



**HAL**  
open science

# Etude du couplage du quark top au boson de Higgs dans l'expérience ATLAS

Arthur Chomont

► **To cite this version:**

Arthur Chomont. Etude du couplage du quark top au boson de Higgs dans l'expérience ATLAS. Physique Nucléaire Expérimentale [nucl-ex]. Université Clermont Auvergne [2017-2020], 2017. Français. NNT : 2017CLFAC082 . tel-01822631

**HAL Id: tel-01822631**

**<https://theses.hal.science/tel-01822631>**

Submitted on 25 Jun 2018

**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.

Université Clermont Auvergne  
ÉCOLE DOCTORALE DES SCIENCES  
FONDAMENTALES

THESE

présentée pour l'obtention du grade de

DOCTEUR D'UNIVERSITÉ EN PHYSIQUE

*Spécialité : particules, interactions, univers*

Présentée par **Arthur CHOMONT**

Master

Étude du couplage du quark top au boson de Higgs dans  
l'expérience ATLAS

Soutenue publiquement le 29 septembre 2017 devant la commission d'examen :

<i>Président</i>	Stéphane MONTEIL	- Professeur (Université Clermont Auvergne)
<i>Rapporteurs</i>	Marumi KADO	- Directeur de recherche CNRS (LAL Orsay)
	Frédéric DELIOT	- Chercheur CEA/IRFU (Saclay)
<i>Examineurs</i>	Sandrine LAPLACE	- Chercheuse CNRS (LPNHE Paris)
	Susan GASCON-SHOTKIN	- Professeur (Université Claude Bernard, Lyon)
<i>Directeur de thèse</i>	Stéphane MONTEIL	- Professeur (Université Clermont Auvergne)
	Djamel BOUMEDIENE	- Chercheur CNRS (LPC Clermont-Ferrand)



# Remerciements

Tout d'abord, je remercie Alain Falvard et Dominique Pallin, directeurs du LPC pour m'avoir accueilli et permis de faire ma thèse au sein du laboratoire.

Je remercie vivement Marumi Kado et Frédéric Déliot pour avoir accepté d'être rapporteurs de ce manuscrit et pour leur commentaires et corrections. Je remercie également Suzanne Gascon-Shotkin, Sandrine Laplace et Stéphane Monteil qui ont accepté de participer à mon jury de thèse.

Je tiens à remercier tout particulièrement Djamel Boumediene, mon directeur de thèse, pour son encadrement et ses conseils précieux depuis mes premiers stages au LPC jusqu'à l'aboutissement de cette thèse.

Je tiens également à exprimer toute ma gratitude aux membres de l'équipe ATLAS de Clermont-Ferrand pour leur accueil, leur soutien et leur aide au cours de ces trois années.

Finalement je remercie tout les membres du laboratoire avec qui j'ai eu la chance d'échanger que ce soit au détour d'un couloir ou lors d'enseignements.



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Contexte théorique : le modèle standard de la physique des particules</b>	<b>3</b>
1.1 Introduction	3
1.2 Particules élémentaires et les interactions fondamentales	3
1.2.1 Les particules élémentaires	3
1.2.2 Les interactions fondamentales	4
1.2.3 Le formalisme du modèle standard	4
1.2.4 Les limites du modèle standard	9
1.3 Le boson de Higgs au LHC	10
1.3.1 Les voies de production du boson de Higgs	10
1.3.2 Les voies de désintégration du boson de Higgs	11
1.3.3 Les couplages du boson de Higgs	11
1.4 Le couplage du quark <i>top</i> au boson de Higgs	12
1.4.1 L'observation directe du couplage du boson de Higgs au quark <i>top</i>	13
1.4.2 Les états finals leptoniques du processus de production associé $t\bar{t}H$	14
1.4.3 Les processus multileptons aux LHC	15
<b>2 Outils expérimentaux : Le LHC et le détecteur ATLAS</b>	<b>17</b>
2.1 Introduction	17
2.2 Le LHC et les collisions proton-proton	17
2.2.1 Le grand collisionneur de hadrons (LHC)	17
2.2.2 Le circuit d'accélération et les propriétés du LHC	18
2.2.3 Les collisions protons-protons au LHC	19
2.3 Le détecteur ATLAS (A Toroidal LHC ApparatuS)	20
2.3.1 Conventions géométriques	21
2.3.2 Le détecteur interne	22
2.3.3 Les calorimètres	24
2.3.4 Le spectromètre à muons	27
2.3.5 Le système de déclenchement	28
<b>3 Étalonnage du calorimètre hadronique à tuiles d'ATLAS avec un système laser</b>	<b>31</b>
3.1 Introduction	31
3.2 Fonctionnement du calorimètre à tuiles d'ATLAS	31
3.2.1 Composante optique	31
3.2.2 Les blocs photomultiplicateurs (PMT)	31
3.2.3 L'électronique de lecture	32
3.2.4 La reconstruction du signal	32
3.3 Les systèmes d'étalonnage du TileCal	33
3.3.1 Le système Césium	34
3.3.2 Le système d'intégration de courants ( <i>Minimum Bias</i> )	34
3.3.3 Le système d'étalonnage par injection de charge	34
3.3.4 Le système Laser	35
3.4 Description du système Laser	35
3.4.1 La partie optique	35
3.4.2 La partie électronique	36

3.5	Stabilité du système Laser . . . . .	36
3.5.1	Piédestaux . . . . .	37
3.5.2	Étalonnage interne des diodes . . . . .	37
3.5.3	Linéarité . . . . .	37
3.6	Calcul des constantes d'étalonnages avec le système Laser . . . . .	38
3.6.1	Données Laser . . . . .	38
3.6.2	Analyse des données Laser . . . . .	38
3.6.3	Précision statistique du système Laser . . . . .	39
3.6.4	Résultat sur la variation de gain des canaux . . . . .	39
3.6.5	Comparaison avec les variations de gain mesurées par le système Césium . . . . .	40
3.7	Algorithme d'étalonnage des canaux du TileCal . . . . .	42
3.8	Conclusion . . . . .	44
<b>4</b>	<b>Simulation des événements dans ATLAS et définition des objets</b>	<b>49</b>
4.1	Introduction . . . . .	49
4.2	Simulation Monte-Carlo des événements dans ATLAS . . . . .	49
4.2.1	Génération des événements . . . . .	50
4.2.2	Simulation des interactions avec le détecteur ATLAS . . . . .	51
4.3	Définition des objets physiques . . . . .	51
4.3.1	Les électrons . . . . .	52
4.3.2	Les muons . . . . .	54
4.3.3	Les jets . . . . .	54
4.3.4	Les jets étiquetés $b$ . . . . .	58
4.3.5	Les taus hadroniques . . . . .	58
4.3.6	L'énergie transverse manquante . . . . .	58
4.3.7	BDT sur l'isolation des électrons et des muons . . . . .	58
<b>5</b>	<b>Mauvaise identification de la charge des électrons</b>	<b>61</b>
5.1	Introduction . . . . .	61
5.2	Stratégie d'estimation des taux de mauvaise identification de la charge des électrons . . . . .	61
5.3	Estimation des taux de mauvaise identification de la charge des électrons par maximum de vraisemblance . . . . .	63
5.4	Choix des intervalles en $\eta$ et $p_T$ pour l'estimation des taux de mauvaise identification de la charge des électrons . . . . .	65
5.5	Estimation des taux de mauvaise identification de la charge à haute impulsion transverse . . . . .	65
5.6	Incertitudes associées à l'estimation par maximum de vraisemblance . . . . .	66
5.7	Tests de consistances . . . . .	66
5.8	Estimation des taux de mauvaise estimation de la charge des électrons à partir des données pour $13.2fb^{-1}$ de données à 13TeV . . . . .	68
5.8.1	Validation de la méthode par maximum de vraisemblance . . . . .	68
5.8.2	Incertitudes systématiques . . . . .	68
5.8.3	Taux observés . . . . .	70
5.8.4	Test de fermeture . . . . .	70
5.9	Estimation des taux de mauvaise estimation de la charge des électrons à partir des données pour $36.1fb^{-1}$ de données à 13TeV . . . . .	71
5.9.1	Validation de la méthode par maximum de vraisemblance . . . . .	71
5.9.2	Incertitudes systématiques . . . . .	71
5.9.3	Taux observés . . . . .	72
5.9.4	Test de fermeture . . . . .	74

