne contraint que faiblement les paramètres de nuisance, ce qui est conforme aux attentes.

L'effet des principaux paramètres de nuisance sur le signal est représenté sur la figure 5.14 à l'aide des rectangles bleus associés à chaque paramètre de nuisance. Ces rectangles représentent l'impact $\Delta\mu$ d'une variation du paramètre de nuisance sur la valeur ajustée de la force du signal, dans le cas où le paramètre est diminué d'une fois sa largeur ($-\Delta\theta$ ou $-\Delta\hat{\theta}$) ou augmenté d'une fois sa largeur ($+\Delta\theta$ ou $+\Delta\hat{\theta}$). La largeur prise en compte est celle avant ajustement ($\Delta\theta$) pour les rectangles creux et celle après ajustement ($\Delta\hat{\theta}$) pour les rectangles pleins. Les paramètres de nuisance les plus importants sont associés aux incertitudes sur l'identification des jets de quarks b (préfixe « FT_EFF_Eigen_ ») et sur les jets en général (préfixe « JET_ »), ainsi que les incertitudes sur le bruit de fond de faux leptons (noms contenant le terme « Fakes » ou avec le suffixe « _Stat »). L'incertitude systématique dominante représente environ 20 % de l'incertitude totale sur μ . La valeur ajustée de la force du signal est donnée dans le tableau 5.10. Les valeurs ajustées de force du signal sont également données pour quelques modèles BSM étudiés.

5.7.3 Résultats sur la production de quatre quarks top selon le modèle standard

Les limites supérieures calculées sur la production de quatre quarks top selon le MS sont présentées dans le tableau 5.11. Les limites observée et attendue sont respectivement 69 fb et 29 fb. La prédiction théorique utilisée est d'environ 9,2 fb, donc les limites observée et attendue correspondent respectivement à 7,5 et 3,2 fois la prédiction théorique. La limite observée est nettement supérieure à la limite attendue étant donné l'excès de données mentionné auparavant. En effet un nombre d'événements plus élevé que prévu

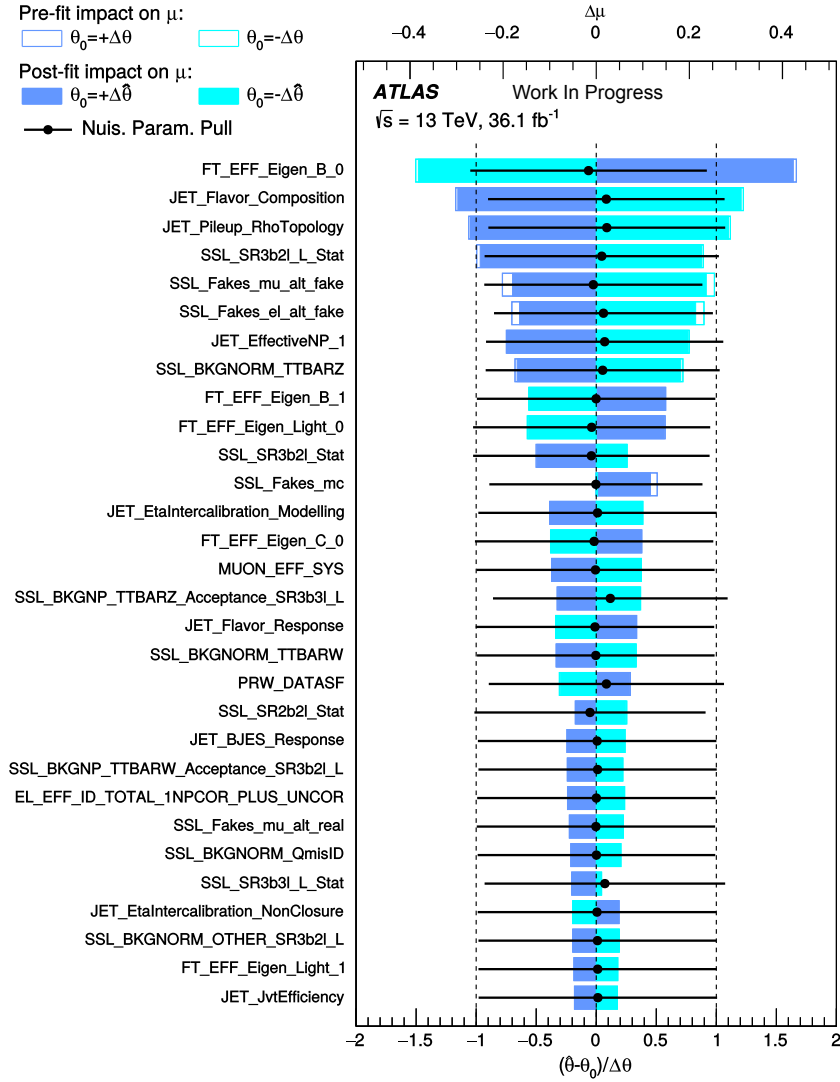


Figure 5.14 – Effet des paramètres de nuisance les plus influents sur la force du signal de production de quatre quarks tops dans le cadre du modèle standard. Le point noir et la barre noire associés à un paramètre de nuisance représentent respectivement le déplacement de sa valeur par rapport à sa valeur a priori et la contrainte apportée par l’ajustement. Les rectangles bleus représentent l’impact $\Delta\mu$ d’une variation du paramètre de nuisance sur la valeur ajustée de la force du signal, dans le cas où le paramètre est diminué d’une fois sa largeur ($-\Delta\theta$ ou $-\Delta\hat{\theta}$) ou augmenté d’une fois sa largeur ($+\Delta\theta$ ou $+\Delta\hat{\theta}$). La largeur prise en compte est celle avant ajustement ($\Delta\theta$) pour les rectangles creux et celle après ajustement ($\Delta\hat{\theta}$) pour les rectangles pleins.

Observable	Médiane attendue avec intervalle à 1σ	Observée
section efficace MS [fb]	$29.0^{+12.2}_{-8.1}$	69.2
section efficace CI [fb]	$20.8^{+12.2}_{-8.1}$	38.6
couplage CI $ C_{4t} /\Lambda^2$ [TeV $^{-2}$]	$1.9^{+1.2}_{-0.7}$	2.6

Table 5.11 – Limites supérieures attendues et observées sur la production de quatre quarks top dans le modèle standard MS et le modèle à interaction de contact CI.

rend plus plausible des valeurs élevées de section efficace, entraînant une limite plus élevée et donc moins contraignante sur la section efficace.

5.7.4 Résultats sur les modèles de nouvelle physique

Des limites sont également calculées pour les modèles de nouvelle physique décrits dans la section 4.2. Ces limites sont représentées sur différentes figures dans cette sous-section respectant le même code couleur. La ligne pointillée ainsi que la bande colorée verte (jaune) représentent la limite attendue avec sa bande d'incertitude à $\pm 1\sigma$ ($\pm 2\sigma$). La courbe noire en trait plein représente la limite observée. La zone hachurée en rouge est exclue par l'analyse à un niveau de confiance de 95 %. Les modèles 2UED et 2HDM étant générés pour certaines valeurs de masse uniquement, les lignes noires sont alors déterminées par interpolation entre les résultats obtenus pour chacune de ces valeurs.

Dans le cas du modèle CI (voir section 4.2.1), les limites observée et attendue sont respectivement 39 fb et 21 fb, comme présenté dans le tableau 5.11. Ce tableau montre également la limite observée (attendue) sur le paramètre central du modèle, à savoir le rapport $|C_{4t}|/\Lambda^2$, qui est d'environ 2,6 TeV $^{-2}$ (1,9 TeV $^{-2}$). Le modèle CI est également étudié dans le plan $(\Lambda, |C_{4t}|)$, comme présenté dans la figure 5.15. Les grandeurs Λ et C_{4t} correspondent à l'échelle d'énergie du phénomène et à la valeur de couplage du modèle, respectivement. La région de l'espace des phases exclue (zone hachurée) correspond à une échelle d'énergie basse et à un fort couplage.

Une limite sur la valeur du paramètre m_{KK} est calculée dans le cadre du modèle 2UED (voir section 4.2.2). Il est question ici de la configuration par défaut, c'est-à-dire $\xi = 1$. Ce résultat est représentée sur la figure 5.16 (gauche). Il est ainsi possible d'exclure les valeurs de masse m_{KK} inférieures à 1,45 TeV. La limite attendue est de 1,48 TeV. Une étude est également effectuée dans le plan (m_{KK}, ξ) , et présentée sur la figure 5.16 (droite). Le bas de la figure de droite correspond à la valeur $\xi = 1$ et donc au résultat de la figure de gauche.

Enfin, les limites sur la production de quatre quarks top dans le modèle 2HDM (voir section 4.2.3) sont présentées sur la figure 5.17. Deux scénarios sont étudiés : l'un pour

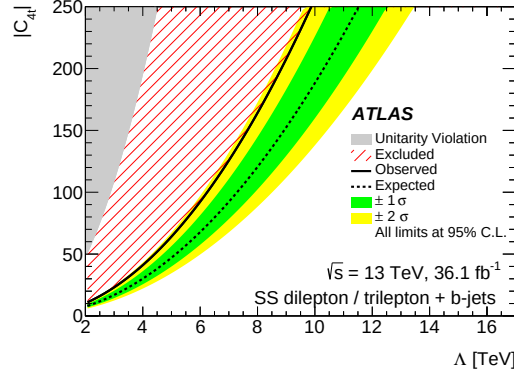


Figure 5.15 – Contraintes sur le modèle à interaction de contact CI dans le plan $(\Lambda, |C_{4t}|)$. La zone grisée est exclue par la théorie et la zone hachurée est exclue par l’analyse présentée dans ce chapitre.

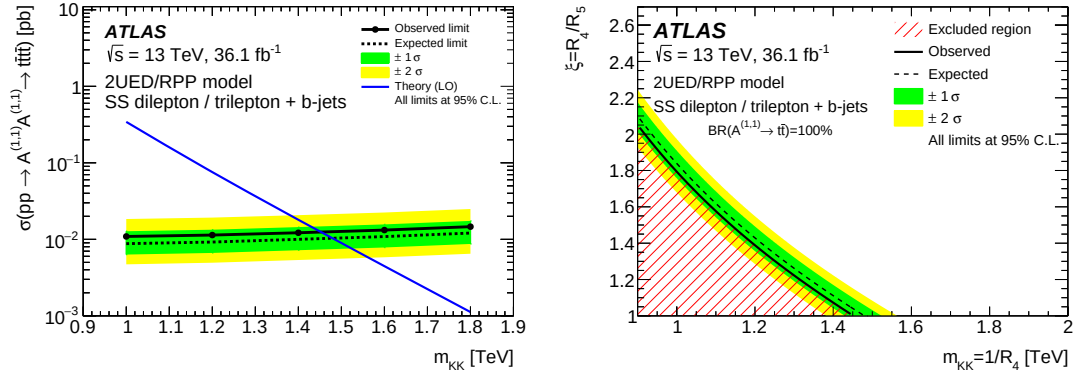


Figure 5.16 – Limite attendue et observée sur la section efficace (gauche) et dans le plan (m_{KK}, ξ) (droite), dans le cadre du modèle à deux dimensions supplémentaires compactifiées 2UED. La droite bleue représente la section efficace théorique de production de quatre quarks top via ce modèle.

lequel seul le boson lourd H contribue au processus de production de quatre quarks top (figure 5.17 en haut à gauche), l’autre où les deux bosons H et A contribuent (figure 5.17 en haut à droite). Dans les deux cas, les limites obtenues sont présentées dans le plan $(m_H, \tan \beta)$. La limite sur la section efficace du processus 2HDM est montrée sur la figure 5.17 (bas). Les prédictions théoriques pour trois valeurs du paramètre $\tan \beta$ sont présentées. Pour les valeurs les plus élevées du paramètre $\tan \beta$, il n’est pas possible de calculer une limite sur la masse m_H . La figure 5.17 (bas) montre en effet le cas $\tan \beta = 1,0$ pour lequel l’intersection entre la droite bleue de prédiction théorique et la droite noire d’exclusion expérimentale semble se situer en dessous de 400 GeV. Les simulations du

signal 2UED ne sont effectuées qu'à partir de cette valeur, c'est pourquoi la limite sur la masse m_H n'est pas calculable. Dans le cas de figure $\tan\beta = 0,5$, les valeurs de m_H inférieures à environ 400 GeV sont exclues, et dans le cas $\tan\beta = 0,3$ la limite vaut environ 750 GeV.

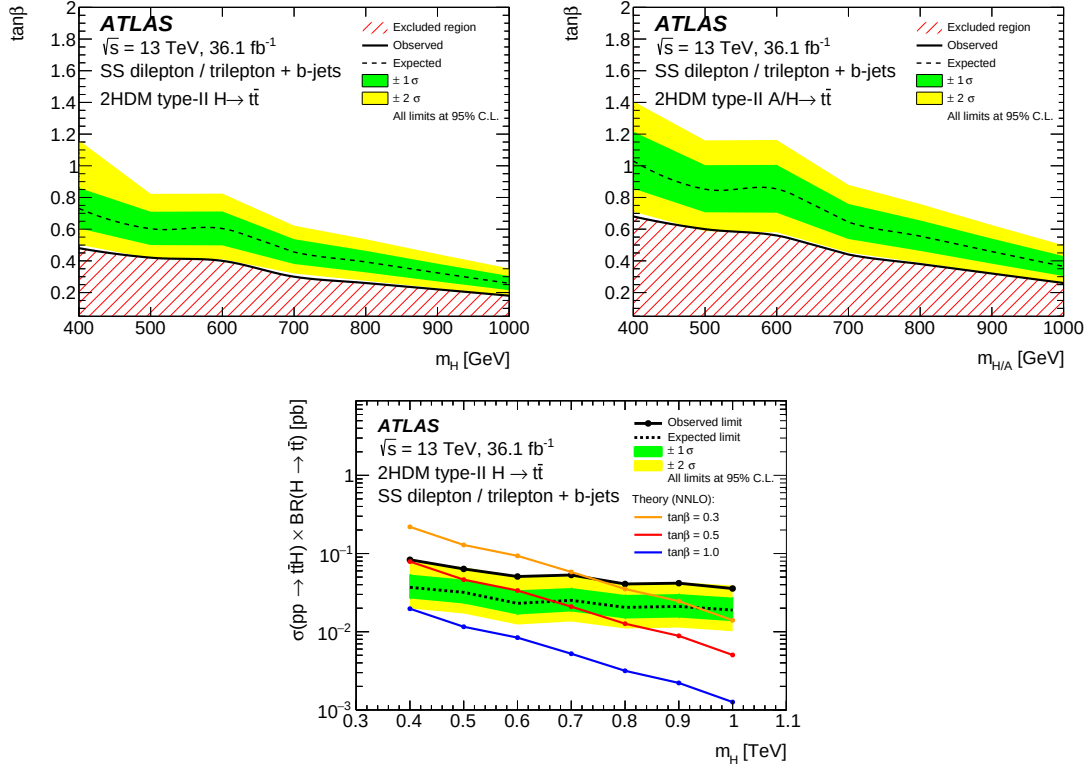


Figure 5.17 – Limites sur le modèle à deux doublets de Higgs 2HDM, dans le plan $(m_H, \tan\beta)$ (haut), et sur la section efficace du processus (bas). Le cas pour lequel seul le boson H contribue au processus de production de quatre quarks top (haut, gauche) est considéré, ainsi que celui pour lequel les bosons H et A contribuent à ce processus (haut, droite). La limite sur la section efficace est donnée pour trois valeurs de $\tan\beta$ dans le cas où seul le le boson H contribue au processus. Les droites bleue, rouge et orange représentent les sections efficaces théoriques de production de quatre quarks top via ce modèle, pour les trois valeurs de $\tan\beta$.

5.7.5 Résultats combinés avec l'analyse à un seul lepton

Les résultats présentés précédemment concernent la production de quatre quarks top dans les événements comportant au moins deux leptons de même charge électrique. Un autre canal est considéré pour l'étude du processus $pp \rightarrow t\bar{t}t\bar{t}$ comme décrit dans

la section 4.1.2, utilisant les événements avec un seul lepton ou deux leptons de charge opposée. Les résultats des deux canaux sont combinés et présentés dans la référence décrivant l'analyse dans le canal à un seul lepton [8].

La sélection des objets et des événements dans les deux analyses sont similaires, cependant il existe quelques différences dans les critères appliqués sur les leptons. C'est pourquoi les deux analyses ne sont pas parfaitement orthogonales. Autrement dit il est possible que certains événements passent les coupures des deux analyses à la fois. L'orthogonalité a été vérifiée sur des événements simulés de signal avant de combiner les résultats. Le recouvrement maximal obtenu entre deux régions de signal des deux analyses est de 5 %. Cette valeur est faible, ainsi les recouvrements entre régions de signal sont négligés.

Les incertitudes systématiques expérimentales des deux analyses sont corrélées entre les régions de signal des deux analyses, tandis que les incertitudes de modélisation des bruits de fond sont décorrélées. En effet la composition du bruit de fond ainsi que les méthodes d'estimation des bruits de fond instrumentaux sont très différentes d'une analyse à l'autre. L'ensemble des paramètres de nuisance des deux analyses, ainsi que toutes les régions de signal des deux canaux sont introduits dans un nouveau modèle statistique qui est ajusté sur les données avec la même méthode que celle détaillée dans la section 5.6. Un ajustement de la meilleure valeur de μ est réalisé, et une limite sur $\sigma_{t\bar{t}t\bar{t}}$ est calculée. D'autres configurations pour cette combinaison sont considérées, elles diffèrent de la configuration standard par la décorrélation de l'incertitude principale associée à l'identification des jets de quarks b , ou bien de toutes les incertitudes d'identification des jets de quarks b , ou encore par la décorrélation de toutes les incertitudes. Dans tous ces cas de figure l'ajustement diffère de moins de 5 %.

La figure 5.18 (gauche) présente les deux limites ainsi que la limite combinée. Les lignes pointillées ainsi que les bandes colorées vertes (jaunes) représentent les limites attendues avec leur bande d'incertitude à $\pm 1\sigma$ ($\pm 2\sigma$). Les lignes pointillées rouges représentent les limites attendues en présence d'un signal conforme au modèle standard, c'est-à-dire pour $\mu = 1$. Les lignes noires en trait plein représentent les limites observées. La figure 5.18 (droite) présente les meilleurs ajustements de la force du signal μ dans chaque cas, ainsi que le meilleur ajustement de μ combiné. Les barres bleu clair (respectivement foncé) représentent les incertitudes statistiques (respectivement totales).

La limite supérieure observée (attendue) sur la section efficace est de 49 fb (19 fb) après combinaison des résultats, correspondant à 5,3 (2,1) fois la prédiction théorique. Le meilleur ajustement combiné de la force du signal μ est de $28,5^{+12}_{-11}$ fb, pour une prédiction théorique de $9,2^{+2,9}_{-2,4}$ fb. On peut remarquer que l'analyse dans le canal à deux leptons de même charge est dominée par les incertitudes statistiques, tandis que l'analyse dans le canal à un seul lepton est dominée par les incertitudes systématiques.

La compatibilité entre les deux canaux a également été vérifiée, en particulier puisque

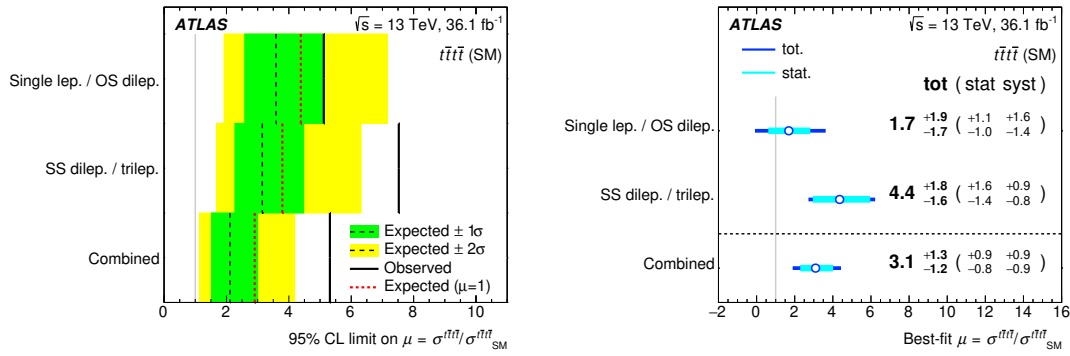


Figure 5.18 – Combinaison des résultats obtenus avec ceux de l’analyse dans les canaux à un seul lepton et à deux leptons de même charge électrique. Combinaison des limites (gauche) et des meilleurs ajustements (droite).

l’excès d’événements observés est plus grand dans le canal à deux leptons de même charge que dans le canal à un seul lepton. Pour cela, un ajustement avec deux forces de signal différentes a été effectué, et les deux forces de signal résultantes ont été comparées à la force du signal commune obtenue via la combinaison. La probabilité d’obtenir une différence entre les deux forces de signal supérieure ou égale à la différence observée vaut environ 31 %. Les deux résultats peuvent donc être considérés comme compatibles.

Une combinaison des résultats pour le modèle CI est également réalisée et place une limite observée (attendue) de 21 fb (15 fb) sur le processus, ou encore une limite observée (attendue) sur le paramètre libre du modèle $|C_{4t}|/\Lambda^2 < 1,9 \text{ TeV}^{-2}$ ($1,6 \text{ TeV}^{-2}$).

5.8 Conclusion

Une analyse des données du Run 2 prises en 2015 et en 2016 par le détecteur ATLAS au LHC étudie les événements comportant au moins deux leptons ayant la même charge électrique et des jets de quarks b . Le jeu de données représente une luminosité intégrée de $36,1 \text{ fb}^{-1}$ à une énergie de 13 TeV dans le centre de masse. L’objectif de cette analyse est de sonder différents modèles BSM dont certains prédisent l’augmentation de la section efficace de production de quatre quarks top $\sigma_{t\bar{t}t\bar{t}}$. Cette analyse s’intéresse également au processus de production de quatre quarks top conforme au modèle standard.

La sélection des événements est effectuée dans huit régions de signal, communes à tous les modèles testés. Les événements comportant deux leptons de même charge, et ceux avec au moins trois leptons sont traités séparément. Seuls les événements avec au moins un jet de quark b sont considérés, et la sensibilité de l’analyse aux signaux de production de quatre quarks top augmente avec le nombre de jets de quarks b .

Les bruits de fond principaux de l’analyse sont les processus physiques de production

d'une paire de quarks top associée à un boson W , Z ou H , ainsi que les bruits de fond dits réductibles, provenant d'erreurs d'identification de particules. Les bruits de fond physiques sont estimés via des simulations Monte-Carlo, tandis que les bruits de fond réductibles sont estimés à partir de données. Le nombre d'événements dus aux faux leptons et aux leptons non-prompts est estimé via la méthode de la matrice, fondée sur la mesure du taux de leptons identifiés par des coupures lâches qui passent les coupures d'identification strictes. Le bruit de fond associé à la mauvaise identification de la charge d'un électron est quand à lui estimé à partir des événements avec deux leptons de charge opposée ainsi que la mesure des taux de mauvaise identification.

L'accord entre l'estimation du bruit de fond et les données est vérifié dans des régions de validation, puis un modèle statistique est construit à partir des nombres d'événements dans les régions de signal sous forme d'un produit de vraisemblances. Les incertitudes systématiques sont introduites dans le modèle en tant que paramètres de nuisance pouvant être contraints lors de son ajustement. Une limite supérieure est alors calculée sur la section efficace de chaque modèle étudié, à partir d'un rapport de vraisemblances. Les valeurs des limites sont données avec un niveau de confiance de 95 %.

La limite supérieure observée (attendue) sur la section efficace de production de quatre quarks top conforme au modèle standard $\sigma_{tttt}^{\text{MS}}$ est de 69 fb (29 fb). L'analyse décrite dans ce chapitre est combinée avec l'analyse utilisant les canaux à un seul lepton et à deux leptons de même charge électrique, et la limite combinée observée (attendue) sur $\sigma_{tttt}^{\text{MS}}$ vaut 49 fb (19 fb). L'analyse combinée effectue également une mesure de $\sigma_{tttt}^{\text{MS}} = 28,5^{+12}_{-11} \text{ fb}$. Ce résultat est le plus précis au moment de sa publication.

L'analyse s'intéressant aux événements avec deux leptons ayant la même charge électrique est dominée par l'incertitude statistique. L'étude des données prises par le détecteur ATLAS durant l'entièreté du Run 2 permettra donc une amélioration conséquente de ce résultat. Une nouvelle analyse menée sur les données du Run 2 est décrite dans le chapitre 6. Elle s'intéresse exclusivement à la production de quatre quarks top, et vise à améliorer les méthodes d'estimation des bruits de fond majeurs ainsi que la synergie entre les canaux à un seul lepton et à deux leptons de même charge électrique.

Références

- [1] ATLAS COLLABORATION. “Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Journal of High Energy Physics* 2018.12 (déc. 2018). DOI : 10.1007/jhep12(2018)039.
- [2] ATLAS COLLABORATION. “Analysis of events with b-jets and a pair of leptons of the same charge in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In : *Journal of High Energy Physics* 2015.10 (oct. 2015). DOI : 10.1007/jhep10(2015)150.
- [3] ATLAS COLLABORATION. *Search for new physics using events with b-jets and a pair of same charge leptons in 3.2fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Rapp. tech. ATLAS-CONF-2016-032. Geneva : CERN, juin 2016. URL : <https://cds.cern.ch/record/2161545>.
- [4] CMS COLLABORATION. “Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In : *The European Physical Journal C* 77.9 (sept. 2017). DOI : 10.1140/epjc/s10052-017-5079-z.
- [5] CMS COLLABORATION. “Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In : *The European Physical Journal C* 78.2 (fév. 2018). DOI : 10.1140/epjc/s10052-018-5607-5.
- [6] CMS COLLABORATION. “Search for the production of four top quarks in the single-lepton and opposite-sign dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In : (6 juin 2019). arXiv : 1906.02805v1 [hep-ex].
- [7] CMS COLLABORATION. *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Rapp. tech. CMS-PAS-TOP-18-003. Geneva : CERN, 2019. URL : <https://cds.cern.ch/record/2668710>.
- [8] ATLAS COLLABORATION. “Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Physical Review D* 99.5 (mar. 2019). DOI : 10.1103/physrevd.99.052009.
- [9] ATLAS COLLABORATION. *Electron identification measurements in ATLAS using $\sqrt{s} = 13$ TeV data with 50 ns bunch spacing*. Rapp. tech. ATL-PHYS-PUB-2015-041. Geneva : CERN, sept. 2015. URL : <https://cds.cern.ch/record/2048202>.

- [10] ATLAS COLLABORATION. “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ ”. In : *The European Physical Journal C* 76.5 (mai 2016). DOI : 10.1140/epjc/s10052-016-4120-y.
- [11] Matteo CACCIARI, Gavin P SALAM et Gregory SOYEZ. “The anti-kt jet clustering algorithm”. In : *Journal of High Energy Physics* 2008.04 (avr. 2008), p. 063-063. DOI : 10.1088/1126-6708/2008/04/063.
- [12] ATLAS COLLABORATION. *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*. Rapp. tech. ATLAS-CONF-2015-029. Geneva : CERN, juil. 2015. URL : <https://cds.cern.ch/record/2037702>.
- [13] ATLAS COLLABORATION. *Tagging and suppression of pileup jets with the ATLAS detector*. Rapp. tech. ATLAS-CONF-2014-018. Geneva : CERN, mai 2014. URL : <https://cds.cern.ch/record/1700870>.
- [14] ATLAS COLLABORATION. “Performance of b -jet identification in the ATLAS experiment”. In : *Journal of Instrumentation* 11.04 (avr. 2016), P04008-P04008. DOI : 10.1088/1748-0221/11/04/p04008.
- [15] ATLAS COLLABORATION. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ ”. In : *The European Physical Journal C* 78.11 (nov. 2018). DOI : 10.1140/epjc/s10052-018-6288-9.
- [16] J. ALWALL et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In : *Journal of High Energy Physics* 2014.7 (juil. 2014). DOI : 10.1007/jhep07(2014)079.
- [17] Torbjörn SJÖSTRAND et al. “An introduction to PYTHIA 8.2”. In : *Computer Physics Communications* 191 (juin 2015), p. 159-177. DOI : 10.1016/j.cpc.2015.01.024.
- [18] T GLEISBERG et al. “Event generation with SHERPA 1.1”. In : *Journal of High Energy Physics* 2009.02 (fév. 2009), p. 007-007. DOI : 10.1088/1126-6708/2009/02/007.
- [19] Erich W. VARNES. “A Poisson likelihood approach to fake lepton estimation with the matrix method”. In : (2016). arXiv : 1606.06817 [hep-ex].
- [20] ATLAS COLLABORATION. “Luminosity determination in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ using the ATLAS detector at the LHC”. In : *The European Physical Journal C* 76.12 (nov. 2016). DOI : 10.1140/epjc/s10052-016-4466-1.
- [21] G. AVONI et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In : *Journal of Instrumentation* 13.07 (juil. 2018), P07017-P07017. DOI : 10.1088/1748-0221/13/07/p07017.

- [22] A L READ. “Presentation of search results : the CL_s technique”. In : *Journal of Physics G : Nuclear and Particle Physics* 28.10 (sept. 2002), p. 2693-2704. DOI : 10.1088/0954-3899/28/10/313.
- [23] Glen COWAN et al. “Asymptotic formulae for likelihood-based tests of new physics”. In : *The European Physical Journal C* 71.2 (fév. 2011), p. 1-19. DOI : 10.1140/epjc/s10052-011-1554-0.

6

Vers la mise en évidence de la production modèle standard de quatre quarks tops avec le détecteur ATLAS

La production de quatre quarks top dans les collisions à 13 TeV au LHC a été étudiée par les collaborations ATLAS et CMS avec les données prises en 2015 et en 2016 représentant une luminosité intégrée d'environ 36 fb^{-1} . De nouvelles analyses sont en cours de réalisation sur l'intégralité des données du Run 2 s'étendant de 2015 à 2018, et correspondant à une luminosité intégrée d'environ 139 fb^{-1} . Un résultat préliminaire a notamment été obtenu par la collaboration CMS [1] avec les données du Run 2. Ce chapitre décrit l'analyse du processus de production de quatre quarks top conforme au modèle standard avec le détecteur ATLAS pour l'ensemble des données du Run 2. Cette analyse est en cours et n'est pas encore publique. On s'intéresse plus particulièrement au canal avec au moins deux leptons ayant la même charge électrique dans l'état final, dans la continuité de l'analyse présentée dans le chapitre précédent [2].

Cette nouvelle analyse s'intéresse exclusivement au processus de production de quatre quarks top prédit par le modèle standard, contrairement à l'analyse utilisant $36,1 \text{ fb}^{-1}$ qui étudiait principalement des modèles de nouvelle physique. De plus une synergie est recherchée entre d'un côté le canal constitué des états finaux avec un seul lepton ou avec deux leptons ayant la même charge électrique, et de l'autre le canal constitué des états finaux avec deux leptons ayant la même charge électrique ou trois leptons. Le premier canal est appelé « canal à un lepton » dans la suite de ce chapitre, tandis que le second canal est qualifié de « canal à deux leptons de même charge » et est parfois noté « SSML ». Les leptons τ ne sont pas reconstruits donc le terme « lepton » et le symbole ℓ désignent un électron e ou un muon μ .