5.10 Conclusion . . . . .	75
<b>6 Interprétation statistique des résultats</b>	<b>77</b>
6.1 Modèle Statistique . . . . .	77
6.2 Incertitudes et paramètres de nuisances . . . . .	77
6.2.1 Incertitude statistique . . . . .	77
6.2.2 Incertitude systématiques . . . . .	78
6.2.3 Incertitude systématiques sur la normalisation . . . . .	78
6.2.4 Incertitude systématiques sur la forme des distributions . . . . .	78
6.3 Test statistique et Profile Likelihood ratio . . . . .	78
6.4 Données Asimov . . . . .	79
6.5 Limite d'exclusion observées et attendues par méthode CLs . . . . .	80
6.6 Représentation de l'impact des paramètres de nuisance sur la valeur de la force du signal . . . . .	81
<b>7 Etude du couplage du boson de Higgs au quark <i>top</i> dans l'expérience ATLAS</b>	<b>83</b>
7.1 Introduction . . . . .	83
7.2 Objets physiques et présélection . . . . .	83
7.2.1 Objets physiques . . . . .	83
7.2.2 Exclusion du double comptage des objets . . . . .	84
7.2.3 Présélection . . . . .	84
7.3 Échantillons de données, de simulation et description des bruits de fonds irréductibles	85
7.3.1 Données utilisées . . . . .	85
7.3.2 Simulation : Échantillons de signal . . . . .	85
7.3.3 Définition des bruits de fond et échantillons utilisés . . . . .	86
7.4 Définition de la région enrichie en signal . . . . .	87
7.5 Estimation des bruits de fonds instrumentaux . . . . .	87
7.5.1 Estimation du nombre d'événements avec des électrons dont la charge électrique est mal reconstruite . . . . .	89
7.5.2 Estimation du nombre d'événements avec des leptons secondaires . . . . .	89
7.6 Incertitudes systématiques . . . . .	92
7.6.1 Les incertitudes théoriques . . . . .	92
7.6.2 Incertitudes systématiques instrumentales . . . . .	94
7.7 Validation de la simulation des bruits de fond . . . . .	94
7.8 Résultats . . . . .	94
7.9 Conclusion . . . . .	102
<b>8 Reconstruction cinématique des événements</b>	<b>107</b>
8.1 Introduction . . . . .	107
8.2 Combinatoire dans la reconstruction des quarks <i>top</i> . . . . .	107
8.3 Ajustement cinématique . . . . .	108
8.3.1 Méthode . . . . .	108
8.3.2 Définition des résolutions et des biais sur les objets . . . . .	108
8.4 Résultat . . . . .	109
8.5 Conclusion . . . . .	111
<b>9 Extensions et améliorations de l'analyse</b>	<b>115</b>
9.1 Introduction . . . . .	115
9.2 Motivation . . . . .	115
9.3 Ajout de nouveaux canaux . . . . .	115
9.4 Nouvelle définition des objets physiques et présélection . . . . .	116



9.4.1	Système de déclenchement . . . . .	116
9.4.2	Objets physiques . . . . .	116
9.4.3	Présélection . . . . .	117
9.5	Estimation et validation des bruits de fond . . . . .	117
9.5.1	Estimation des événements avec un électron dont la charge électrique est mal reconstruite . . . . .	117
9.5.2	Estimation des événements avec des leptons non-prompts par méthode de la matrice . . . . .	119
9.6	Sélection des événements . . . . .	123
9.6.1	Analyse multivariée dans le canal $2\ell SS0\tau$ . . . . .	123
9.6.2	Catégorisation des événements dans la région $2\ell SS0\tau$ . . . . .	123
9.6.3	Résumé des régions de signal . . . . .	123
9.7	Incertitudes systématiques . . . . .	128
9.8	Résultats . . . . .	129
9.8.1	Nombres d'événements attendus et observés dans les régions de signal . . . . .	129
9.8.2	Ajustement aux données . . . . .	129
9.9	Conclusion . . . . .	130
<b>10</b>	<b>Combinaison des analyses <math>t\bar{t}H</math></b> . . . . .	<b>141</b>
10.1	Introduction . . . . .	141
10.2	Recherche du processus $t\bar{t}H$ avec $H \rightarrow b\bar{b}$ . . . . .	141
10.3	Recherche du processus $t\bar{t}H$ avec $H \rightarrow \gamma\gamma$ . . . . .	141
10.4	Résultat de la combinaison . . . . .	141
10.4.1	Orthogonalité des analyses . . . . .	141
10.4.2	Schéma de corrélation . . . . .	142
10.4.3	Résultats pour l'analyse à $13.2fb^{-1}$ . . . . .	142
10.4.4	Résultats pour l'analyse à $36.1fb^{-1}$ . . . . .	144
10.5	Conclusion . . . . .	146
	<b>Conclusion</b> . . . . .	<b>147</b>
<b>A</b>	<b>Validation des bruits de fond l'analyse à <math>36.1fb^{-1}</math></b> . . . . .	<b>149</b>
A.1	Validation des leptons non-prompts dans les canaux $2\ell SS0\tau$ et $3\ell 0\tau$ . . . . .	149
A.2	Validation du processus $t\bar{t}W$ dans les canaux $2\ell SS0\tau$ et $3\ell 0\tau$ . . . . .	151
A.3	Validation du processus $t\bar{t}Z$ dans les canaux $3\ell 0\tau$ et $4\ell$ . . . . .	152
<b>B</b>	<b>Incertitudes systématiques expérimentales appliquées sur les objets dans l'analyse à <math>36.1fb^{-1}</math></b> . . . . .	<b>155</b>
	<b>Bibliographie</b> . . . . .	<b>159</b>



# Introduction

Le modèle standard de la Physique des Particules est le cadre théorique de la physique de l'infiniment petit, des particules élémentaires aux interactions fondamentales. Il a été développé au cours des années 1960 et décrit précisément trois des quatre forces fondamentales que sont l'interaction forte, l'interaction faible et l'interaction électromagnétique. Toutefois l'absence de description de la dernière force fondamentale ainsi que diverses observations telle que l'existence de matière noire tend à prouver que le modèle standard n'est pas complet. De nombreux modèles tentent de compléter le modèle standard. Ils sont regroupés sous la dénomination de Nouvelle Physique ou Physique au-delà du modèle standard.

Un des outils le plus à même aujourd'hui de tester les prédictions du modèle standard est le LHC, un collisionneur proton-proton qui produit des collisions avec une énergie dans le centre de masse de 13 TeV. L'exemple le plus récent de l'importance du LHC pour tester le modèle standard est la découverte en juillet 2012 au LHC par les expériences ATLAS et CMS du boson de Higgs. Son existence avait été prédite en 1964 par Robert Brout, François Englert et Peter Higgs. Les caractéristiques de cette particule sont maintenant étudiées pour savoir si le boson de Higgs découvert correspond à celui prédit par le modèle standard ou bien à celui d'un scénario de Nouvelle Physique. Parmi les propriétés du boson de Higgs à tester, son couplage aux fermions est très intéressant car lié à la masse de ces particules. Le couplage au quark *top* devrait donc être le plus intense, lui conférant un impact significatif sur les prédictions de la théorie. L'étude du processus  $t\bar{t}H$  est un bon moyen de tester le modèle standard par une mesure directe de ce couplage. Le but de cette thèse est l'étude du couplage du quark *top* au boson de Higgs par la recherche du processus  $t\bar{t}H$  dans un état final avec plusieurs leptons avec une contribution à l'opération du détecteur ATLAS. Une première analyse infructueuse a été faite au Run 1 du LHC pour observer le processus  $t\bar{t}H$  tandis que la mesure du taux de production du boson de Higgs et son observation par sa désintégration en une paire de photon permet une mesure indirecte du couplage.

Le Chapitre 1 est consacré à la description du modèle standard avec une attention particulière apportée aux propriétés du boson de Higgs et comment la mesure du couplage du boson de Higgs aux autres particules est réalisée.

Les Chapitres 2, 3 et 4 présentent le contexte expérimental de ce travail. Le Chapitre 2 décrit le LHC. Le détecteur ATLAS, l'un des quatre détecteurs situés auprès du LHC permettant de recueillir les données des collisions, est présenté avec les sous-détecteurs qui le composent. Le Chapitre 3 détaille les résultats de l'étalonnage de l'un des sous-détecteurs d'ATLAS, le calorimètre hadronique, grâce à un système Laser. Le Chapitre 4 explicite le processus de simulation d'événements dans ATLAS ainsi que la reconstruction des objets physiques nécessaires aux analyses tels que les jets ou les électrons.