L'analyse préliminaire menée par la collaboration CMS avec les données du Run 2 [1] vise le canal à deux leptons de même charge. Elle obtient une signification statistique observée (attendue) de 2,6 (2,7) déviations standard. La section efficace théorique utilisée dans cette analyse ainsi que dans le travail présenté dans ce chapitre provient d'un calcul au second ordre dominant (NLO) incluant les corrections QCD mais également les corrections électrofaibles, comme expliqué dans la section 4.1.1. La valeur de section efficace est $\sigma_{t\bar{t}t\bar{t}}^{\text{SM}} = 12,0^{+2,2}_{-2,5} \text{ fb}$, l'incertitude incluant à la fois les incertitudes sur les PDF et les incertitudes sur les échelles [3].

Mes contributions à l'analyse présentée dans ce chapitre concernent les études d'op-

timisation dans le canal à deux leptons de même charge, en particulier l'établissement des régions de signal dans ce canal qui sont décrites dans la section 6.1.3, la mise en place d'une nouvelle méthode d'estimation du bruit de fond dû aux faux leptons appelée méthode des *templates* et présentée dans la section 6.2, l'établissement du modèle statistique dans le canal à deux leptons de même charge et la préparation à la combinaison statistique des deux canaux de l'analyse. Ce chapitre se concentre donc sur le canal à deux leptons de même charge.

6.1 Vue d'ensemble de l'analyse

6.1.1 Sélection des objets physiques

La sélection des objets de l'analyse pour le canal à deux leptons de même charge est presque identique à la sélection présentée dans la section 5.2. Les coupures sur l'impulsion transverse des objets sont toujours $p_T > 28 \text{ GeV}$ pour les électrons et les muons, et $p_T > 25 \text{ GeV}$ pour les jets. La seule différence concerne les électrons dans les états finaux avec exactement deux leptons ayant la même charge électrique et avec au moins un électron c'est-à-dire les états finaux ee et $e\mu$. L'analyse présentée dans le chapitre 5 applique une coupure spécifique $|\eta| < 1.37$ sur ces électrons, afin de réduire le bruit de fond dû à la mauvaise identification de leur charge. Cette coupure est désormais remplacée par une coupure sur un discriminant multivarié réduisant le bruit de fond de mauvaise identification de la charge. Ce discriminant utilise des grandeurs associées à l'électron : p_T , $|\eta|$, E/p , charge/ d_0 , $\Delta\phi$ entre l'amas calorimétrique et la trace extrapolée au niveau de calorimètre, ainsi que la charge moyenne pondérée par le nombre de coups dans le sous-détecteur SCT.

L'algorithme par défaut pour l'identification des jets de quarks b reste MV2c10 pour une efficacité de 77 %, mais d'autres options sont explorées. Il est envisagé d'utiliser l'algorithme DL1 utilisant un réseau de neurones à apprentissage profond. Cet algorithme prend en entrée les mêmes grandeurs que l'algorithme MV2c10 et fournit des résultats multidimensionnels correspondant aux probabilités pour un jet de provenir d'un quark b , c , ou d'un quark léger, grâce à plusieurs couches de neurones artificiels interconnectées [4]. Une autre piste considérée dans cette analyse en développement est l'utilisation de plusieurs points de fonctionnement d'identification des jets de quarks b , dite identification pseudo-continue. Les jets sont classés dans cinq catégories selon leur score d'identification, au lieu d'être simplement identifiés ou non. Par exemple les jets passant le critère le plus strict à 60 % d'efficacité constituent la première catégorie, puis les jets restants passant le critère à 70 % d'efficacité constituent la seconde catégorie, et ainsi de suite jusqu'aux jets ne passant aucun critère.

6.1.2 Pré-sélection des événements

Les données utilisées correspondent aux données du Run 2 en collisions proton-proton au LHC avec une énergie dans le centre de masse de 13 TeV. Ces données ont été prises entre 2015 et 2018, et représentent une luminosité intégrée de $139,4 \pm 2,4 \text{ fb}^{-1}$. L'empilement moyen μ varie entre 13 et 38 selon l'année de la prise de données et vaut environ 34 pour le jeu de données dans son intégralité. La section efficace théorique du processus $pp \rightarrow t\bar{t}t\bar{t}$ vaut environ 12 fb [3], ainsi on s'attend à ce qu'environ 1 700 événements à quatre quarks top aient été produits.

Les périodes de prises de données considérées sont celles pour lesquelles tous les sous-détecteurs ont fonctionné de manière nominale, et pour lesquelles le LHC était en mode de collision stable. Les événements utilisés par l'analyse doivent comporter un vertex primaire associé à au moins deux particules chargées vérifiant $p_T > 0,4 \text{ GeV}$. Les événements sélectionnés doivent correspondre à un mode de déclenchement à un lepton de haut p_T . Plusieurs déclenchements sont pris en compte, sans critère d'isolation du lepton sauf pour les déclenchements utilisant les valeurs de p_T les plus basses. Ces valeurs sont 26 GeV (24 GeV) pour les électrons et 26 GeV (20 GeV) pour les muons, pour les données prises en 2016, 2017, et 2018 (2015).

Les événements sont pré-sélectionnés et classés par catégorie. Les différents critères de sélection sont détaillés dans ce paragraphe et résumés dans le tableau 6.1. Les événements comportant un seul lepton ou bien deux leptons de charge opposée sont classés dans les catégories correspondant au canal à un lepton, s'il contiennent au moins deux jets de quarks b et cinq jets (quatre jets pour les événements à deux leptons de charge opposée). Les catégories sont notées e , μ , $e^\pm e^\mp$, $e^\pm \mu^\mp$, $\mu^\pm \mu^\mp$. Pour le canal à deux leptons de même charge, les événements doivent comporter au moins un jet de quark b et au moins deux leptons ayant la même charge électrique. Si l'événement ne comporte que deux leptons de même charge, il est catégorisé suivant le type des leptons : $e^\pm e^\pm$, $e^\pm \mu^\pm$, $\mu^\pm \mu^\pm$, et de même si l'événement comporte au moins trois leptons : eee , $ee\mu$, $e\mu\mu$, $\mu\mu\mu$. Les événements avec quatre leptons ou plus sont pris en compte comme des événements à trois leptons. Les événements avec deux électrons de même charge subissent une coupure additionnelle de rejet du boson Z basée sur la masse invariante des deux électrons et la masse du boson Z : $|m_{\ell\ell} - m_Z| > 10 \text{ GeV}$, avec $m_Z = 91 \text{ GeV}$. Les événements avec trois leptons comportant une paire de leptons de même saveur (e ou μ) et de charge opposée subissent la même coupure basée sur la masse invariante de ces deux leptons afin de réduire le bruit de fond $t\bar{t}Z$.

6.1.3 Définition des régions

Les événements pré-sélectionnés sont ensuite classés dans des régions de contrôle, de validation, ou de signal. Les régions de contrôle sont utilisées afin d'estimer des bruits

Catégorie(s)	Critère sur N_b	Critère additionnel
e, μ	$N_b \geq 2$	$N_j \geq 5$
$e^\pm e^\mp, e^\pm \mu^\mp, \mu^\pm \mu^\mp$	$N_b \geq 2$	$N_j \geq 4$
$e^\pm e^\pm$	$N_b \geq 1$	$ m_{ee} - m_Z > 10 \text{ GeV}$
$e^\pm \mu^\pm, \mu^\pm \mu^\pm$	$N_b \geq 1$	
$eee, ee\mu, e\mu\mu, \mu\mu\mu$	$N_b \geq 1$	$ m_{\ell\ell} - m_Z > 10 \text{ GeV}$

Table 6.1 – Résumé des catégories de l’analyse et de la pré-sélection des événements. Pour les événements avec deux électrons de même charge, le critère additionnel utilise la masse invariante des deux électrons m_{ee} . Pour les événements avec trois leptons, le critère additionnel s’applique s’il existe une paire de leptons de même saveur et de charge opposée, $m_{\ell\ell}$ étant leur masse invariante.

de fond et contraindre des incertitudes systématiques, tandis que les régions de signal visent un haut rapport signal sur bruit. Les régions de validation ont pour rôle de vérifier la modélisation des bruits de fond en comparant leur estimation aux données dans une région de l’espace des phases proche des régions de signal mais contenant peu de signal. Les grandeurs utilisées pour définir ces régions doivent permettre une bonne distinction entre le signal $t\bar{t}t\bar{t}$ et les bruits de fond. Le bruit de fond principal dans le canal à un lepton est la production d’une paire de quarks top $t\bar{t}$ accompagnée de jets, tandis que le canal à deux leptons de même charge a pour bruits de fond principaux la production d’une paire de quarks top accompagnée d’un boson ($t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$). Dans les deux canaux on s’intéresse donc aux événements ayant un grand nombre de jets noté N_j , un grand nombre de jets de quarks b noté N_b , puisque ces grandeurs sont en moyenne plus élevées pour le signal que pour les bruits de fond. De plus l’énergie totale dans l’événement doit être grande pour produire quatre quarks top, ainsi la grandeur H_T est également utilisée. Il s’agit de la somme des impulsions transverses des leptons et de jets reconstruits dans l’événement.

Dans le canal à deux leptons de même charge, les régions sont définies à partir des grandeurs N_b et N_j , mais aussi H_T . Les définitions s’inspirent des régions de l’analyse utilisant $36,1 \text{ fb}^{-1}$, avec une catégorisation suivant le nombre de jets de quarks b : 2, 3 ou ≥ 4 . C’est l’augmentation de la luminosité intégrée qui permet d’envisager l’utilisation de la région $N_b \geq 4$. De plus les événements à deux leptons de même charge et les événements à trois leptons sont classés dans des régions distinctes, sauf pour les événements avec quatre jets de quarks b qui sont très peu nombreux. La coupure sur la grandeur H_T a été placée assez basse puisque la distribution de cette grandeur est utilisée comme discriminant dans les régions de signal. Le critère choisi est $H_T > 500 \text{ GeV}$. Enfin, les coupures sur le nombre de jets ont été optimisées à partir d’une analyse statistique

Nom	Catégorie	N_b	N_j	$H_T[\text{GeV}]$
SR2b2 ℓ	$\ell^\pm \ell^\pm$	2	≥ 6	> 500
SR3b2 ℓ	$\ell^\pm \ell^\pm$	3	≥ 6	> 500
SR2b3 ℓ	lll	2	≥ 6	> 500
SR3b3 ℓ	lll	3	≥ 6	> 500
SR4b	$\ell^\pm \ell^\pm$ et lll	≥ 4	≥ 6	> 500

Table 6.2 – Définition des régions de signal du canal à deux leptons de même charge. La catégorie $\ell^\pm \ell^\pm$ désigne les événements avec deux leptons de même charge tandis que la catégorie lll désigne les événements avec trois leptons.

simplifiée en maximisant la signification statistique. Les coupures obtenues sont $N_j \geq 6$ pour les événements à deux leptons de même charge, et $N_j \geq 5$ pour les événements à trois leptons. Une coupure globale $N_j \geq 6$ est finalement définie étant donné que la différence entre les deux coupures dans le cas des événements à trois leptons est très faible. Ceci permet de simplifier la définition des régions de contrôle et de validation dans le cadre de la méthode des templates décrite dans la section 6.2. Les critères des régions de signal sont résumés dans le tableau 6.2. Ces définitions sont préliminaires et donc susceptibles d'évoluer.

6.1.4 Stratégie d'extraction du signal

Le modèle statistique est construit suivant la même méthode que celle présentée dans la section 5.6. La différence principale avec la méthode présentée dans le chapitre 5 est que l'ajustement est effectué sur des distributions de variables discriminantes dans les régions de signal et de contrôle, plutôt que sur les simples nombres d'événements dans ces régions. La grandeur H_T est utilisée à l'heure actuelle. À terme il est prévu de remplacer la grandeur H_T par un classificateur multivarié, par exemple un arbre de décision boosté. L'ajustement sera alors effectué sur les distributions des valeurs de sortie de ce classificateur dans les régions de signal. Le classificateur pourrait notamment prendre en entrée les valeurs du discriminant utilisé pour identifier les jets de quarks b . Ainsi au lieu d'une distinction binaire entre les jets identifiés comme provenant d'un quark b et les autres, il serait possible d'utiliser le continuum de valeurs du discriminant permettant l'identification.

L'utilisation de distributions de variables discriminantes au lieu de simples nombres d'événements entraîne l'apparition d'incertitudes de forme des distributions. L'incertitude sur certains paramètres physiques peuvent en effet déplacer un ou plusieurs événements d'un intervalle à un autre dans la distribution et ainsi modifier sa forme. Il faut également tenir compte des corrélations entre les différents intervalles d'une même distribution.

Au stade actuel de l'analyse, les données ne sont utilisées que dans les régions où l'on attend une quantité très faible de signal afin de s'assurer que la recherche du signal n'est pas biaisée par le signal lui-même. Pour le canal à deux leptons de même charge, il s'agit des événements vérifiant $H_T < 500$ GeV ou respectant les critères $N_b \leq 2$ et $N_j \leq 5$. Les autres régions de l'espace des phases peuvent contenir une quantité non-négligeable de signal et sont donc opérées en aveugle c'est-à-dire que les données ne sont pas utilisées, seules les simulations Monte-Carlo le sont.

6.1.5 Méthodes d'estimation des bruits de fond

La composition du bruit de fond dans le canal à deux leptons de même charge a déjà été présentée dans la section 5.4. Plusieurs processus physiques peuvent entrer dans les régions de signal, et les bruits de fond principaux sont la production d'une paire de quarks top accompagnée d'un boson : $t\bar{t}W$, $t\bar{t}Z$, et $t\bar{t}H$. Ces bruits de fond, ainsi que d'autres processus secondaires, sont estimés à partir de simulations Monte-Carlo. Les générateurs utilisés sont MADGRAPH5_AMC@NLO [5] ($t\bar{t}W$, $t\bar{t}Z$) ou POWHEGBOX [6] ($t\bar{t}H$) et PYTHIA [7]. La modélisation des processus $t\bar{t}W$ et $t\bar{t}Z$ peut toutefois présenter des difficultés, surtout à grand nombre de jets de quarks b . Ainsi deux régions de contrôle pour les bruits de fond $t\bar{t}W$ et $t\bar{t}Z$ sont définies comme précisé dans la section 6.2. Il est en particulier envisagé d'ajuster la normalisation du bruit de fond $t\bar{t}W$ durant l'ajustement du modèle statistique, puisqu'il s'agit d'un bruit de fond majeur qui influence le résultat, mais aussi parce que les études récentes de ce processus mesurent une valeur de section efficace légèrement supérieure à la valeur théorique. Par exemple l'analyse effectuée par la collaboration ATLAS avec les données prises en 2015 et en 2016 mesure un rapport entre la section efficace observée pour le processus $t\bar{t}W$ et la section efficace théorique de $1,44 \pm 0,32$ [8].

Les autres bruits de fond du canal à deux leptons de même charge proviennent de la mauvaise reconstruction de la charge des électrons d'une part, et des faux leptons d'autre part. Ces bruits de fond réductibles ainsi que leurs méthodes d'estimation utilisées dans l'analyse présentée dans le chapitre 5 sont décrits dans la section 5.4.2. L'estimation du bruit de fond de mauvaise identification de la charge des électrons est réalisée de manière similaire à celle présentée dans la section 5.4.2. Cette estimation est préliminaire et n'est pas encore interfacée avec le reste de l'analyse, il n'a donc pas été possible de l'utiliser dans l'étude présentée dans ce chapitre. La figure 6.1 montre l'estimation des taux de mauvaise identification en fonction des grandeurs p_T et $|\eta|$ dans les données et les simulations Monte-Carlo. Les estimations sont toutes compatibles, mais les critères de sélection utilisés ne sont pas encore finaux et le discriminant multivarié présenté dans la section 6.1.1 permettant de réduire le bruit de fond de mauvaise identification n'est pas encore inclus dans cette estimation.

En ce qui concerne le bruit de fond dû aux faux leptons, deux options sont envisagées.

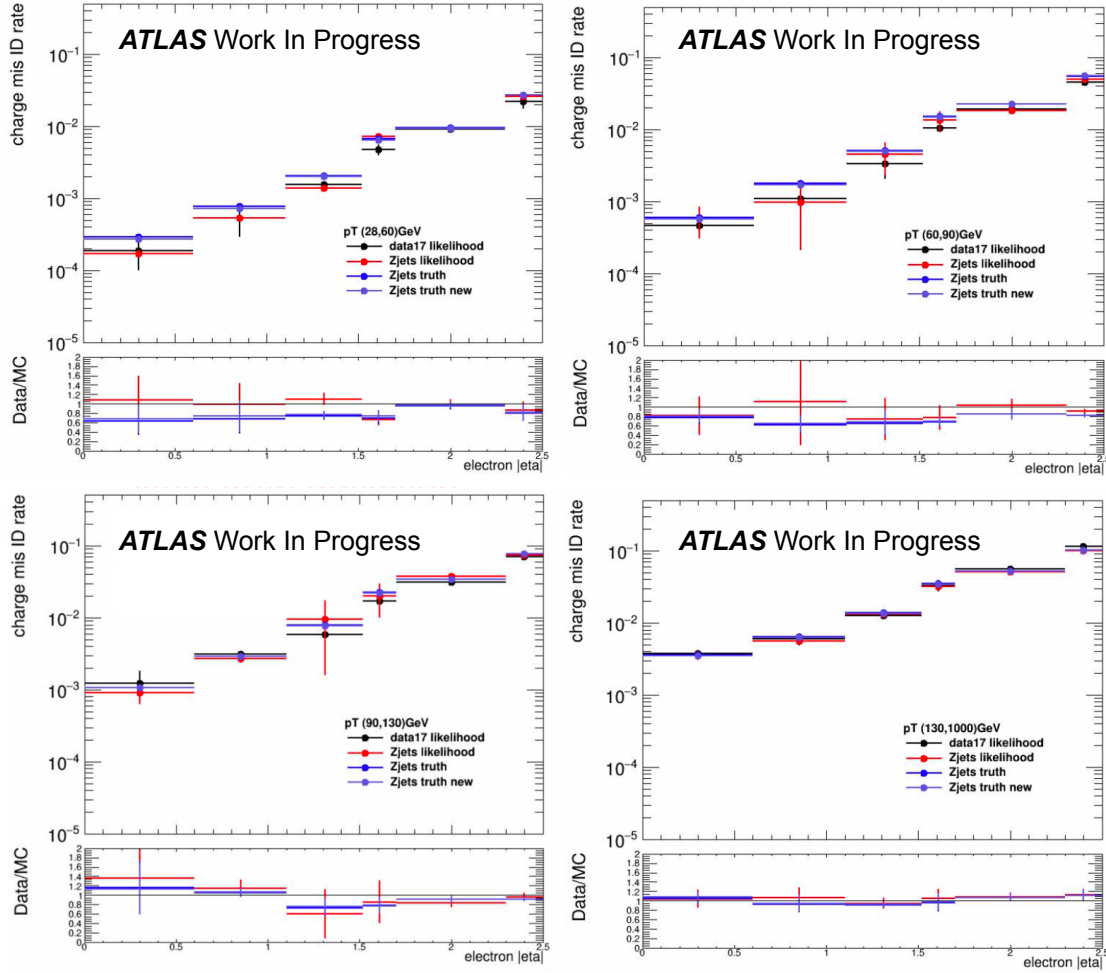


Figure 6.1 – Taux de mauvaise identification de la charge des électrons en fonction de $|\eta|$ pour quatre intervalles de p_T . Les taux utilisant la méthode standard sur les données sont représentés en rouge, et ceux estimés sur simulation sont représentés en noir. Ils sont comparés aux taux estimés à partir d'informations Monte-Carlo, représentés en bleu et en gris.

La première option est d'utiliser la méthode de la matrice c'est-à-dire la même méthode que celle présentée dans la section 5.4.2, et la seconde option est une nouvelle méthode appelée méthode des templates. Cette dernière est détaillée dans la section 6.2.

6.2 Méthode des templates

Le bruit de fond dû aux faux leptons est engendré par plusieurs processus physiques distincts. Le terme « faux » leptons est utilisé ici et dans la suite de manière générique pour faire référence à toutes les composantes de ce bruit de fond. Ces leptons sont soit des leptons non-prompts c'est-à-dire des leptons provenant de désintégrations secondaires et non de l'interaction initiale, soit des faux leptons c'est-à-dire des particules hadroniques reconstruites en tant que lepton. Afin d'obtenir une estimation de ce bruit de fond dans les régions de signal de l'analyse, la méthode dite des templates est envisagée.

La méthode des templates utilise à la fois les simulations Monte-Carlo et les données. Son principe est de se reposer sur les simulations Monte-Carlo pour décrire la cinématique des événements contenant des faux leptons, mais de corriger leur normalisation grâce aux données. Les composantes du bruit de fond provenant de faux leptons sont traitées de manière indépendante, et autant de facteurs de normalisation sont définis dans le modèle statistique. Ces facteurs sont contraints grâce aux données dans des régions de contrôle optimisées pour certains types de faux leptons. L'estimation finale du bruit de fond dû aux faux leptons dans les régions de signal correspond donc à la somme des composantes de ce bruit de fond telles que modélisées par Monte-Carlo, et normalisées par un facteur qui est ajusté sur les données en même temps que le reste du modèle statistique. Des incertitudes systématiques doivent être introduites sur cette méthode, pour tenir compte de l'extrapolation des normalisations depuis les régions de contrôle vers les régions de signal.

Les événements entrant dans les régions de signal à cause des faux leptons peuvent provenir de nombreux processus physiques différents comme expliqué dans la section 5.4.2. Ces événements proviennent toutefois majoritairement du processus $t\bar{t}$. Dans le travail présenté ici, des simulations des processus suivants sont utilisées pour estimer le bruit de fond réductible :

- la production d'une paire de quarks top $t\bar{t}$;
- la production d'un seul quark top qui se présente accompagné d'un quark (tq) ou d'un boson (tW), en excluant cependant les processus tZ ou tWZ qui peuvent produire deux leptons de même charge et sont donc des bruit de fond irréductibles.

6.2.1 Composition du bruit de fond dû aux faux leptons

Il est nécessaire de distinguer les différents types de processus physiques produisant des faux leptons et de les traiter séparément. Plusieurs catégories sont définies à partir

des informations Monte-Carlo :

- les leptons « prompts » sont les leptons standards qui ne sont pas des faux leptons ;
- les électrons « trident » proviennent d'un processus trident $e \rightarrow e\gamma \rightarrow eee$;
- les « conversions » désignent les leptons provenant d'un photon, on distingue les conversions issues de photons émis par interaction avec le matériau du détecteur des conversions issues de photons virtuels émis au niveau de l'interaction ;
- les leptons « HF » proviennent de la désintégration semi-leptonique d'un jet de quark b ou c ;
- les électrons « light » proviennent de l'identification d'un jet en tant qu'électron ;
- enfin, les leptons « autres » sont ceux qui ne correspondent à aucune des catégories listées ci-dessus, par exemple quand on n'arrive pas à déterminer leur origine.

Les électrons prompts et trident peuvent voir leur charge mal reconstruite. Des événements contenant de tels leptons doivent être ignorés pour éviter de les compter deux fois, puisque le bruit de fond dû à la mauvaise identification de la charge est estimé via une méthode dédiée, similaire à celle présentée dans la section 5.4.2. Pour le moment ces leptons ne sont pas ignorés et sont donc inclus dans l'estimation du bruit de fond par la méthode des templates. Les électrons trident qui ont une charge bien reconstruite ne sont pas considérés comme faux puisqu'ils sont sélectionnés lors de l'étude des bruits de fond irréductibles. Ils ont certes une impulsion transverse inférieure à l'électron d'origine et les événements associés ont une cinématique différente des événements ne contenant que des leptons prompts, mais il n'a pas été décidé de leur réserver un traitement particulier. Ils représentent une faible proportion des électrons.

Après cette classification par lepton, une classification au niveau de l'événement est effectuée selon le type de leptons qu'il contient. Par exemple les événements contenant uniquement des leptons « prompts » sont dits « physiques » et sont ignorés puisqu'ils ne contiennent aucun faux lepton. Les événements qui contiennent au moins deux faux leptons sont classés dans la catégorie appelée « autre », ils sont présents en quantité négligeable ($\lesssim 0,1\%$). Les événements qui contiennent exactement un faux lepton et dont tous les autres leptons sont prompts sont classés selon le type de faux lepton. On obtient ainsi quatre ou cinq catégories :

- conversion ;
- HF ;
- light ;
- mauvaise identification de la charge (temporaire, en attente d'une estimation dédiée) ;
- autre.

Les deux types de conversions sont classées dans la même catégorie puisqu'ils ont une cinématique similaire, de même les leptons HF provenant de quarks b et c appartiennent à la même catégorie. La catégorie appelée light contient en réalité un autre type de

processus que celui mentionné plus haut, il s'agit des leptons émis lors de désintégrations semi-leptoniques de jet légers, c'est-à-dire tous les types de jets sauf les jets de quarks b et c . Ces leptons sont différents des jets reconstruits en tant qu'électrons, mais ils ont une cinématique proche et il s'agit d'une composante négligeable, c'est pourquoi ils ne sont pas traités séparément.

La composition du bruit de fond simulé dans les régions de signal du canal à deux leptons de même charge est montrée sur les figures 6.2 et 6.3 en fonction de H_T , et est résumée sur la figure 6.4. Ce sont les processus $t\bar{t}W$, $t\bar{t}Z$, et $t\bar{t}H$ qui dominent le bruit de fond dans les régions de signal. Ils représentent environ 60 à 70 % du bruit de fond dans les régions avec $N_b = 2$ et $N_b = 3$, et environ 40 % du bruit de fond dans la région SR4b. La composante de mauvaise identification de la charge (« QmisID ») représente quant à elle entre 10 et 20 % du bruit de fond des régions de signal. Puis les composantes conversion (« CO+ γ^* ») et HF représentent environ 10 % du bruit de fond, le taux variant légèrement d'une région à l'autre. Dans la région SR4b la composante de conversion représente notamment plus de 20 % du total. Finalement les autres composantes du bruit de fond représentent chacune moins de 5 % du bruit de fond total.

6.2.2 Définition des régions de contrôle

Parmi les quatre catégories d'événements définies auparavant, deux d'entre elles se voient affecter des facteurs de normalisation, il s'agit des catégories conversion et HF. Ce sont les composantes principales du bruit de fond dû aux faux leptons. La composante associée à la mauvaise identification de la charge des électrons ne se voit pas affecter de facteur de normalisation puisqu'elle est estimée par une méthode dédiée. Pour le moment la prédiction des simulations Monte-Carlo est utilisée pour ce bruit de fond. Une configuration avec un facteur de normalisation a été testée mais fait apparaître plusieurs corrélations entre les paramètres ajustés. La composante light est très faible dans les régions de signal, c'est pourquoi un facteur de normalisation supplémentaire n'est pas nécessaire. La catégorie autre est également laissée telle quelle puisqu'elle est composée d'événements non classés et potentiellement hétérogènes. Elle est faible dans toutes les régions.

Comme expliqué auparavant le processus $t\bar{t}W$ a été mesuré comme ayant une section efficace différente de sa valeur théorique, c'est pourquoi la normalisation de ce processus est également ajustée grâce aux données grâce à un facteur de normalisation dédié. Il y a donc un total de quatre facteurs de normalisation dans le modèle statistique :

- la force du signal μ , comme dans le modèle statistique présenté dans le chapitre 5 ;
- un facteur de normalisation pour le processus $t\bar{t}W$;
- un facteur de normalisation pour les événements conversion ;
- un facteur de normalisation pour les événements HF.

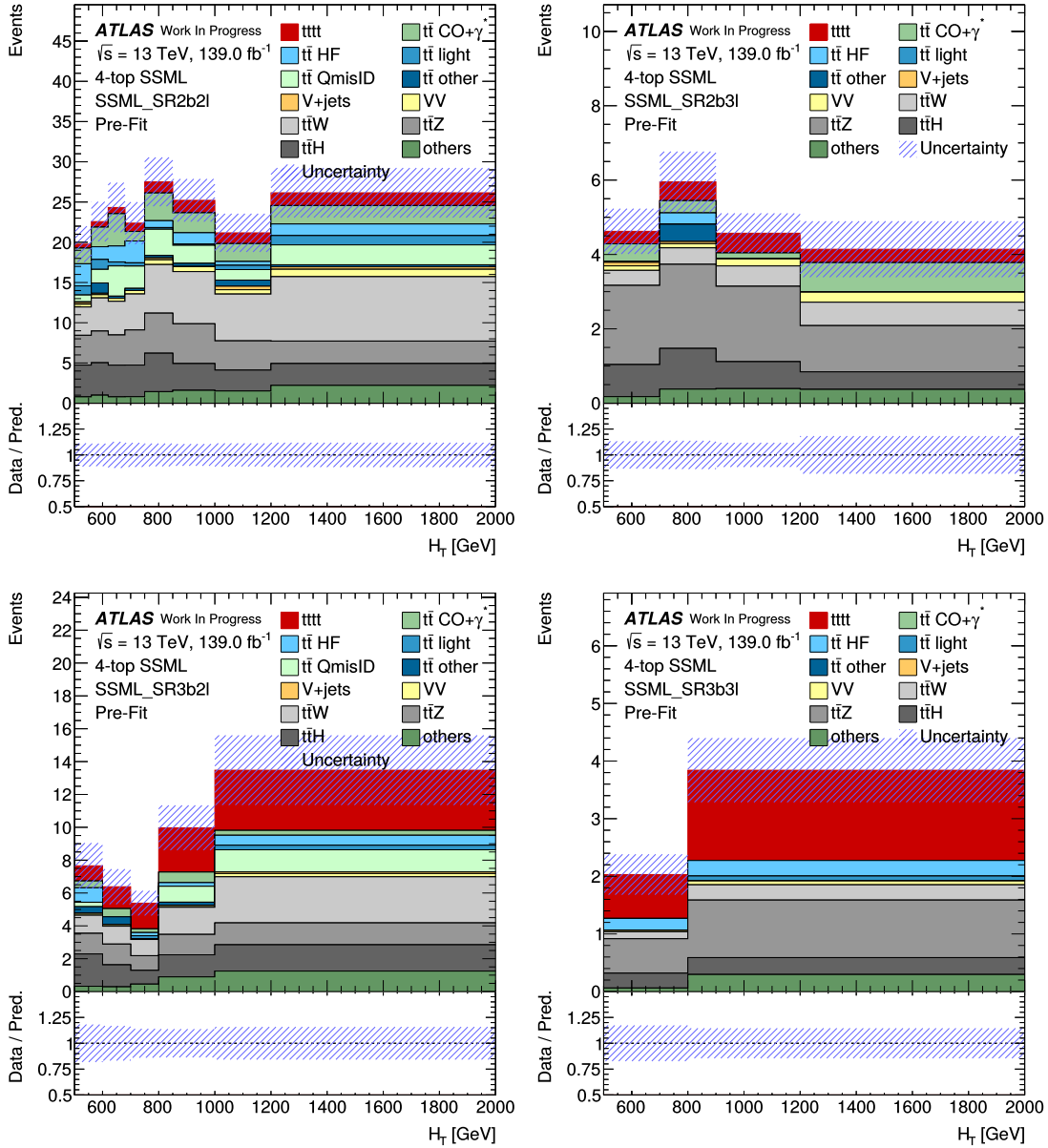


Figure 6.2 – Nombre d'événements du bruit de fond simulé dans les régions de signal à deux leptons de même charge (gauche) et trois leptons (droite). Les régions à deux jets de quarks b sont en haut et celles à trois jets de quarks b en bas. Le signal est représenté en rouge, et les autres couleurs représentent les différents types de bruits de fond.