Les Chapitres 5, 6, 7, 8 et 9 s'attachent à la description des différents éléments composant l'analyse des données pour la recherche du processus  $t\bar{t}H$ . Le Chapitre 5 est consacré à la description de l'estimation de l'un des bruits de fonds de l'analyse qui est dû à la mauvaise reconstruction de la charge des électrons. Les éléments statistiques nécessaires à l'analyse des données sont discutés succinctement dans le Chapitre 6. Une première analyse avec les données enregistrées par ATLAS en 2015 et au début de l'année 2016 est présentée dans le Chapitre 7. Le Chapitre 8 décrit une étude de la cinématique des événements à deux leptons de même signe visant à améliorer les performances de l'analyse. Une analyse plus optimisée avec le lot complet de données disponibles est montrée dans le Chapitre 9. Finalement, le résultat de la combinaison de toutes les analyses cherchant le processus  $t\bar{t}H$  dans ATLAS par diverses voies de désintégration est discuté dans le Chapitre 10.

Le travail de ces trois années de thèse est plus particulièrement décrit dans les Chapitres 3, 5 et 7 à 10. Concernant les analyses présentées dans les Chapitres 7 à 9, mon travail personnel a porté sur l'étude du canal avec deux leptons de même charge électrique par l'estimation de l'un des bruits de

fond ainsi que le traitement statistique des données pour la combinaison de l'ensemble des canaux. En outre pour l'analyse présentée dans le Chapitre 9, j'ai participé à la mise en place de l'analyse par catégorisation dans le canal avec deux leptons de même charge. Enfin pour la combinaison j'ai travaillé à l'harmonisation du traitement des bruits de fond et des incertitudes systématiques avec les autres analyses de recherche du processus  $t\bar{t}H$ .

# Contexte théorique : le modèle standard de la physique des particules

---

## 1.1 Introduction

Le modèle standard de la physique des particules est à l'heure actuelle la meilleure description des particules élémentaires et de leurs interactions. La découverte du boson de Higgs en 2012 est la dernière pièce confortant cette théorie.

Dans ce chapitre, le modèle standard est décrit avec ses réussites mais aussi ses faiblesses ainsi que l'apport que l'étude et la mesure des propriétés du boson de Higgs peuvent avoir sur la connaissance des limites du modèle standard.

## 1.2 Particules élémentaires et les interactions fondamentales

### 1.2.1 Les particules élémentaires

Les **fermions** sont des particules de spin demi-entier qui obéissent à la statistique de Fermi-Dirac et constituent la matière. Ils peuvent être séparés en deux sous-groupes : les **leptons** et les **quarks**. Les premiers portent une charge électrique entière de -1 pour les leptons chargés ( $e^-$ ,  $\mu^-$ ,  $\tau^-$ ) et nulle pour les neutrinos. Les quarks ont des charges électriques fractionnaire :  $\frac{2}{3}$  pour les quarks *up*, *charm*, *top* ( $u$ ,  $c$  et  $t$ ) et  $\frac{1}{3}$  pour les quarks *down*, *strange*, *bottom* ( $d$ ,  $s$  et  $b$ ). Les quarks ont aussi la propriété d'être sensibles à l'interaction forte car porteurs de charges de couleur (voir la partie 1.2.3.3), ce qui n'est pas le cas des leptons. Un quark ne peut être libre et il va donc s'associer avec d'autres quarks pour former des structures de couleur neutre appelées hadrons : les mésons formés d'un quark et d'un antiquark et les baryons formés de trois quarks. La seule exception va être le quark *top* qui se désintègre dans un temps caractéristique plus court que celui du processus d'hadronisation à cause de sa masse très importante.

Il existe trois familles dans chacun de ces deux sous-groupes, la matière ordinaire étant constituée exclusivement de fermions de la première famille. Les particules de cette famille sont pour les quarks, le quark *up* et le quark *down* qui sont les bases des protons ( $uud$ ) et neutrons ( $udd$ ) et pour les leptons, l'électron et le neutrino électronique. Les autres quarks et leptons de deuxième et troisième générations sont listés dans la Table 1.1. À chacune de ces particules est associée une antiparticule de même spin et masse mais de charge électrique opposée.

Les **bosons** ont pour leur part un spin entier et obéissent à la statistique de Bose-Einstein. Ces particules portent les interactions entre les différents fermions. Il existe au moins un boson par interaction fondamentale décrite dans la suite du chapitre (excepté pour la gravitation dont l'existence d'un boson lié à cette interaction n'est qu'hypothétique). Ces bosons vecteurs (de spin égal à un) sont appelés les bosons  $W^\pm$ ,  $Z$ , le photon et les gluons. Il existe un autre boson, scalaire (de spin nul), le boson de Higgs, qui représente l'interaction entre les particules et le champ de Higgs, cette interaction donnant une masse aux particules.

	Particule	Masse (MeV)	Charge
Quarks	<i>up</i> $u$	2.2	2/3
	<i>down</i> $d$	4.7	1/3
	<i>charm</i> $c$	$1.27 \times 10^3$	2/3
	<i>strange</i> $s$	96	1/3
	<i>top</i> $t$	$173.21 \times 10^3$	2/3
	<i>bottom</i> $b$	$4.2 \times 10^3$	1/3
Leptons	electron $e$	0.511	-1
	Neutrino électronique $\nu_e$	$< 2 \times 10^{-6}$	0
	muon $\mu$	105	-1
	Neutrino muonique $\nu_\mu$	$< 0.19$	0
	Tau $\tau$	$1.78 \times 10^3$	-1
	Neutrino tauique $\nu_\tau$	$< 18.2$	0

TABLE 1.1 – Récapitulatif des fermions et leurs propriétés [1]

### 1.2.2 Les interactions fondamentales

Les interactions fondamentales sont au nombre de 4 : forte, électromagnétique, faible et gravitationnelle. Le modèle standard décrit les trois premières quantiquement tandis que la gravité, dont l'amplitude à l'échelle subatomique est inférieure à celle de l'interaction faible par 34 ordres de grandeurs, va être négligée.

L'interaction forte agit à très courte portée et est responsable de la cohésion des noyaux et hadrons. Ses particules médiatrices sont les gluons au nombre de huit. Elle agit sur les particules portant une charge de couleurs (quarks, gluons). L'interaction faible agit sur toutes les particules à courte distance et se manifeste par exemple dans la radioactivité  $\beta$ . Ses médiateurs sont les bosons  $W^\pm$  et  $Z^0$ . La dernière interaction fondamentale décrite dans le modèle standard est l'interaction électromagnétique qui agit sur les particules chargées électriquement (leptons chargés, quarks, bosons  $W^\pm$ ). Le photon est la particule médiatrice de l'interaction électromagnétique. La Table 1.2 résume les caractéristiques des 4 forces fondamentales.

En outre une dernière particule est prédite dans le modèle standard, il s'agit du boson de Higgs. Cette particule est un boson scalaire, de spin nul, qui rendrait compte de l'interaction des particules avec le champs de Higgs et confère ainsi une masse aux fermions et bosons de jauge. Une particule de masse  $125.09 \pm 0.21 \pm 0.11 \text{ GeV}/c^2$  correspondant à cette description a été découverte en 2012 au LHC par les équipes ATLAS [2] et CMS [3].

Interaction	Bosons porteurs	Charge	Portée (en m)	Intensité ( à $10^{-18} \text{ m}$ )
Faible	$W^\pm / Z$	Isospin faible	$\leq 10^{-16}$	0.8
Electromagnétique	Photon ( $\gamma$ )	Electrique	$\infty$	1
Forte	Gluons	Couleur	$\leq 10^{-18}$	25
Gravitation	Graviton (hypothétique)	Masse	$\infty$	$10^{-41}$

TABLE 1.2 – Interactions fondamentales

### 1.2.3 Le formalisme du modèle standard

Le Modèle standard regroupe la mécanique quantique et la relativité restreinte dans un formalisme commun que l'on appelle une théorie quantique des champs [4]. Dans ce cadre théorique une dualité onde-particule va être formalisée avec pour toute particule un champ  $\psi(\vec{x}, t)$  associé, où  $\vec{x}$  et  $t$  sont respectivement le vecteur position et le temps. Des opérateurs de création et/ou d'an-

nihilation dans cette théorie permettent la description d'un nombre variable de particules. Pour construire le modèle standard, le formalisme de Lagrange est utilisé. On peut alors accéder à la dynamique des systèmes physiques considérés par la maximisation de l'action  $S = \int \mathcal{L} dx^4$  avec  $\mathcal{L}$  le lagrangien. La construction du lagrangien va être contrainte par la présence de symétries dans la nature. D'après le théorème de Noether [5] une grandeur conservée au cours de l'évolution du système peut être associée à chacune de ces symétries. Pour la physique des particules on peut prendre l'exemple de la charge de couleur qui doit être conservée dans le cadre de l'interaction forte. Chaque interaction fondamentale repose sur un groupe de symétrie particulier :  $SU(3)_C$  pour l'interaction forte,  $SU(2)_L$  pour l'interaction faible et  $U(1)_Y$  pour l'interaction électromagnétique. Le lagrangien du modèle standard est invariant pour toute transformation du groupe de symétrie  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ .