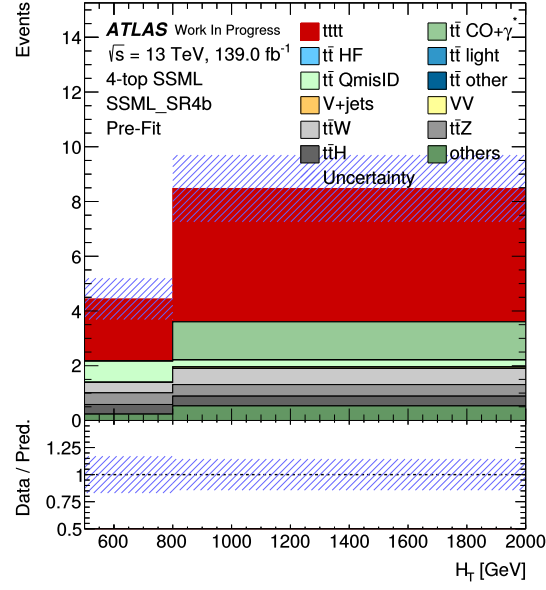


Figure 6.3 – Nombre d’événements du bruit de fond simulé dans la région de signal SR4b. Le signal est représenté en rouge, et les autres couleurs représentent les différents types de bruits de fond.

Le tableau 6.3 présente la liste des régions de contrôle définies pour la méthode des templates, ainsi que les coupures associées. La région de contrôle ciblant le processus $t\bar{t}W$ est appelée $CR_{t\bar{t}W}$. Elle est définie à partir des événements avec deux leptons de même charge et au moins deux jets de quarks b , mais un faible nombre de jets : $2 \leq N_j < 6$. Des coupures additionnelles assurent l’orthogonalité avec les autres régions comme indiqué dans le tableau 6.3. Durant l’ajustement du modèle statistique, c’est la distribution de la grandeur N_b qui est ajustée dans la région $CR_{t\bar{t}W}$ puisqu’on s’attend à des difficultés de modélisation associées au nombre de jets de quarks b additionnels.

La région $CR_{t\bar{t}CO}$ cible les conversions en utilisant la grandeur de rayon de conversion R_{CO} calculée à partir de la reconstruction d’un vertex de conversion. Les événements avec deux leptons de même charge et au moins un électron c’est-à-dire les catégories ee et $e\mu$ sont sélectionnés. Ce rayon est en général non défini en l’absence de conversion, ce qui est représenté par une valeur inférieure à -100 mm. Elle permet ainsi une bonne discrimination entre les conversions et les autres processus. La coupure $-100 \text{ mm} < R_{CO} < 50 \text{ mm}$ a été choisie après observation de la distribution de la grandeur R_{CO} . Les événements sont sélectionnés s’ils ont au moins un jet de quark b afin de maximiser le nombre d’événements disponibles. D’autres coupures sont ajoutées afin d’assurer l’orthogonalité entre les régions, comme indiqué dans le tableau 6.3. Dans cette région, c’est la distribution de la grandeur M_{CO}^{ee} qui est ajustée. Il s’agit de la masse des deux traces de conversion reconstruite

Nom	Catégorie(s)	N_b	N_j, H_T [GeV], R_{CO} [mm]
CR $t\bar{t}W$	$\ell^\pm\ell^\pm$	≥ 2	$2 \leq N_j < 6$ $H_T > 500$ pour $N_b = 2$ $H_T < 500$ pour $N_b \geq 3$
CR $t\bar{t}CO$	ee et $e\mu$	≥ 1	$2 \leq N_j < 6$ $200 < H_T < 500$ $-100 < R_{CO} < 50$
CR1 $b3\ell$	$\ell\ell\ell$	1	$200 < H_T < 500$
CR $t\bar{t}Z$	$\ell\ell\ell$, pic du Z	≥ 2	$H_T > 500$

Table 6.3 – Définition des régions de contrôle pour la méthode des templates. La catégorie $\ell^\pm\ell^\pm$ désigne les événements avec deux leptons de même charge tandis que la catégorie $\ell\ell\ell$ désigne les événements avec trois leptons. Les catégories ee et $e\mu$ désignent les événements avec deux leptons de même charge et au moins un électron. L’expression « pic du Z » signifie que la condition de rejet du boson Z décrite dans la section 6.1.2 est inversée.

au niveau du vertex primaire. On peut remarquer que la région CR $t\bar{t}CO$ n’est pas parfaitement orthogonale à la région CR $t\bar{t}W$. Ceci sera modifié dans une prochaine version de l’analyse.

La région CR1 $b3\ell$ sélectionne les événements avec trois leptons et un seul jet de quark b , pour de valeurs modérées d’énergie $200 < H_T < 500$. Elle cible le bruit de fond HF, mais contient également une grande proportion de bruit de fond $t\bar{t}W$. La distribution de la grandeur H_T est ajustée dans cette région.

Finalement une région de contrôle est définie pour le bruit de fond $t\bar{t}Z$. En effet, même si la normalisation de ce bruit de fond n’est pas ajustée comme c’est le cas pour $t\bar{t}W$ puisqu’elle est supposée correcte, il est nécessaire de contraindre les incertitudes systématiques associées à ce bruit de fond en particulier dans la région comportant quatre jets de quarks b . La région CR $t\bar{t}Z$ est construite en sélectionnant les événements qui ne respectent pas le critère de rejet du boson Z décrit dans la section 6.1.2, et en appliquant une sélection proche des régions de signal : $N_b \geq 2$, $H_T > 500$ GeV. C’est la distribution du nombre d’événements en fonction du nombre de jets de quarks b qui est ajustée puisqu’on s’attend à des difficultés de modélisation croissants avec la grandeur N_b .

Nom	Catégorie	N_b	N_j et H_T [GeV]
VR1b2 ℓ	$\ell^\pm \ell^\pm$	1	$N_j \geq 6$ et $H_T > 500$
VR2b2 ℓ	$\ell^\pm \ell^\pm$	2	$N_j \geq 6$ et $200 < H_T < 500$
VR3b2 ℓ	$\ell^\pm \ell^\pm$	≥ 3	$N_j \geq 6$ et $200 < H_T < 500$
VR1b3 ℓ	$\ell\ell\ell$	1	$N_j \geq 6$ et $H_T > 500$
VR2b3 ℓ	$\ell\ell\ell$	≥ 2	$N_j \geq 6$ ou $200 < H_T < 500$

Table 6.4 – Définition des régions de validation pour la méthode des templates. La catégorie $\ell^\pm \ell^\pm$ désigne les événements avec deux leptons de même charge tandis que la catégorie $\ell\ell\ell$ désigne les événements avec trois leptons.

6.2.3 Définition des régions de validation

Afin de vérifier les facteurs de normalisation ajustés par la méthode des templates, des régions de validation sont définies, et l'accord entre données et simulation avant et après l'ajustement du modèle statistique est comparé. Les régions de validation sont définies de manière à être proches des régions de signal mais pauvres en signal. Elles sont définies similairement aux régions de signal suivant leur nombre de jets de quarks b et le canal, c'est-à-dire deux leptons de même charge ou trois leptons. Leur définition est résumée dans le tableau 6.4. La composition du bruit de fond dans ces régions, ainsi que dans les autres régions de l'analyse, est montrée sur la figure 6.4.

6.3 Résultats préliminaires

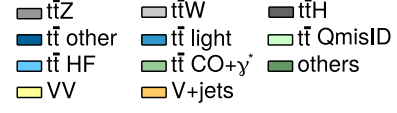
Cette section décrit les résultats préliminaires obtenus dans le canal à deux leptons de même charge. La liste des incertitudes systématiques est présentée, et les résultats de la configuration la plus simple avec seulement les cinq régions de signal sont donnés. Seule la simulation est présente dans les régions de signal qui sont opérées en aveugle à ce stade de l'analyse. Enfin, les résultats obtenus avec la méthode des templates sont décrits.

6.3.1 Incertitudes systématiques et premiers résultats

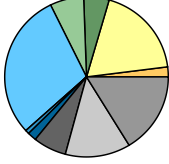
Cette étude préliminaire inclut la plupart des incertitudes systématiques qui ont été décrites dans la section 5.5. Les incertitudes sur la reconstruction des objets sont présentes sauf celles associées à l'échelle d'énergie des jets. L'incertitude sur la luminosité est de 1,7 %, estimée via la calibration de l'échelle de luminosité utilisant des scans Van de Meer, similairement à la méthode présentée dans la référence [9], ainsi que grâce au

ATLAS $\sqrt{s} = 13$ TeV
4-top SSML

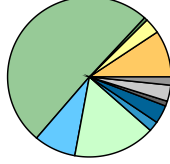
Work In Progress



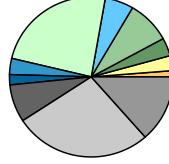
CR1b3l



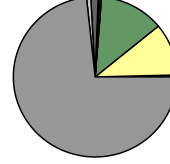
CRttbarCO



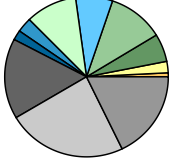
CRttW



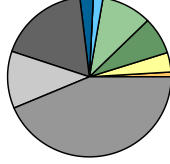
CRttZ



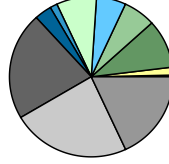
SR2b2l



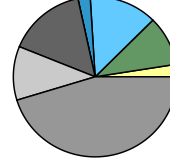
SR2b3l



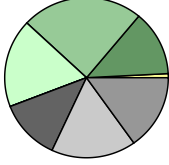
SR3b2l



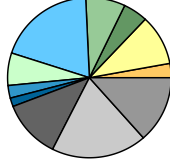
SR3b3l



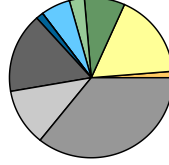
SR4b



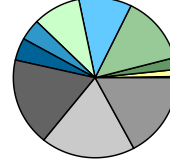
VR1b2l



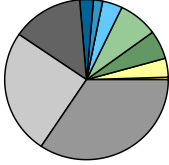
VR1b3l



VR2b2l



VR2b3l



VR3b2l

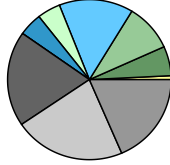


Figure 6.4 – Composition du bruit de fond des différentes régions de contrôle, de validation, et de signal.

détecteur LUCID-2 [10]. Elle est appliquée de manière corrélée à toutes les régions de signal.

Des incertitudes sur les bruits de fond irréductibles sont introduites, sauf pour les composantes dont la normalisation est ajustée. Elles correspondent aux incertitudes sur la section efficace utilisée pour leur normalisation. Des valeurs approximatives sont utilisées pour le moment :

- 10 % pour les processus VV et $t\bar{t}H$, V désignant W ou Z ;
- 15 % pour $t\bar{t}Z$;
- 30 % pour la production d'un boson vecteur accompagné de jets, noté $V + \text{jets}$;
- 30 % pour les composantes du bruit de fond dû aux faux leptons (processus $t\bar{t}$, tq , tW) qui n'ont pas de facteur de normalisation : light, autre, et mauvaise identification de la charge ;
- 50 % pour les autres bruits de fond.

Enfin des incertitudes additionnelles sont appliquées sur les processus $t\bar{t}W$, $t\bar{t}Z$, et $t\bar{t}H$ afin de prendre en compte les incertitudes de la modélisation Monte-Carlo aux grandes valeurs de N_b :

- 10 % pour $N_b = 2$;
- 30 % pour $N_b = 3$;
- 50 % pour $N_b \geq 4$.

Ces incertitudes sont très préliminaires et n'ont pas encore fait l'objet d'une étude poussée, elles sont donc susceptibles d'être modifiées.

La configuration la plus simple est celle avec seulement les cinq régions de signal opérées en aveugle. La signification statistique obtenue sur le signal de production de quatre quarks top dans cette configuration est de $2,4\sigma$ en utilisant des pseudo-données générées à partir des simulations Monte-Carlo. Les incertitudes sur les bruits de fond irréductibles sont toutefois très préliminaires donc ce résultat est susceptible de changer. La signification statistique obtenue sans les incertitudes systématiques est de $3,1\sigma$. Puisque l'ajustement est réalisé sur simulation uniquement, aucune anomalie n'est observée dans le modèle statistique.

6.3.2 Résultats avec la méthode des templates

On s'intéresse maintenant aux résultats obtenus en ajoutant les régions de contrôle ainsi que les facteurs de normalisation de la méthode des templates. Les données sont toujours absentes des régions de signal, mais elles sont incluses dans les autres régions. Les valeurs ajustées des cinq facteurs de normalisation sont représentées sur la figure 6.5. Les facteurs de normalisation sont tous proches de 1, même si celui de la composante conversion est éloigné de 1 d'un peu plus qu'une déviation standard.

Il existe des corrélations entre les facteurs de normalisation, mais aussi avec certains paramètres de nuisance, comme présenté sur la figure 6.6. Ces corrélations sont problé-

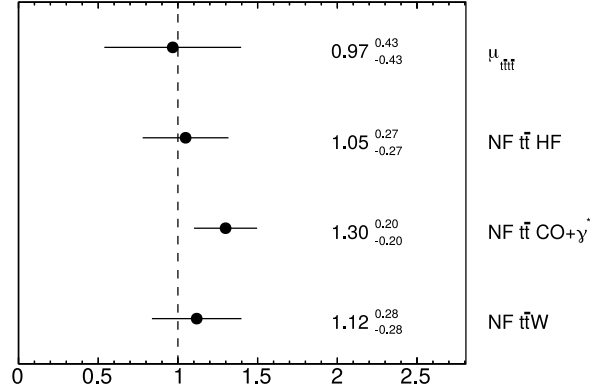


Figure 6.5 – Valeurs ajustées des facteurs de normalisation de la méthode des templates.

matiques car peuvent potentiellement déplacer les valeurs de certains paramètres sans contrainte physique réelle. Quelques études préliminaires utilisant de nouvelles variables dans les régions de contrôle parviennent à réduire significativement ces corrélations.

Les figures 6.7 et 6.8 représentent les distributions des variables d’ajustement dans les différentes régions de contrôle, tandis que les figures 6.9 et 6.10 représentent les distributions de la grandeur H_T dans les régions de validation. Les régions VR2b2 ℓ et VR3b2 ℓ ne sont pas représentées car il s’agit de deux distributions à un seul intervalle c’est-à-dire un simple nombre d’événements. Les parties gauche (droite) des figures correspondent aux distributions avant (après) l’ajustement du modèle statistique. La figure 6.11 montre les nombres d’événements dans les régions de contrôle (haut) et dans les régions de validation (bas), incluant les régions VR2b2 ℓ et VR3b2 ℓ .

On peut remarquer que l’ajustement se comporte bien dans les régions de contrôle, toutefois l’accord entre données et simulation est déjà bon avant l’ajustement, sauf dans la région de contrôle CRt $\bar{t}$ CO. En ce qui concerne les régions de validation, on n’observe pas d’amélioration significative de l’accord entre données et simulation avant et après l’ajustement, l’accord étant déjà plutôt bon avant ajustement, sauf dans la région VR2b3 ℓ .