### 1.2.3.1 Électrodynamique quantique

Le lagrangien correspondant à un fermion chargé libre de masse  $m$ , représenté par le champ  $\Psi(x)$  s'écrit :

$$\mathcal{L} = \bar{\Psi}(x)[i\gamma^\mu \partial_\mu - m]\Psi(x) \quad (1.1)$$

où  $\gamma^\mu$  sont les matrices de Dirac. On considère une transformation de symétrie telle que :

$$\Psi \rightarrow e^{i\theta(x)}\Psi. \quad (1.2)$$

La dépendance en  $x$  de  $\theta$  confère un caractère local à la transformation. La dérivée devient :

$$\partial_\mu \Psi(x) = e^{i\theta(x)}[\partial_\mu - i\partial_\mu \theta(x)]\Psi(x). \quad (1.3)$$

Le second terme de l'équation aboutit à la violation de l'invariance du lagrangien sous cette transformation. Pour rétablir l'invariance on définit la dérivée covariante  $D_\mu$  qui remplace la dérivée :

$$D_\mu = \partial_\mu - ieA_\mu \quad \text{et} \quad D_\mu \Psi(x) = e^{i\theta(x)}D_\mu \Psi(x). \quad (1.4)$$

$A_\mu$  est appelé champ de jauge. L'invariance du lagrangien repose sur les propriétés de transformation qui doivent compenser les termes supplémentaires. Ainsi :

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu \theta(x). \quad (1.5)$$

La propagation du champ de jauge est assurée par l'introduction dans le lagrangien d'un terme cinétique pour ce champ invariant de jauge :

$$\mathcal{L}_{propagation} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad \text{avec} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (1.6)$$

Finalement le lagrangien s'écrit :

$$\begin{aligned} \mathcal{L} &= \bar{\Psi}(x)[i\gamma^\mu D_\mu - m]\Psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \underbrace{\bar{\Psi}(x)[i\gamma^\mu \partial_\mu - m]\Psi(x)}_{\text{Particule libre}} - \underbrace{\bar{\Psi}(x)eA_\mu\Psi(x)}_{\text{Interaction}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{Propagation du champ de jauge}}. \end{aligned} \quad (1.7)$$

Dans l'expression 1.7, les interactions entre les fermions chargés sont liées à un boson de jauge  $A_\mu$  qui correspond au photon. En outre, aucun terme de masse n'est présent pour ce boson car dans le cas contraire la symétrie de jauge  $U(1)$  serait brisée.

### 1.2.3.2 Unification électrofaible

L'interaction faible entre en jeu dans les désintégrations  $\beta$ . Une caractéristique de cette interaction est que la parité<sup>1</sup> est maximale violée dans l'interaction faible tel que montré par les expériences de C.S.Wu [6] et n'agit que sur les particules de chiralité gauche<sup>2</sup>. Au début des années 1960, Glashow, Salam et Weinberg proposent l'unification des théories de l'interaction faible et de l'électromagnétisme dans un même formalisme conduisant à une unique interaction : l'interaction électrofaible [7] [8]. Le groupe de symétrie locale de l'interaction électrofaible est  $SU(2)_L \otimes U(1)_Y$ . L'invariance locale de jauge conduit à définir quatre champs de jauge  $W^i$  ( $i=1,2,3$ ) et  $B$ . Les trois premiers découlent de la symétrie  $SU(2)_L$  où l'indice  $L$  indique que cette interaction n'agit que sur les composantes de chiralité gauche des particules. Le dernier vient de l'unique générateur du groupe  $U(1)_Y$  où l'indice  $Y$  correspond à l'hypercharge faible. Dans la construction du lagrangien de l'interaction électrofaible de nouvelles dérivées covariantes tenant compte des états de chiralité sont définies :

$$D_\mu^L = \partial_\mu + igY B_\mu + ig' \frac{\sigma_a}{2} W_\mu^a \quad \text{et} \quad D_\mu^R = \partial_\mu + igY B_\mu \quad (1.8)$$

avec  $g$  et  $g'$  les constantes de couplage des deux interactions et  $\sigma_a$  ( $a=1,2,3$ ) les matrices de Pauli, génératrices du groupe  $SU(2)_L$ . Des termes cinétiques pour les champs de jauge  $W^{1,2,3}$  et  $B$  sont aussi introduits :

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad \text{et} \quad W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g' \varepsilon_{abc} W_\mu^b W_\nu^c \quad (1.9)$$

où  $\varepsilon_{abc}$  sont les constantes de structure du groupe  $SU(2)_L$ . Le lagrangien électro-faible s'écrit alors :

$$\mathcal{L}_{\text{electrofaible}} = \underbrace{i\bar{\Psi}_L \gamma_\mu D_\mu^L \Psi_L}_{\text{fermions gauche}} + \underbrace{i\bar{\Psi}_R \gamma_\mu D_\mu^R \Psi_R}_{\text{fermions droits}} - \underbrace{\frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}}_{\text{cinétique des champs de jauge}}. \quad (1.10)$$

Selon l'équation 1.10 les bosons  $W^a$  n'interagissent qu'avec les composantes de chiralité gauche tandis que le boson  $B$  interagit avec l'ensemble des composants. Les bosons  $W^\pm$ ,  $Z^0$  et le photon ne sont pas directement ces champs  $W^a$  et  $B$  mais s'obtiennent à partir de combinaisons linéaires de ceux-ci :

$$\begin{aligned} W^\pm &= (W_\mu^1 \mp W_\mu^2) / \sqrt{2}, \\ \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} &= \begin{pmatrix} \cos \theta_w & \sin \theta_w \\ -\sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \end{aligned} \quad (1.11)$$

avec  $\theta_w$  l'angle de mélange. A ce stade les bosons  $W^\pm$  et  $Z^0$  n'ont pas de masse.

### 1.2.3.3 L'interaction forte : chromodynamique quantique

L'interaction agit sur les particules porteuses de charge de couleur (quarks, gluons). La chromodynamique quantique (ou QCD) [9] est une théorie de jauge non-abeélienne<sup>3</sup>, construite à partir

---

1. Symétrie de renversement des coordonnées d'espace (miroir).  
 2. Dans un cadre ultra-relativiste, la chiralité peut être confondue avec l'hélicité qui est définie comme la projection du spin selon l'axe de l'impulsion. Une particule sera d'hélicité droite si ses spin et impulsion sont dans la même direction.  
 3. La commutation des opérateurs de rotation avec les différents champs n'est pas réalisée



du groupe de symétrie  $SU(3)_C$ . Chaque quark s'écrit sous  $SU(3)_C$  comme un triplet de couleur :

$$\Psi = \begin{pmatrix} \Psi_R \\ \Psi_G \\ \Psi_B \end{pmatrix} \text{ qui se transforme comme } \Psi \rightarrow e^{-\frac{i}{2}\lambda^a\theta^a(x)}\Psi \quad (1.12)$$

avec  $\lambda_a$  les 8 matrices de Gell-Mann génératrices du groupe  $SU(3)_C$ . Une réflexion similaire sur les conditions d'invariance du lagrangien à celles menées pour la QED implique l'introduction de la dérivée covariante suivante :

$$D_\mu = \partial_\mu - ig_s \frac{\lambda_a}{2} G_\mu^a \quad (1.13)$$

avec  $G_\mu^a$  qui représente les 8 champs de jauge de l'interaction forte associés aux gluons et  $g_s$  la constante de couplage fort. Le lagrangien final de la QCD est donné par :

$$\begin{aligned} \mathcal{L}_{QCD} = & -\frac{1}{4}(\partial_\mu G_\nu^a - \partial_\nu G_\mu^a)(\partial^\mu G_a^\nu - \partial^\nu G_a^\mu) + \bar{\Psi}(i\gamma_\mu\partial^\mu - m)\Psi \\ & + ig_s \bar{\Psi}\lambda_a G_\mu^a \gamma_\mu \Psi \\ & + \frac{g_s}{2} f_a^{bc}(\partial_\mu G_\nu^a - \partial_\nu G_\mu^a)G_b^\mu G_c^\nu - \frac{g_s^2}{2} f_{abc} f^{ade} G_\mu^b G_\nu^c G_d^\mu G_e^\mu. \end{aligned} \quad (1.14)$$

La première ligne correspond aux propagateurs des gluons et quarks. La seconde ligne décrit l'interaction quark-gluon. La dernière ligne représente le couplage entre les gluons avec  $f^{abc}$  les constantes de structure du groupe de jauge.

Un élément particulier de la QCD est le fait que sa constante de couplage  $\alpha_S$  diminue avec l'échelle d'énergie considérée. Ainsi deux effets apparaissent pour les quarks : la liberté asymptotique et le confinement. La liberté asymptotique est le fait que les quarks, à très courte distance, se comportent comme des particules libres. Le confinement des quarks est le fait qu'un quark produit après une interaction ne peut rester libre, il va donc s'hadroniser pour former des états liés.

### 1.2.3.4 Le mécanisme de Brout-Englert-Higgs

Dans la théorie de l'interaction électrofaible telle que décrite précédemment les particules n'ont pas de masse ce qui est contraire aux observations faites des différentes particules : de l'observation de l'électron par Thomson en 1897 à la découverte des bosons  $W^\pm$  et  $Z^0$  par les expériences UA1 [10][11] et UA2 [12][13] au CERN au début des années 1980.

En 1964, François Englert et Robert Brout d'un côté et Peter Higgs de l'autre proposent un mécanisme permettant l'introduction de termes de masse pour les bosons de jauge [14][15]. Dans ce mécanisme la symétrie de jauge électrofaible est brisée tout en conservant intacte la symétrie électromagnétique. Cette brisure est réalisée en introduisant un doublet de champs scalaires complexes :

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \text{avec} \quad \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.15)$$

où les champs  $\phi_i$  sont réels. Ce doublet va être décrit par le lagrangien :

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \quad (1.16)$$

avec  $D_\mu$  la dérivée covariante telle que décrite pour l'interaction électrofaible et  $V$  le potentiel de Higgs qui s'écrit :

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \quad (1.17)$$

avec  $\mu$  et  $\lambda$  deux paramètres libres. L'état fondamental du champ  $\phi$ , correspondant à l'état du vide, est défini par la valeur du champ au minimum du potentiel  $V(\phi)$ . Le cas où  $\mu^2$  est positif conduit

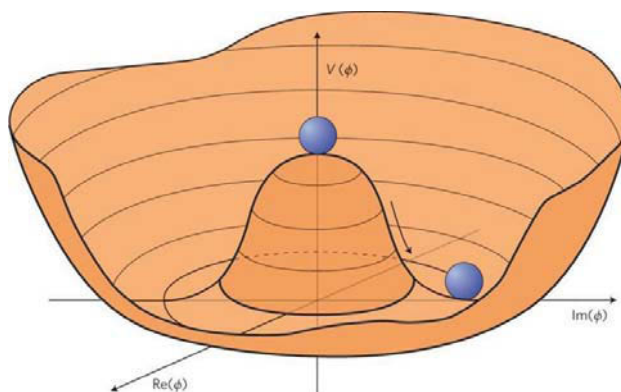


FIGURE 1.1 – Forme du potentiel  $V(\phi)$  pour  $\lambda > 0$  et  $\mu^2 < 0$ .

à avoir comme solution  $\phi = 0$  et des bosons de jauge qui restent sans masse. Le potentiel  $V(\phi)$  est représenté sur la Figure 1.1 dans le cas de  $\mu^2 < 0$ .

Le minimum est atteint pour un nombre infini de combinaisons correspondant à un cercle centré en 0 et de rayon  $v$ , la valeur attendue du vide (*v.e.v.*). On peut choisir arbitrairement  $v = \sqrt{\frac{-\mu^2}{\lambda}}$ , ce choix brisant la symétrie de jauge. On obtient alors une valeur dans le vide du champ  $\phi$  telle que :

$$\phi_{min} = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \quad (1.18)$$

Les fluctuations du champ autour de son minimum modifient son expression et on obtient :

$$\phi = \frac{e^{i\theta_a(x)\frac{\sigma_a}{2}}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (1.19)$$

où les  $\theta_a$  correspondent à trois bosons de Goldstone de masse nulle. En choisissant une jauge unitaire, les degrés de liberté associés à ces bosons peuvent être absorbés comme modes longitudinaux des bosons de jauge  $W^\pm$  et  $Z$ . Le champ scalaire restant,  $H$ , est associé au boson de Higgs. L'introduction de cette expression de  $\phi$  dans le lagrangien de Higgs permet de faire apparaître des termes de masse pour les bosons  $W^\pm$ ,  $Z$  et  $H$  tout en conservant le photon sans masse :

$$\begin{aligned} m_{W^\pm} &= \frac{1}{2}vg'. \\ m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2}. \\ m_H &= \sqrt{2\mu^2}. \end{aligned}$$

Après la brisure spontanée de symétrie une masse peut être ajoutée aux fermions en ajoutant un terme d'interaction supplémentaire, dit de Yukawa, dans le lagrangien :

$$\mathcal{L} = -(1 + \frac{H}{v}) \sum_{i,j} [(\bar{q}_d)_L^i (M_d)_{ij} (q_d)_R^j + (\bar{q}_u)_L^i (M_u)_{ij} (q_u)_R^j + (\bar{l})_L^i (M_l)_{ij} (l)_R^j + HC] \quad (1.20)$$

avec les  $M_{ij}$  des matrices  $3 \times 3$ , appelées matrices de Yukawa. Le premier terme de l'équation 1.20 correspond au terme de masse des fermions quand le second terme (en  $\frac{H}{v}$ ) décrit l'interaction d'un fermion et du boson de Higgs. Cette interaction est proportionnelle à la masse du fermion et suit

une constante de couplage qui vaut  $k = \sqrt{2} \frac{m_f}{v}$ .

Les matrices de Yukawa ne sont pas diagonales ce qui signifie que les états propres de masse ne sont pas les états propres de l'interaction faible. Dans le cas des quarks la diagonalisation des matrices de Yukawa permet d'extraire une matrice unitaire de mélange des trois familles de quarks. Cette matrice est appelée matrice de Cabibbo, Kobayashi, Maskawa (CKM) [16][17] :

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (1.21)$$

Le caractère unitaire de la matrice CKM permet de réduire son nombre de paramètres libres, dans l'hypothèse de trois familles de quarks, à trois angles de mélange et une phase, cette dernière étant responsable de la violation CP<sup>4</sup> dans les interactions faibles.

### 1.2.4 Les limites du modèle standard

A l'heure actuelle, le modèle standard décrit avec précision et justesse la majorité des phénomènes physiques observés. Ainsi les bosons de  $W^\pm$ ,  $Z$  et le boson de Higgs ont été prédits avant d'être observés. Toutefois certains éléments ne sont pas expliqués par ce modèle tels que :

- le modèle standard ne peut rendre compte de la présence seule de la matière dans l'univers tel que nous le connaissons : l'excédent de matière par rapport à l'antimatière devrait être, selon ce modèle, beaucoup plus faible que celui observé ;
- seules trois des quatre interactions fondamentales sont quantifiées dans ce modèle qui n'inclut pas la gravitation. A l'échelle électrofaible ( $\simeq 100 \text{ GeV}$ ) caractéristique du modèle standard la gravitation est négligeable, or à l'échelle de Planck ( $\Lambda_P \simeq 10^{19} \text{ GeV}$ ) la gravitation doit être prise en compte. Cet écart entre les deux échelles d'énergie et le fait que le modèle standard ne décrit que l'effet électrofaible définissent ce qui est appelé le problème de hiérarchie ;
- une autre limite du modèle standard est le problème de naturalité de la masse du boson de Higgs. Le problème porte ici sur la masse mesurée du boson de Higgs qui va prendre en compte l'existence de corrections radiatives dues à des boucles de fermions ou bosons. Dans ce cas la masse mesurée  $m$  du boson de Higgs sera dans le cas d'une boucle de fermions :

$$m^2 = m_0^2 - \frac{\lambda_f^2}{8\pi^2} \Lambda^2 + \dots \quad (1.22)$$

où  $m_0$  correspond à la masse nue du boson de Higgs,  $\lambda_f$  le couplage d'un fermion au boson de Higgs,  $\Lambda$  l'échelle d'énergie jusqu'à laquelle le modèle standard est valide. À la suite des contraintes obtenues par la mesure de la masse du boson de Higgs, on obtient que  $m_0$  et  $\Lambda$  doivent être ajustés à 32 décimales près ;

- une autre problématique est la présence de la matière noire (matière massive non prédite par le modèle standard) prouvée par les observations astrophysiques sur les masses des galaxies et leur courbe de vitesses de rotation ;

Beaucoup de modèles théoriques de Nouvelle Physique sont développés à l'heure actuelle pour tenter de combler les lacunes du modèle standard. Toutefois ces modèles ne prédisent pas l'échelle d'énergie de cette Nouvelle Physique. Une méthode pour tester le modèle standard et l'échelle d'énergie jusqu'à laquelle il est valable est de faire des tests de précision sur les paramètres du modèle standard tels que les propriétés du boson de Higgs.