Un autre résultat récent n’utilisant pas la méthode des templates mais simplement un ajustement de la grandeur H_T dans les régions de signal obtient une signification statistique d’environ $2,1\sigma$. L’ensemble des incertitudes systématiques sur les jets est inclus. L’effet des principaux paramètres de nuisance sur l’ajustement du modèle statistique est représenté sur la figure 6.12 à l’aide des rectangles bleus associés à chaque paramètre de nuisance, similairement à la figure 5.14.

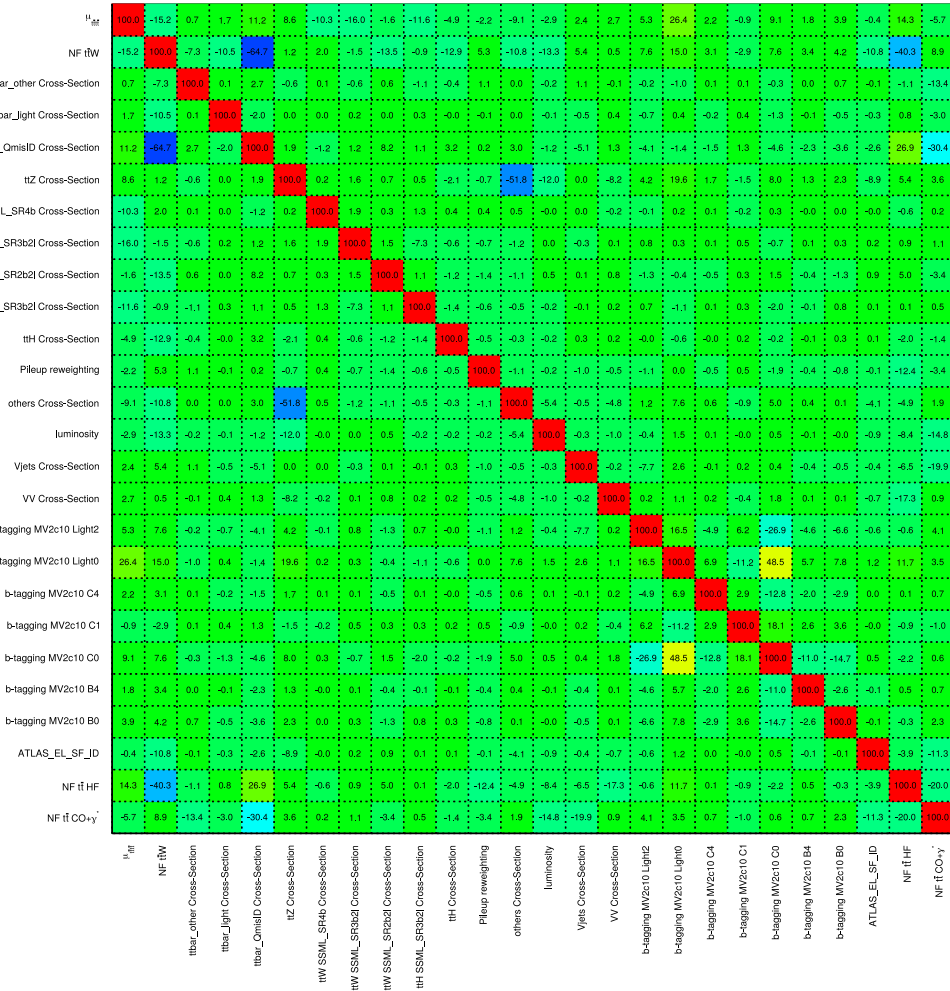


Figure 6.6 – Matrice de corrélation des paramètres de nuisance et des facteurs de normalisation. Seuls les paramètres présentant au moins une corrélation supérieure à 10 % sont représentés. La couleur rouge désigne une corrélation, la couleur bleue une anti-corrélation et la couleur verte une absence de corrélation.

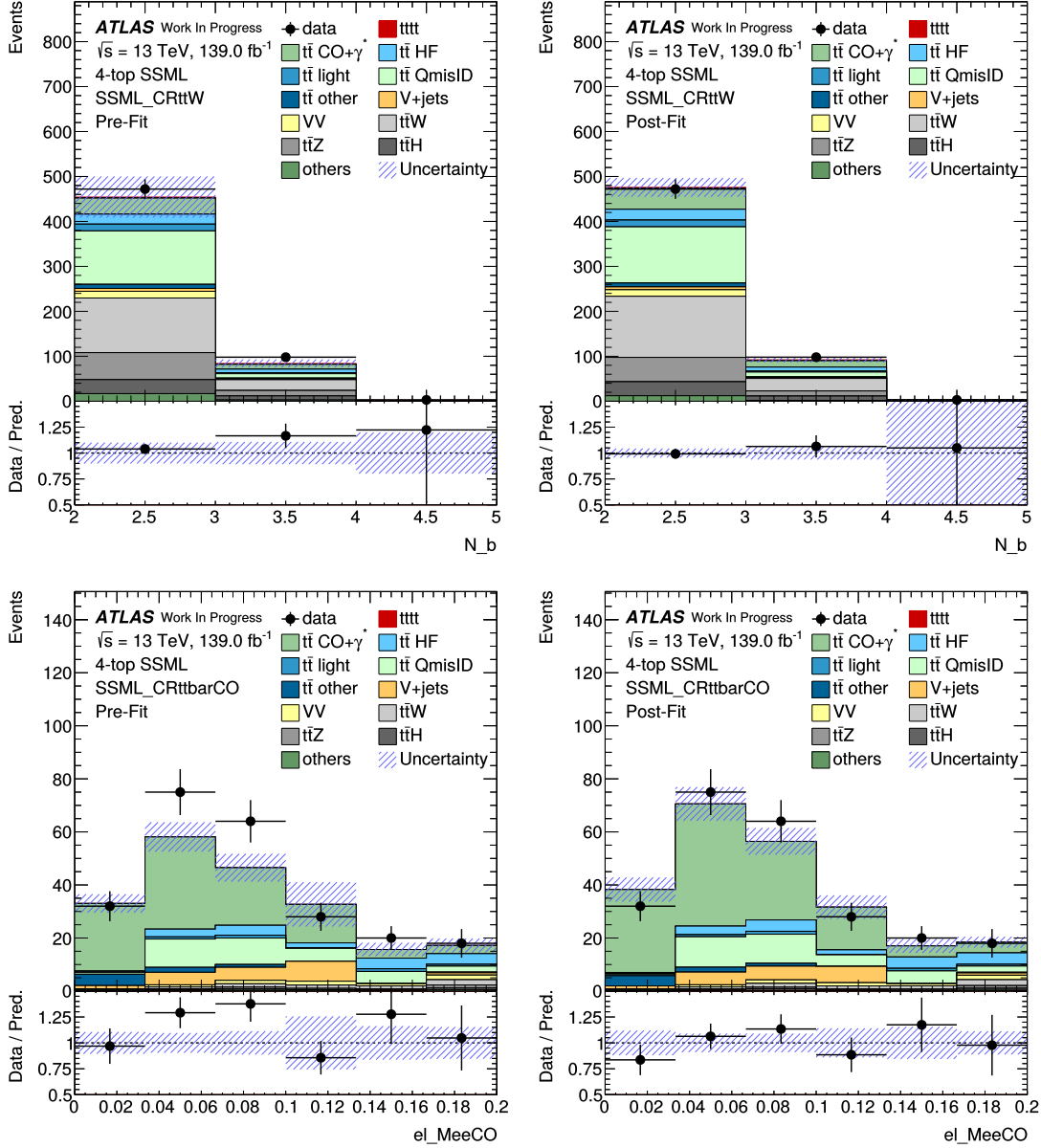


Figure 6.7 – Distributions des variables d'ajustement dans les régions de contrôle CR $t\bar{t}W$ (haut) et CR $t\bar{t}CO$ (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.

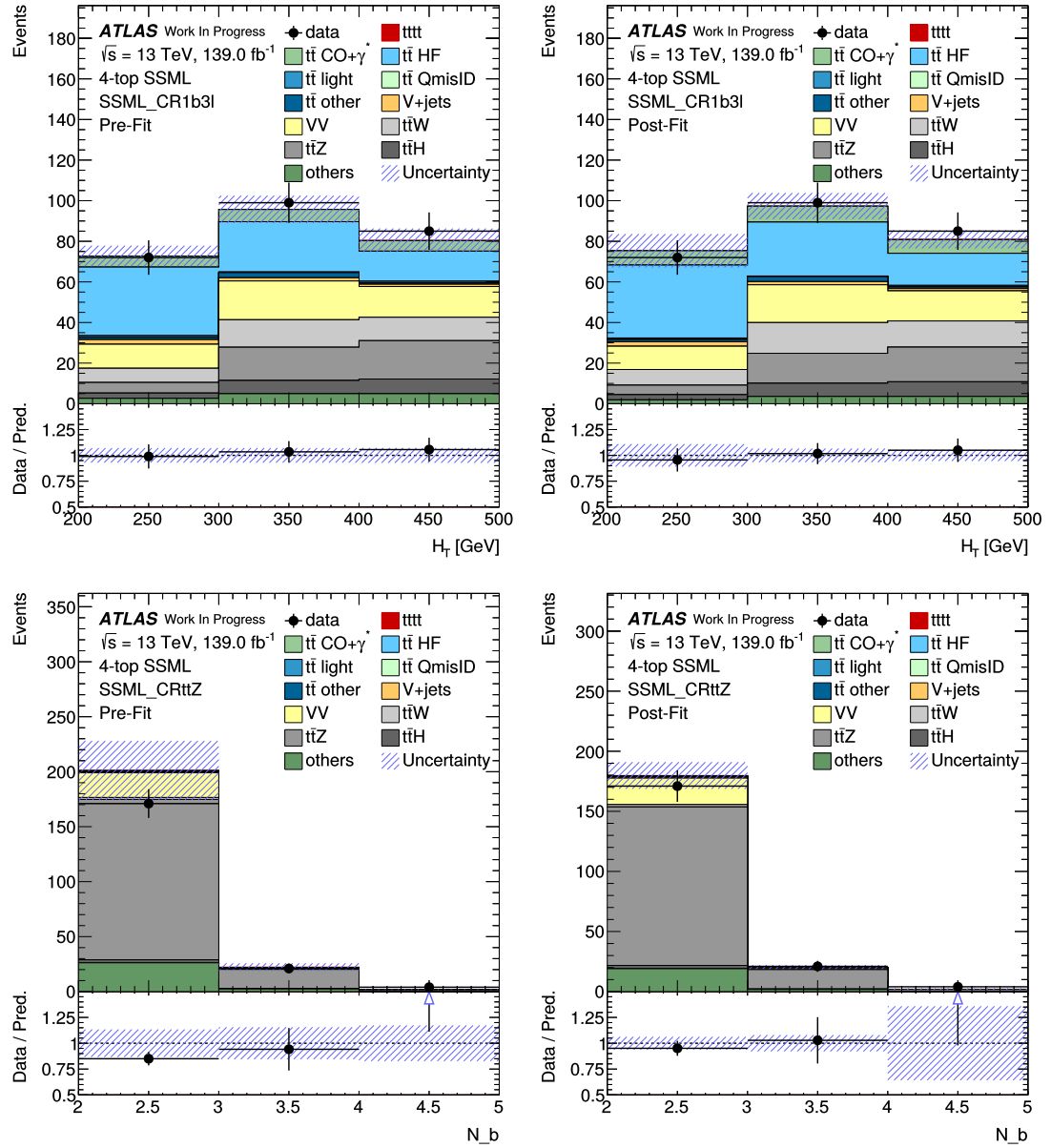


Figure 6.8 – Distributions des variables d'ajustement dans les régions de contrôle CR1b3l (haut) et CRtZ (bas), avant (gauche) et après (droite) l'ajustement du modèle statistique.

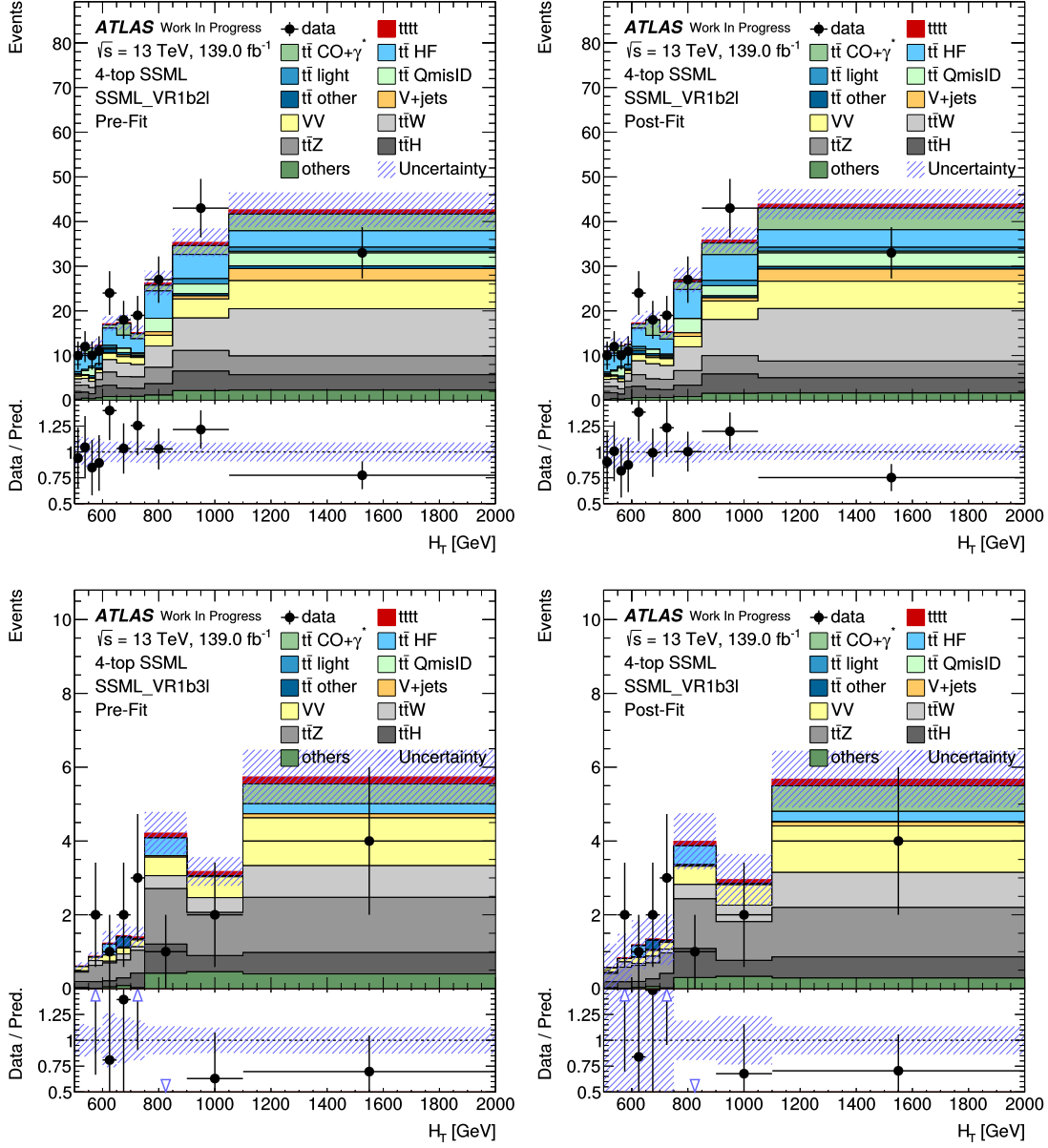


Figure 6.9 – Distributions de la grandeur H_T dans les régions de validation VR1b2 ℓ (haut) et VR1b3 ℓ (bas), avant (gauche) et après (droite) l’ajustement du modèle statistique.