4. La violation de la symétrie Charge-Parité a été observée en 1964 dans la désintégration des kaons neutres [18]

### 1.3 Le boson de Higgs au LHC

La découverte du boson de Higgs (avec une masse d'environ  $125.09 \text{ GeV}/c^2$  [1]) au LHC en juillet 2012 ouvre un nouveau champ de recherche pour tester la validité du modèle standard et en tester les limites par l'étude des différentes caractéristiques de ce boson. Les premières estimations du spin et de la parité de la particule observée au CERN sont compatibles avec les caractéristiques attendues du boson de Higgs dans le modèle standard : Spin 0, Parité + [19][20]. Le couplage du boson de Higgs aux autres particules est accessible par ses voies de production et de désintégration qui sont décrites dans la suite de cette section.

#### 1.3.1 Les voies de production du boson de Higgs

Il est prévu théoriquement que le boson de Higgs puisse être produit selon différents processus. Pour un boson de Higgs d'une masse de  $125 \text{ GeV}$ , la production se fait majoritairement par fusion de gluons. Les autres voies de production observables au LHC sont par fusion de bosons vecteurs (VBF), par Higgs Strahlung (Higgs produit dans l'état final avec un boson vecteur) et finalement par fusion de quark *top* (l'état final présentant un boson de Higgs associé à une paire de quarks *top-antitop*<sup>5</sup>). Par ces différents processus décrits sur la Figure 1.2, les couplages du boson de Higgs aux autres particules sont accessibles : aux fermions (indirectement dans les boucles nécessaires à la production par fusion de gluons et principalement par boucles de quarks *top*), aux bosons  $W^\pm$  et  $Z$  (par VBF et Higgs Strahlung) et au quark *top* (directement par la production associée avec une paire  $t\bar{t}$ ).

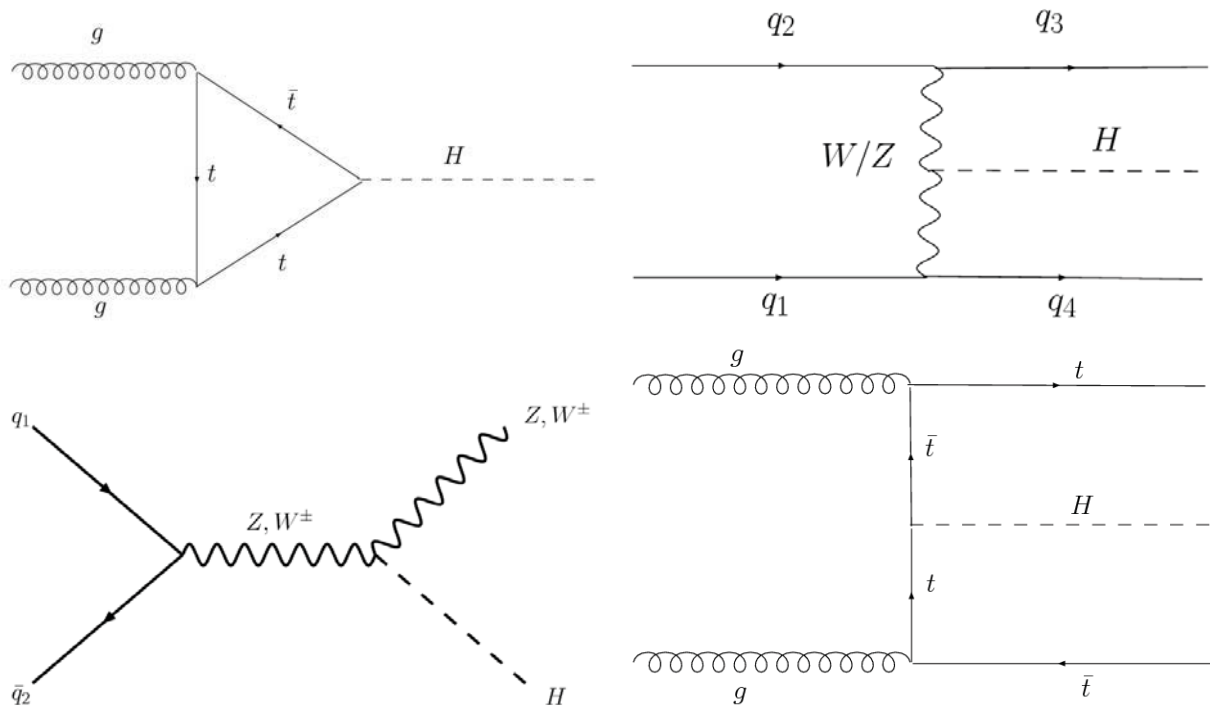


FIGURE 1.2 – Diagrammes de production du boson de Higgs au premier ordre au LHC par fusion de gluons (haut, gauche), par fusion de bosons vecteurs (haut, droite), par Higgs Strahlung (bas, gauche) et par production associée à une paire de quark *top* (bas, droite).

<sup>5</sup>. Ce processus de production existe pour tous les quarks mais seul la production avec des quarks *top* est suffisamment important pour être observée au LHC

### 1.3.2 Les voies de désintégration du boson de Higgs

Le boson de Higgs peut se désintégrer selon plusieurs voies dont les rapports de branchement sont donnés sur la Figure 1.3. Ces désintégrations vont donner accès à de nouveaux couplages tels que celui au quark  $b$ , la désintégration principale du boson de Higgs de 125 GeV étant en paire de quarks  $b\bar{b}$ . À cette masse, la totalité des voies de désintégration du boson de Higgs sont accessibles avec des rapports de branchement suffisamment grands pour être observés. Les désintégrations du boson de Higgs sont les seules voies pour étudier les couplages de Yukawa du Higgs aux fermions à l'arbre (principalement aux quarks  $b$  et  $c$ , aux leptons  $\tau$  et muons) excepté pour le couplage du boson de Higgs au quark  $top$  qui est mesurable au LHC par l'étude de la production associée  $t\bar{t}H$ .

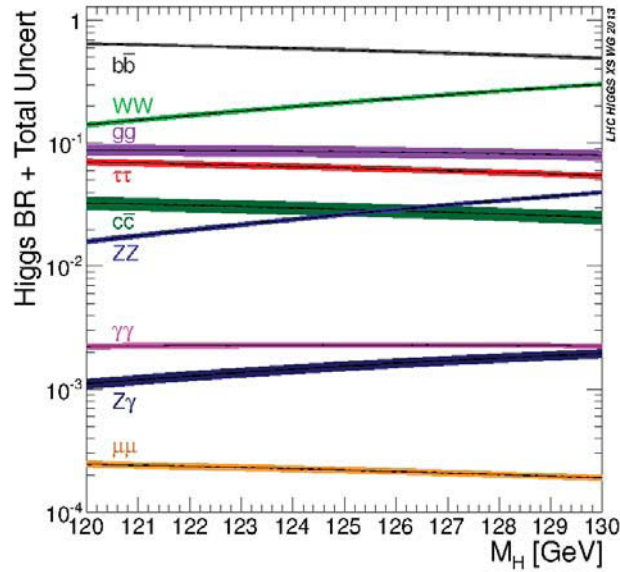


FIGURE 1.3 – Rapports de branchement du boson de Higgs selon les différentes voies de désintégration possibles pour une masse autour de 125 GeV.

### 1.3.3 Les couplages du boson de Higgs

Les premières mesures et estimations des couplages du boson de Higgs aux autres particules par ATLAS sont compatibles avec les prévisions du modèle standard. L'ensemble des résultats décrits dans cette section sont présentés dans [21]. Dans tous les cas le couplage du boson de Higgs avec au moins deux particules entre en jeu et les voies de production et désintégration du boson de Higgs ne peuvent être étudiées indépendamment. Les couplages du boson de Higgs sont estimés par la mesure des facteurs d'échelle,  $\kappa_i$ , appliqués aux couplages prévus dans le modèle standard. Ainsi, dans le cas du modèle standard cette grandeur vaut 1. Ces facteurs  $\kappa_i$  jouent un rôle dans la section efficace  $\sigma(i \rightarrow H \rightarrow f)$  qui peut s'écrire :

$$\sigma(i \rightarrow H \rightarrow f) = \frac{\sigma_i(\kappa_j)\Gamma_f(\kappa_j)}{\Gamma_H(\kappa_j)} \quad (1.23)$$

avec  $\sigma_i$  la section efficace de production du boson de Higgs selon la voie  $i$ ,  $\Gamma_f$  la largeur partielle de désintégration du boson de Higgs dans la voie  $f$  et  $\Gamma_H$  la largeur totale de désintégration du boson de Higgs,  $\kappa_j$  est l'échelle de couplage du boson de Higgs aux particules qui vont participer au processus de production ou désintégration. Une modification de ces facteurs aura donc un impact sur les sections efficaces de production du Higgs et sa probabilité de se désintégrer selon certaines

voies. La Table 1.3 indique pour chaque voie de production et de désintégration du boson de Higgs l'expression de la section efficace, et de la largeur de désintégration en fonction du facteur de couplage  $\kappa$  correspondant. Dans le cas des désintégrations ou des productions du boson de Higgs qui se font en passant par une boucle, un facteur de couplage effectif peut aussi être considéré.

Production	Boucle	Expression selon les échelles de couplages	
$\sigma(\text{ggF})$	Oui	$\kappa_g^2 \sim$	$1.06 \cdot \kappa_t^2 + 0.01 \cdot \kappa_b^2 - 0.07 \cdot \kappa_t \kappa_b$
$\sigma(\text{VBF})$	-	$\sim$	$0.74 \cdot \kappa_W^2 + 0.26 \cdot \kappa_Z^2$
$\sigma(\text{WH})$	-	$\sim$	$\kappa_W^2$
$\sigma(q\bar{q} \rightarrow ZH)$	-	$\sim$	$\kappa_Z^2$
$\sigma(\text{gg} \rightarrow ZH)$	Oui	$\kappa_{\text{ggZH}}^2 \sim$	$2.27 \cdot \kappa_Z^2 + 0.37 \cdot \kappa_t^2 - 1.64 \cdot \kappa_Z \kappa_t$
$\sigma(b\bar{b}H)$	-	$\sim$	$\kappa_b^2$
$\sigma(t\bar{t}H)$	-	$\sim$	$\kappa_t^2$
$\sigma(\text{gb} \rightarrow WtH)$	-	$\sim$	$1.84 \cdot \kappa_t^2 + 1.57 \cdot \kappa_W^2 - 2.41 \cdot \kappa_W \kappa_t$
$\sigma(\text{qb} \rightarrow tHq')$	-	$\sim$	$3.4 \cdot \kappa_t^2 + 3.56 \cdot \kappa_W^2 - 5.96 \cdot \kappa_W \kappa_t$
Largeur de désintégration partielle			
$\Gamma_{b\bar{b}}$	-	$\sim$	$\kappa_b^2$
$\Gamma_{WW}$	-	$\sim$	$\kappa_W^2$
$\Gamma_{ZZ}$	-	$\sim$	$\kappa_Z^2$
$\Gamma_{\tau\tau}$	-	$\sim$	$\kappa_\tau^2$
$\Gamma_{\mu\mu}$	-	$\sim$	$\kappa_\mu^2$
$\Gamma_{\gamma\gamma}$	Oui	$\kappa_{\gamma^2} \sim$	$1.59 \cdot \kappa_W^2 + 0.07 \cdot \kappa_t^2 - 0.66 \cdot \kappa_W \kappa_t$
$\Gamma_{Z\gamma}$	Oui	$\kappa_{Z\gamma}^2 \sim$	$1.12 \cdot \kappa_W^2 + 0.00035 \cdot \kappa_t^2 - 0.12 \cdot \kappa_W \kappa_t$
Largeur de désintégration totale			
$\Gamma_H$	Oui	$\kappa_H^2 \sim$	$0.57 \cdot \kappa_b^2 + 0.22 \cdot \kappa_W^2 + 0.09 \cdot \kappa_g^2 +$ $0.06 \cdot \kappa_\tau^2 + 0.03 \cdot \kappa_Z^2 + 0.03 \cdot \kappa_c^2 +$ $0.0023 \cdot \kappa_\gamma^2 + 0.0016 \cdot \kappa_{Z\gamma}^2 + 0.00022 \cdot \kappa_\mu^2$

TABLE 1.3 – Dépendance des sections efficaces de production du boson de Higgs et de ses largeurs de désintégration en fonction des facteurs de couplage  $\kappa$ . La présence ou non de boucle dans les processus est aussi indiquée. Les valeurs correspondent à une énergie dans le centre de masse de 8 TeV et une masse du boson de Higgs de 125.36 GeV.

A partir des relations entre les sections efficaces de production et les largeurs de désintégration du boson de Higgs mesurées, il est possible de remonter à ces facteurs de couplage et de les comparer à la prédiction du modèle standard. Toutefois, pour mesurer l'un de ces facteurs différentes hypothèses doivent être faites sur les autres paramètres du Modèle.

Dans un premier temps, l'hypothèse est faite que tous les facteurs de couplage des fermions sont égaux à une valeur  $\kappa_f$ . La même hypothèse peut être faite pour les bosons avec une valeur  $\kappa_V$ . Sous ces contraintes, les valeurs observées sont  $\kappa_f = 1.11 \pm 0.16$  et  $\kappa_V = 1.09 \pm 0.07$  [21], compatibles avec la prédiction du modèle standard.

## 1.4 Le couplage du quark *top* au boson de Higgs

Parmi les propriétés du boson de Higgs, son couplage de Yukawa au quark *top* est particulièrement intéressant car il peut donner des informations sur l'échelle d'énergie à laquelle de la Nouvelle Physique pourrait être observée. Le couplage de Yukawa du boson de Higgs au quark *top* est le plus important des couplages du boson de Higgs aux fermions dans le modèle standard et sa valeur affecte fortement d'autres paramètres du Modèle tels que l'autocouplage du boson de Higgs [22]. Ainsi la constante de couplage de Yukawa  $y_t$  du quark *top* au boson de Higgs joue un rôle dans

l'évolution de la constante de couplage trilineaire du boson de Higgs  $\lambda$  et selon la valeur du couplage  $y_t$  on peut trouver une échelle en énergie  $\mu_0$  telle que la constante de couplage  $\lambda$  s'annule. L'existence d'une telle valeur indiquerait que le vide électrofaible est instable contredisant l'observation. L'évolution de la constante  $\lambda$  en fonction de la constante de Yukawa  $y_t$  est illustrée sur la Figure 1.4.

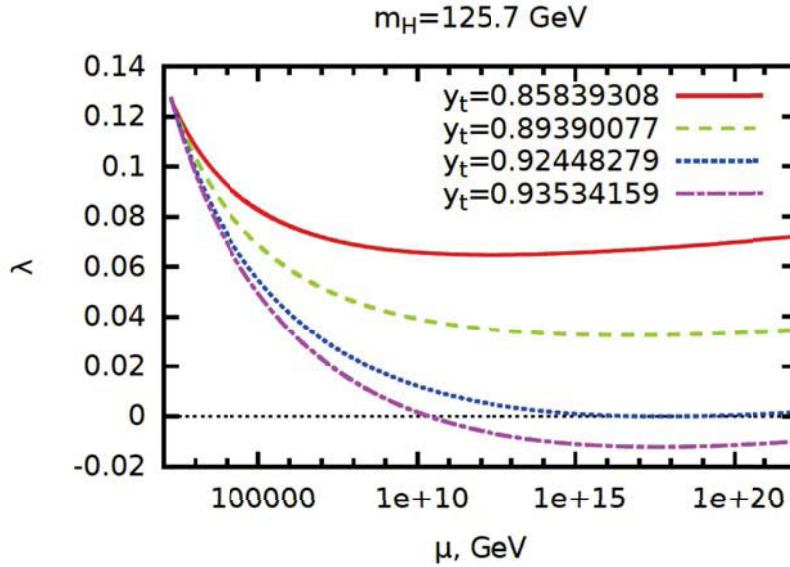


FIGURE 1.4 – Evolution de la constante de couplage du boson de Higgs avec lui-même pour un boson de Higgs de masse  $M_H = 125.7$  GeV et différentes valeurs de la constante de couplage de Yukawa du boson de Higgs au quark *top* [22].