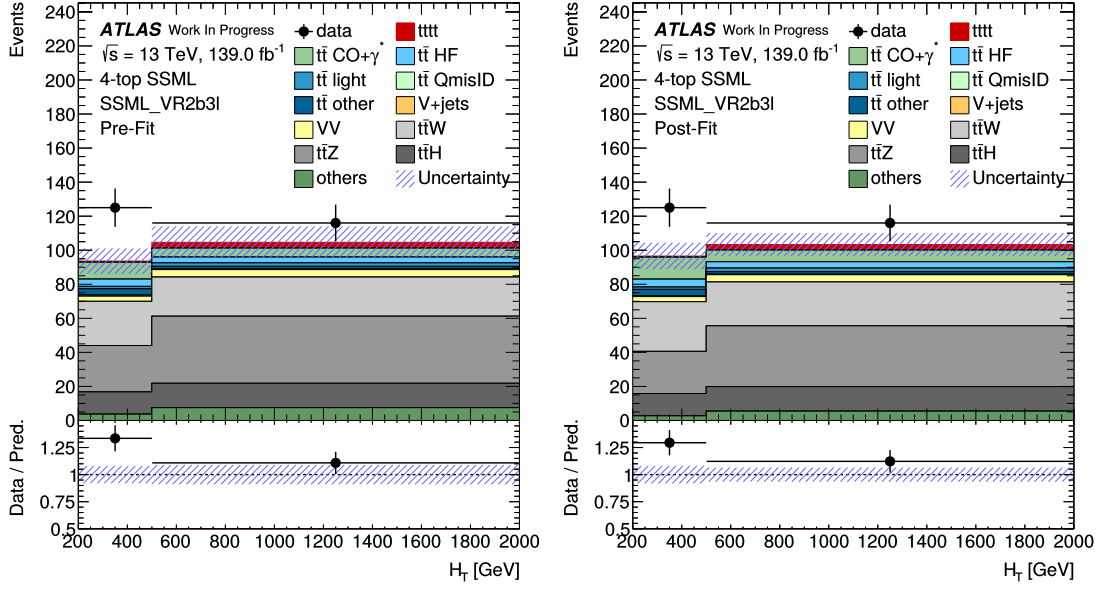


Figure 6.10 – Distribution de la grandeur H_T dans la région de validation VR2b3l, avant (gauche) et après (droite) l’ajustement du modèle statistique.

6.3.3 Perspectives

Le travail présenté dans cette section est très préliminaire, et de nombreuses pistes d’amélioration sont envisagées et sont en cours d’exploration. Parmi les plus évidentes on peut citer l’ajout des incertitudes sur l’échelle d’énergie des jets, puis d’incertitudes systématiques liées à l’extrapolation des résultats de la méthode des templates depuis les régions de contrôle vers les régions de signal. Cette estimation doit prendre en compte les incertitudes de modélisation du processus $t\bar{t}$, et ajouter des incertitudes sur les différentes catégories de faux leptons qui sont traitées séparément. Les incertitudes sur les processus $t\bar{t}W$ et $t\bar{t}Z$ sont également un point très important du modèle statistique. Elles doivent être estimées par une étude spécifique, ce qui n’est pas encore le cas.

Des incertitudes systématiques sur la modélisation du signal doivent également être ajoutées. Elles sont constituées d’incertitudes sur les fonctions de distribution de partons, sur les échelles de renormalisation et de factorisation, ainsi que sur la modélisation des gerbes partoniques et de l’hadronisation. La figure 6.13 montre une comparaison préliminaire de la modélisation au NLO du signal $t\bar{t}t\bar{t}$ avec le générateur MG5_AMC@NLO, couplé avec PYTHIA ou HERWIG pour la modélisation des gerbes partoniques. Les distributions des grandeurs N_b , N_j , H_T , ainsi que du p_T des jets sont présentées. Les événements contiennent deux leptons de charge opposée et vérifient $N_b \geq 2$ et $N_j \geq 2$. Cette comparaison entre générateurs Monte-Carlo permet d’apprécier les

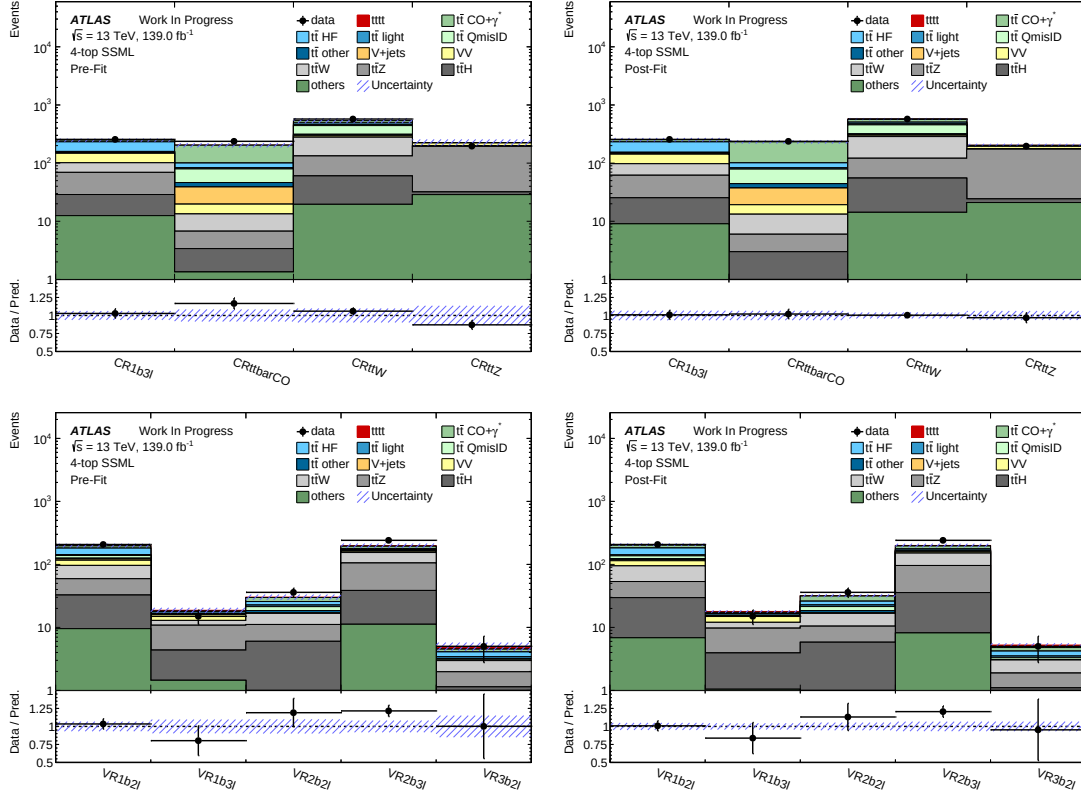


Figure 6.11 – Nombres d’événements dans les régions de contrôle (haut) et dans les régions de validation (bas), avant (gauche) et après (droite) l’ajustement du modèle statistique.

effets de la modélisation des gerbes partoniques et de l’hadronisation sur la cinématique du signal. Les deux générateurs fournissent des résultats similaires, sauf à haut nombre de jets et à haut H_T . La grandeur N_b est peu impactée. Des études similaires doivent être menées dans le canal à deux leptons de même charge.

Une autre vérification à effectuer est la comparaison des résultats entre la méthode des templates et la méthode de la matrice héritée de l’analyse utilisant $36,1 \text{ fb}^{-1}$, ainsi que la comparaison de leurs performances. La méthode de la matrice est cependant encore en développement et une telle comparaison n’est pas possible à ce stade. L’inclusion d’une estimation à partir de données du bruit de fond de mauvaise identification de la charge des électrons est également prévue, comme expliqué précédemment.

On peut également remarquer que les coupures sur la grandeur H_T dans la région CR_{ttW} changent suivant la valeur du paramètre N_b ce qui n’est peut-être pas idéal. Il est envisagé d’ajuster séparément les événements à $N_b = 2$ et à $N_b \geq 3$ si le nombre d’événements disponible le permet.

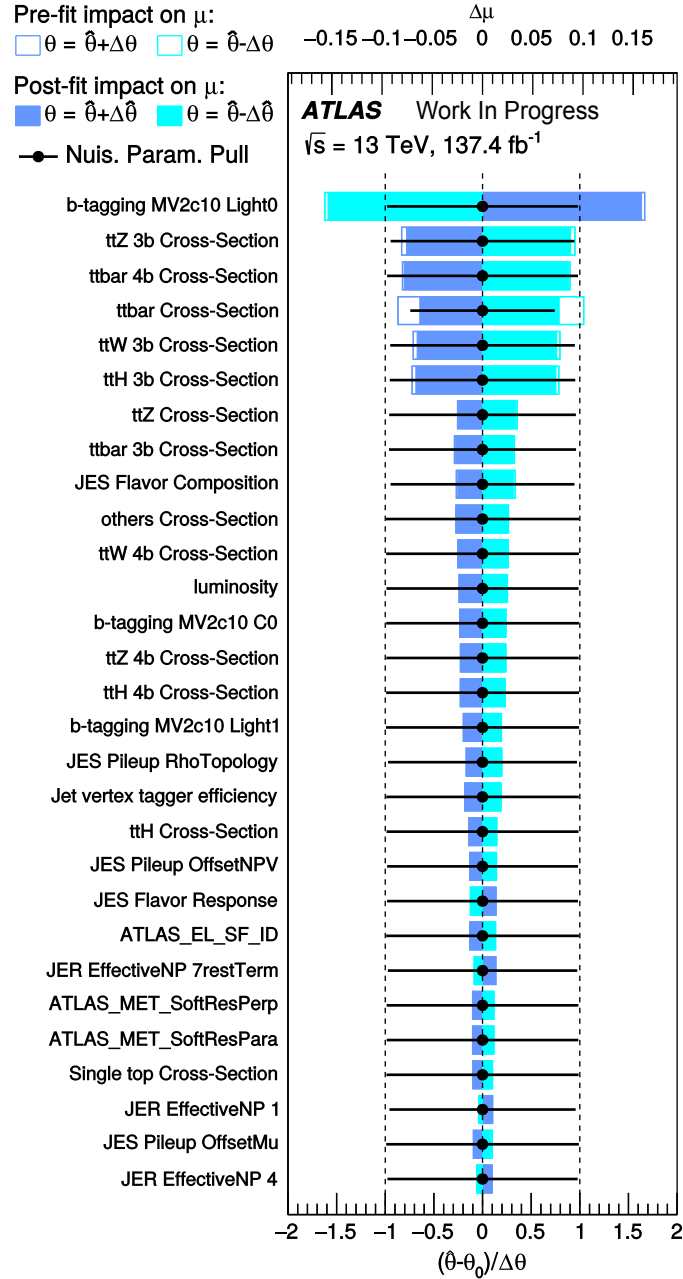


Figure 6.12 – Effet des paramètres de nuisance les plus influents sur la force du signal de production de quatre quarks tops dans le cadre du modèle standard. L’ensemble des incertitudes sur les jets est inclus, et l’estimation du bruit de fond est réalisée par des simulations Monte-Carlo uniquement. La figure est construite de la même façon que la figure 5.14.

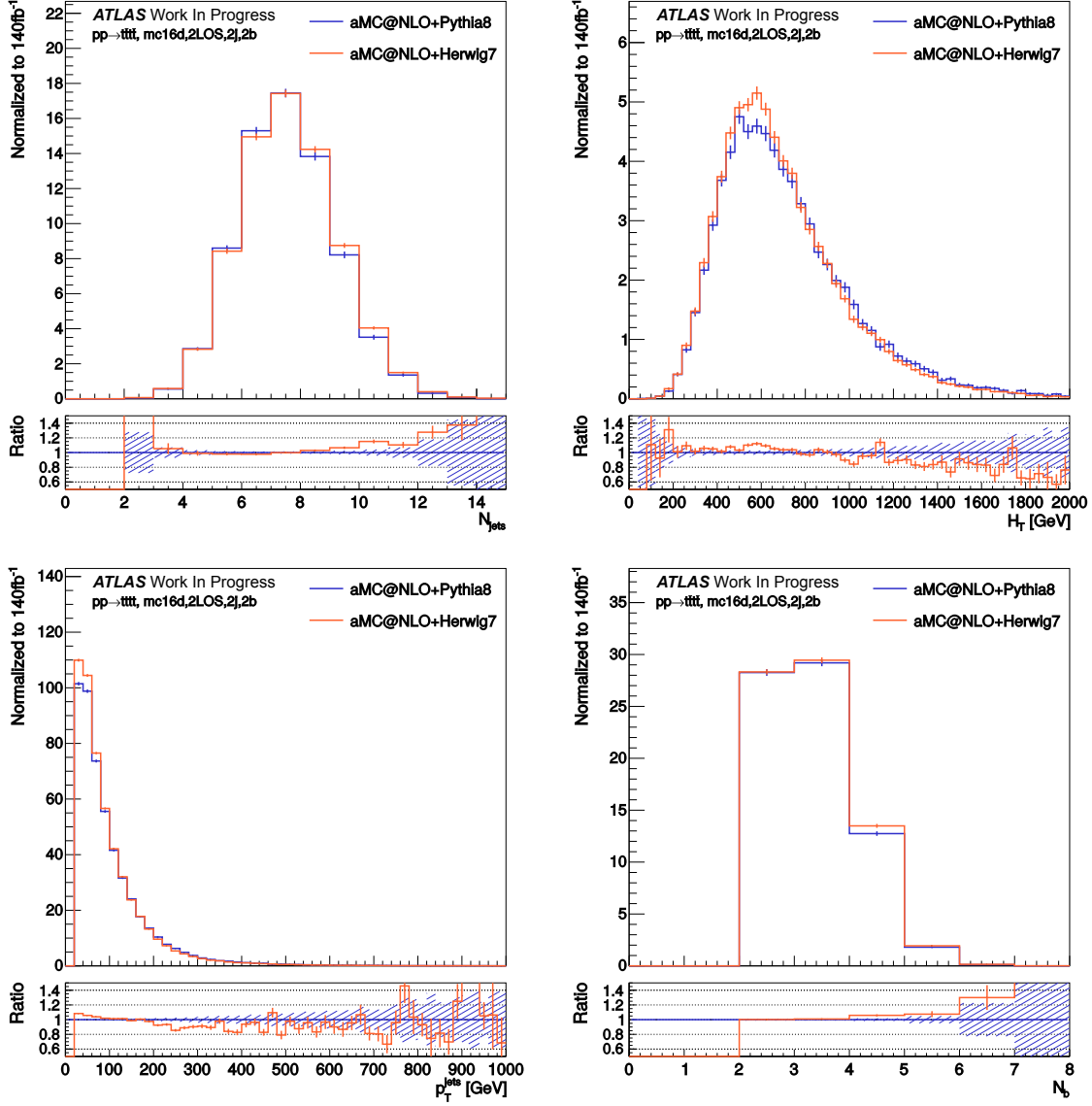


Figure 6.13 – Comparaison de la modélisation au NLO du signal $t\bar{t}t\bar{t}$ avec les générateurs MG5_AMC@NLO+PYTHIA et MG5_AMC@NLO+HERWIG. Les distributions des grandeurs N_b , N_j , H_T , ainsi que du p_T des jets sont présentées. Les événements contiennent deux leptons de charge opposée et vérifient $N_b \geq 2$ et $N_j \geq 2$.

Il est probablement nécessaire de repenser la définition des régions de contrôle, mais aussi des régions de validation. On peut par exemple penser au nombre d'intervalles pour les distributions de H_T dans chacune de ces régions. Les régions de signal peuvent également être repensées, par exemple en séparant les régions à deux leptons de même charge suivant les types de leptons : ee , $e\mu$, et $\mu\mu$.

Les grandeurs d'ajustement utilisées dans les régions de contrôle sont H_T et N_b sauf pour la région ciblant les conversions. Il serait intéressant d'utiliser d'autres grandeurs, et en particulier de trouver des grandeurs qui sont distribuées différemment pour les différentes catégories de faux leptons. Ainsi les contraintes apportées par l'ajustement de modèle statistique sur les catégories de faux leptons pourraient augmenter, et les corrélations entre les différents paramètres pourraient diminuer. Ces tests sont en cours de réalisation et des études préliminaires montrent que le p_T de l'électron ayant le p_T le plus élevé dans l'événement est une grandeur intéressante pour la région CR1b3 ℓ . Pour la région CR $t\bar{t}W$, la somme des charges des deux leptons semble être une grandeur prometteuse, puisque le boson W du processus $t\bar{t}W$ au LHC est plus souvent produit avec une charge positive que négative. Cette asymétrie de charge n'est pas présente pour les autres processus, et permet ainsi de mieux distinguer le processus $t\bar{t}W$ des autres. Une autre piste en cours d'exploration est de séparer la région de contrôle CR $t\bar{t}W$ en deux ou trois régions, selon la catégorie d'événement à deux leptons de même charge, c'est-à-dire ee , $e\mu$, ou $\mu\mu$. On peut s'attendre en effet à une composition de bruit de fond dû aux faux leptons différente. Enfin, une coupure $N_j \geq 6$ dans la région de CR $t\bar{t}Z$ permettrait de sélectionner des événements plus proches des régions de signal et donc de mieux contraindre le bruit de fond $t\bar{t}Z$, même s'il est nécessaire de vérifier que l'augmentation de l'incertitude statistique ne compense pas cet effet positif. Tous les points mentionnés dans ce paragraphe sont en cours d'implémentation et de test, et semblent fournir de meilleurs résultats avec une réduction significative des corrélations entre les paramètres. La signification statistique obtenue par la meilleure configuration actuelle est d'environ $2,7\sigma$.

6.4 Conclusion

La recherche de la production de quatre quarks top au LHC dans les données du Run 2 continue, après l'analyse des données prises en 2015 et en 2016 présentée dans le chapitre 5. Le jeu de données utilisé est désormais l'intégralité du Run 2 qui s'est terminé en 2018. Il est presque quatre fois plus grand et représente environ 139 fb^{-1} à une énergie de 13 TeV dans le centre de masse.

Cette nouvelle analyse se concentre exclusivement sur le processus de production de quatre quarks top conforme au modèle standard, et a pour objectif une mise en évidence du processus, elle vise une signification statistique supérieure à 3σ . Pour cela l'analyse

est réalisée en synergie dans deux canaux : le canal à un lepton et à deux leptons de charge opposée, ainsi que le canal avec au moins deux leptons ayant la même charge électrique. Ces deux canaux présentent un niveau et une composition de bruit de fond très différents, mais suivent une stratégie similaire puisque le signal se situe aux grandes valeurs de nombre de jets, nombre de jets de quarks b , et énergie totale.

Cinq régions de signal préliminaires ont été définies dans le canal à deux leptons de même charge, et plusieurs configurations de régions de contrôle et de régions de validation sont explorées. Une nouvelle méthode appelée méthode des templates est étudiée afin d'estimer le bruit de fond dû aux faux leptons. Elle consiste à décrire la cinématique des différentes catégories de faux leptons avec des simulations Monte-Carlo, puis à ajuster la normalisation des catégories grâce aux données. Quatre régions de contrôle permettant de contraindre les facteurs de normalisation ont été définies et les résultats obtenus sont cohérents et prometteurs. Cependant l'implémentation de cette méthode est encore préliminaire et un travail approfondi sur la définition des régions et les incertitudes est encore nécessaire. Il faut également attendre que certains points de l'analyse soient plus avancés afin de rendre la méthode plus réaliste, par exemple grâce à une estimation à partir des données du bruit de fond de mauvaise identification de la charge des électrons.

À terme l'analyse va se doter d'un discriminant multivarié afin d'améliorer sa sensibilité au signal. Un résultat récent de la collaboration CMS dans le canal à deux leptons de même charge qui utilise l'intégralité des données du Run 2 obtient une signification statistique observée (attendue) de $2,6$ ($2,7\sigma$) [1]. C'est pourquoi les bons résultats préliminaires de l'analyse présentée dans ce chapitre sont encourageants.

La production de quatre quarks top sera un processus intéressant à étudier au LHC dans les prochaines années. Au Run 3 et au HL-LHC, la section efficace du processus est estimée à $15,8^{+18}_{-21}\%$ fb pour une énergie dans le centre de masse de 14 TeV [3]. Une étude prospective de la collaboration ATLAS [11] prédit une signification statistique de 5σ atteinte avec 300 fb^{-1} c'est-à-dire durant le Run 3. Au HL-LHC une mesure utilisant 3000 fb^{-1} de luminosité intégrée prévoit d'atteindre une précision d'environ 11 % sur la section efficace $\sigma_{t\bar{t}t\bar{t}}$, et une signification statistique bien au-delà 5σ . Il sera également intéressant de se tourner vers les interprétations en termes de théorie des champs effectifs, ou EFT, pour la recherche de nouvelle physique grâce au processus à quatre tops. Avec les données du Run 3 et surtout du HL-LHC, des mesures différentielles de section efficace pourront être effectuées, par exemple en fonction du p_T le plus élevé parmi les quarks top dans l'événement, afin de contraindre les paramètres EFT davantage qu'avec une mesure de section efficace inclusive.

Références

- [1] CMS COLLABORATION. *Search for standard model production of four top quarks in final states with same-sign and multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*. Rapp. tech. CMS-PAS-TOP-18-003. Geneva : CERN, 2019. URL : <https://cds.cern.ch/record/2668710>.
- [2] ATLAS COLLABORATION. “Search for new phenomena in events with same-charge leptons and b-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Journal of High Energy Physics* 2018.12 (déc. 2018). DOI : 10.1007/jhep12(2018)039.
- [3] Rikkert FREDERIX, Davide PAGANI et Marco ZARO. “Large NLO corrections in $t\bar{t}W$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions”. In : *Journal of High Energy Physics* 2018.2 (fév. 2018). DOI : 10.1007/jhep02(2018)031.
- [4] ATLAS COLLABORATION. *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*. Rapp. tech. ATL-PHYS-PUB-2017-013. Geneva : CERN, juil. 2017. URL : <https://cds.cern.ch/record/2273281>.
- [5] J. ALWALL et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In : *Journal of High Energy Physics* 2014.7 (juil. 2014). DOI : 10.1007/jhep07(2014)079.
- [6] Simone ALIOLI et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX”. In : *Journal of High Energy Physics* 2010.6 (juin 2010). DOI : 10.1007/jhep06(2010)043.
- [7] Torbjörn SJÖSTRAND et al. “An introduction to PYTHIA 8.2”. In : *Computer Physics Communications* 191 (juin 2015), p. 159-177. DOI : 10.1016/j.cpc.2015.01.024.
- [8] ATLAS COLLABORATION. “Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In : *Physical Review D* 99.7 (avr. 2019). DOI : 10.1103/physrevd.99.072009.
- [9] ATLAS COLLABORATION. “Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC”. In : *The European Physical Journal C* 76.12 (nov. 2016). DOI : 10.1140/epjc/s10052-016-4466-1.
- [10] G. AVONI et al. “The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS”. In : *Journal of Instrumentation* 13.07 (juil. 2018), P07017-P07017. DOI : 10.1088/1748-0221/13/07/p07017.

- [11] ATLAS COLLABORATION. *HL-LHC prospects for the measurement of the Standard Model four-top-quark production cross-section*. Rapp. tech. ATL-PHYS-PUB-2018-047. Geneva : CERN, déc. 2018. URL : <https://cds.cern.ch/record/2651870>.

Conclusion

Le premier aspect étudié dans cette thèse concerne le projet *New Small Wheel*. De grands détecteurs Micromegas sont fabriqués à l'Institut de Recherche sur les lois Fondamentales de l'Univers. Une partie de leur contrôle qualité est effectuée sur un banc de test utilisant les muons cosmiques. Un code d'analyse des données issues du banc de test a été développé, partant des données brutes pour aller jusqu'aux informations de caractérisation des modules : pistes problématiques, gain, efficacité. Les deux premiers modules ont été étudiés en détail afin d'améliorer les méthodes de caractérisation et de permettre aux équipes de production d'améliorer leur procédés. Le code d'analyse est désormais utilisé pour caractériser les nouveaux modules qui seront produits dans les mois à venir. Il reste une trentaine de modules à fabriquer avant de pouvoir remplacer les petites roues du détecteur ATLAS.

Dans un deuxième temps, le processus de production de quatre quarks top a été étudié. Il s'agit d'un processus rare qui est sensible à de nombreux modèles de nouvelle physique. La mesure de sa section efficace pourrait également permettre de contraindre le couplage de Yukawa du quark top. La recherche de ce processus a été effectuée dans le canal à deux leptons de même charge électrique, qui permet la meilleure sensibilité. Ce travail de thèse s'est principalement concentré sur l'analyse du processus à quatre tops utilisant les données prises en 2015 et en 2016 par le détecteur ATLAS, publiée en 2018 dans *Journal of High Energy Physics*. D'abord des régions de signal ont été définies et l'optimisation de leurs coupures a été étudiée. Ces régions ont été optimisées pour l'étude simultanée de plusieurs processus physiques. Des régions de signal dédiées au processus à quatre tops ont également été définies, ce qui a permis d'augmenter très significativement la sensibilité de l'analyse à ce processus. Les incertitudes sur les bruits de fond irréductibles majeurs ont été estimées, puis le dispositif d'analyse statistique a été mis en place. L'analyse statistique a enfin été menée afin de produire les résultats finaux. Une combinaison avec l'analyse à un lepton a été effectuée et a conduit à la meilleure limite sur le processus à quatre tops au moment de sa publication.

Une nouvelle analyse dédiée au processus à quatre tops et utilisant toutes les données du Run 2 est en cours. L'objectif est d'atteindre une signification statistique de 3σ sur le processus. Après des études d'optimisation dans le canal à deux leptons de même charge,

ce travail de thèse s'est tourné vers une nouvelle méthode d'estimation du bruit de fond dû aux faux leptons. Ce bruit de fond est particulièrement difficile à modéliser, et la méthode dite du *template fit* pourrait fournir une meilleure modélisation que la méthode usuelle dite de la matrice. La modélisation est pour le moment préliminaire, et plusieurs pistes d'amélioration sont poursuivies. Un calcul d'incertitudes systématiques doit également être effectué. Les premiers résultats sont cohérents et encourageants. Finalement l'analyse va se doter d'un discriminant multivarié afin d'améliorer sa sensibilité. Les résultats préliminaires ainsi que les résultats récents provenant de la collaboration CMS amènent à penser que cette analyse remplira ses objectifs.

Titre : Caractérisation de grands détecteurs Micromegas pour le projet New Small Wheel et recherche de la production de quatre quarks top avec le détecteur ATLAS au LHC

Mots clés : Analyse de données, Physique des particules, LHC, ATLAS, Quark top, Nouvelle physique

Résumé : Ce travail de thèse a été effectué au sein de la collaboration ATLAS au LHC. Il s'intéresse dans un premier temps au développement de nouveaux détecteurs Micromegas pour le projet NSW, et dans un second temps à la recherche d'un processus physique très rare : la production de quatre quarks top au LHC, à une énergie dans le centre de masse de 13 TeV.

Le programme de physique du LHC prévoit une augmentation significative de la luminosité de l'accélérateur dans les années à venir, en particulier pour la phase haute-luminosité dont le démarrage est prévu en 2026. Certains composants du détecteur ATLAS doivent être mis à jour, notamment les petites roues du spectromètre à muons qui vont devoir supporter une irradiation bien supérieure au niveau actuel. De nouvelles petites roues (NSW) sont en cours de développement et une partie des détecteurs Micromegas pour ces petites roues est construite par l'Irfu. Les performances de ces détecteurs sont validées sur un banc de test utilisant des muons cosmiques.

Dans cette thèse, l'objectif a été de développer un code d'analyse des données de ce nouveau banc de test afin de caractériser les détecteurs Micromegas pour le projet NSW. Le code fournit notamment des informations de gain et d'efficacité de détection, et sera utilisé jusqu'à la fin du projet. Les efficacités mesurées sont globalement en accord avec les attentes sauf sur certaines parties des détecteurs qui doivent être investiguées.

Le quark top est un élément central de plusieurs théories

au-delà du modèle standard, comme par exemple les modèles avec des dimensions supplémentaires, ou encore qui incluent une modification du secteur du boson de Higgs. Certains de ces modèles prédisent une augmentation de la section efficace de production de quatre quarks top qui est un phénomène très rare dans le modèle standard et dont la section efficace ne vaut qu'environ 12 fb à 13 TeV.

L'étude du processus à quatre tops dans le canal à deux leptons de même charge électrique est particulièrement intéressant car le niveau de bruit de fond y est très faible. Ce travail de thèse s'intéresse d'abord à l'analyse des événements à quatre tops avec le détecteur ATLAS, portant sur les données prises en 2015 et 2016. L'étude de ce processus est un défi de par le faible nombre d'événements disponibles, mais aussi à cause du bruit de fond dû aux faux leptons qui est difficile à estimer. Plusieurs limites sont posées sur des modèles de nouvelle physique, et sur la production conforme au modèle standard. Cette dernière limite est combinée avec un résultat provenant du canal à un lepton pour fournir la meilleure limite au moment de sa publication.

Le processus de production de quatre quarks top conforme au modèle standard est à nouveau étudié avec l'entièreté des données prises entre 2015 et 2018, dans le but d'en effectuer une première observation. L'analyse de données est en cours, et cette thèse se concentre sur le développement d'une nouvelle méthode d'estimation du bruit de fond dû aux faux leptons.

Title : Characterization of large Micromegas detectors for the New Small Wheel project and search for four top quark production with the ATLAS detector at the LHC

Keywords : Data analysis, Particle physics, LHC, ATLAS, Top quark, New physics

Abstract : This doctoral work was carried out as a member of the ATLAS collaboration at the LHC, an accelerator with a center-of-mass energy of 13 TeV. The development of new Micromegas detectors for the NSW project has been studied, along with the search for a very rare physics process : the production of four top quarks at the LHC, at a center-of-mass energy of 13 TeV.

The physics programme of the LHC plans a significant increase of the luminosity in the coming years, especially for the high-luminosity phase beginning in 2026. Some parts of the ATLAS detectors must be upgraded to cope with the increased radiation levels. The small wheels of the muon spectrometer will be replaced by new small wheels (NSW) using Micromegas detectors. Irfu is responsible for producing some of the detectors, which performances are being validated on a test bench using cosmic muons.

During this doctoral work, the aim was to develop an analysis code for the test bench data, in order to characterize the Micromegas detectors of the NSW project. The code provides gain and detection efficiency informations, and will be used until the end of the project. The measured efficiencies are in agreement with the expectations except in some parts of the detectors that are being investigated.

The top quark plays a central role in many theories beyond

the standard model, for instance in models with extra dimensions, or with a modification of the Higgs boson sector. Many of these models predict an enhanced cross section for the very rare four top quark process, which is predicted to be approximately 12 fb at 13 TeV.

The study of the four top quark process is particularly interesting in the channel with leptons having the same electric charge because of the very low background level. This doctoral work is firstly carried out on the data taken in 2015 and 2016 by the ATLAS detector, in order to study the four top quark process and some new physics models. This process is very challenging given the very low number of events, and the difficult estimation of the fake lepton background. Several limits are set on new physics models and on the standard model four top quark production. The results on the standard model process are combined with the results from the single lepton channel to set the tightest limit at the time of its publication.

The four top quark production is being studied again, using the data taken from 2015 to 2018 this time, with the aim of making a first observation of this process. The doctoral work is focused on a new method for estimating the fake lepton background.