#### 1.4.1 L'observation directe du couplage du boson de Higgs au quark *top*

Le couplage du boson de Higgs au quark *top* est présent dans le principal mode de production du boson de Higgs par fusion de gluons. La boucle de fermions étant majoritairement des quarks *top* dans le modèle standard à cause du couplage important de cette particule au boson de Higgs. Ensuite, l'observation de la désintégration du boson de Higgs en deux photons offre la possibilité d'une mesure indirecte du couplage du boson de Higgs au quark *top* par la présence de boucles de quarks *top* dans le diagramme correspondant. Cette voie de désintégration soulève le problème de la dépendance des mesures à une hypothèse de Modèle. En effet, la sensibilité de la boucle à des phénomènes non prévus dans le modèle standard pouvant s'y compenser ne permet pas une mesure du couplage du quark *top* au boson de Higgs indépendante du modèle. L'étude de la production associée  $t\bar{t}H$  offre l'opportunité d'une mesure directe de ce couplage par la mesure de la section efficace de production.

Plusieurs états finals peuvent faire suite à la production associée  $t\bar{t}H$ . Les désintégrations possibles du boson de Higgs ont été abordées dans la partie précédente et de même le quark *top* va se désintégrer quasi-exclusivement en un boson  $W$  associé à un quark  $b$  dans plus de 99% des cas. Le boson  $W$  peut ensuite se désintégrer dans deux tiers des cas en une paire de quarks  $q\bar{q}$  et dans un tiers des cas en lepton et neutrino ( $l, \nu_l$ ). Quatre catégories d'états finals sont donc disponibles pour un boson de Higgs associé à une paire  $t\bar{t}$ .

- Un état final avec une désintégration du boson de Higgs en une paire de quarks  $b$ .
- Un état final leptonique avec les désintégrations des bosons  $W^\pm$  et  $Z$  venant du boson de Higgs en leptons ou la désintégration du boson de Higgs en paire  $\tau^+\tau^-$  et des désintégrations

leptoniques des quarks *top*. Un tel état final est dit multilepton.

- Un état final avec une désintégration du boson de Higgs en deux photons.
- Un état final avec une désintégration des bosons  $W^\pm$  et  $Z$  venant du boson de Higgs en quarks légers ( $u, d, s$  ou  $c$ ).

### 1.4.2 Les états finals leptoniques du processus de production associé $t\bar{t}H$

La suite de ce document présente la recherche du processus  $t\bar{t}H$  par l'étude des états finals leptoniques et en particulier l'état final avec deux leptons de même charge électrique sans  $\tau$  hadronique. Pour décrire ces états finals il faut prendre en compte les probabilités de désintégrations des bosons  $W, Z$  et du lepton  $\tau$  en leptons ou hadrons<sup>6</sup>. Ces valeurs sont donnés dans la Table 1.4. Les états finals leptoniques considérés sont :

- deux leptons (électrons ou muons) de même charge électrique avec ou sans  $\tau$  hadronique (notés  $2\ell SS1\tau$  et  $2\ell SS0\tau$ )
- trois leptons avec ou sans  $\tau$  hadronique (notés  $3\ell 1\tau$  et  $3\ell 0\tau$ )
- quatre leptons (noté  $4\ell$ )
- deux leptons de charges électriques opposées avec un  $\tau$  hadronique (noté  $2\ell OS1\tau$ )
- un lepton et deux  $\tau$  hadroniques de charges électriques opposées (noté  $1\ell 2\tau$ )

Les voies de désintégrations du boson de Higgs conduisant à ces états finals sont  $H \rightarrow WW, H \rightarrow ZZ$  et  $H \rightarrow \tau\tau$ . Une caractéristique supplémentaire de ces états finals leptoniques est la présence de deux quarks  $b$  provenant des désintégrations des quarks *top* et de quarks supplémentaires provenant des désintégrations des quarks *top* et du Higgs.

TABLE 1.4 – Probabilité de désintégration des bosons  $W$  et  $Z$  et du lepton  $\tau$  en leptons,  $\tau$  ou hadrons [1].

Voie de désintégration	W	Z	$\tau$
$e$ ou $\mu$	21%	7%	35%
$\tau$	11%	3%	–
Hadrons (ou invisible)	68%	90%	65%

Dans le cas d'un état final avec deux leptons de même charge électrique on attend deux quarks  $b$  et six quarks dans le cas d'une désintégration du boson de Higgs en une paire  $W^+W^-$  telle que décrite sur la Figure 1.5.

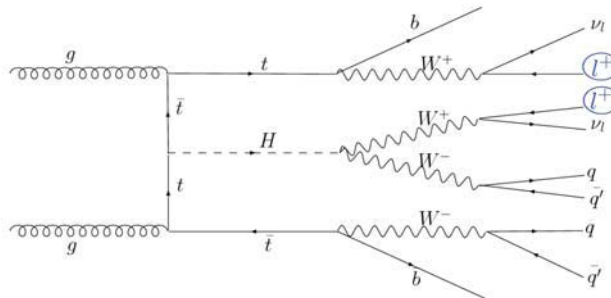


FIGURE 1.5 – Diagramme de désintégration des quarks *top* et du boson de Higgs aboutissant à un état final avec deux leptons de même charge électrique.

Les états finals leptoniques du processus  $t\bar{t}H$  font intervenir différents couplages. On a pour la section efficace de production  $\sigma(t\bar{t}H) \sim \kappa_t^2$  tandis que les voies de désintégration du boson de Higgs

6. Dans le cas du lepton  $\tau$ , on distingue  $\tau$  leptonique pour une désintégration en lepton et neutrinos et  $\tau$  hadronique pour une désintégration vers des hadrons.



donnent accès à  $\kappa_W$ ,  $\kappa_Z$  et  $\kappa_\tau$ . Ainsi dans le cas où les particules sont parfaitement reconstruites dans le détecteur, les voies de désintégration concernées par chaque état final et leur participation relative sont décrites dans la Table 1.5. Les deux voies de désintégration les plus importantes sont  $H \rightarrow WW$  et  $H \rightarrow \tau\tau$ .

TABLE 1.5 – Contribution des différents modes de désintégration du boson de Higgs pour les différents états finals  $t\bar{t}H$ .

Signature	$H \rightarrow WW$	$H \rightarrow \tau\tau$	$H \rightarrow ZZ$
$2\ell SS0\tau$	100%	–	–
$3\ell$	74%	21%	5%
$4\ell$	52%	29%	19%
$1\ell 2\tau$	18%	76%	6%
$2\ell SS1\tau$	34%	64%	2%
$3\ell 1\tau$	32%	60%	8%

Ainsi, selon l'état final considéré, les constantes de couplage du Higgs avec le boson W, le boson Z ou le fermion  $\tau$  sera accessible en association avec le couplage du Higgs au quark  $t$ . Dans l'environnement du détecteur ATLAS, la reconstruction des objets n'étant pas parfaite chacune des voies de désintégration participera quelque soit l'état final, de même que la désintégration du boson de Higgs en paire  $b\bar{b}$ .

### 1.4.3 Les processus multileptons aux LHC

Une problématique liée à l'observation du processus  $t\bar{t}H$  au LHC est la présence d'autres processus prédits par le modèle standard et qui vont aussi produire des états finals avec plusieurs leptons.

Le premier de ces processus est la **production d'une paire  $t\bar{t}$** . Ce processus est caractérisé par un état final avec deux bosons W de signes opposés et deux quarks  $b$ . Les désintégrations de la paire  $t\bar{t}$  peuvent être :

- hadronique : les deux bosons W se désintègrent en quarks.
- semi-leptonique : l'un des deux bosons W se désintègre en lepton et neutrino, le second en quarks.
- dileptonique : les deux bosons W se désintègrent en une paire lepton-neutrino.

Les modes de production de paire  $t\bar{t}$  au LHC à l'arbre sont présentés sur la Figure 1.6. Un second

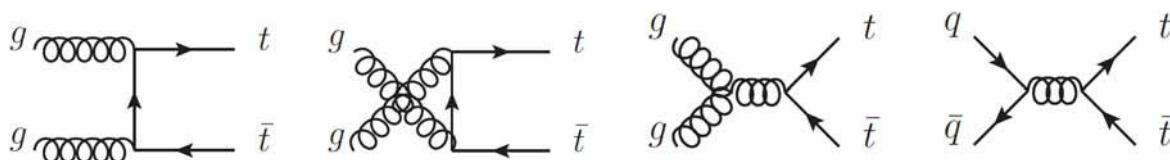


FIGURE 1.6 – Diagrammes de production LO de paire de quarks *top-antitop* au LHC.

processus donnant des événements avec plusieurs leptons au LHC est la **production de deux bosons de jauge**. Les diagrammes correspondants aux différents processus  $W^\pm W^\pm$ ,  $W^\pm Z$ ,  $ZZ$  sont présentés sur la Figure 1.7. Les états finals de ces processus peuvent avoir un nombre de leptons importants, toutefois la principale différence avec le processus  $t\bar{t}H$  est l'absence de quark *top*.

Un autre processus pouvant donner des événements avec plusieurs leptons est la **production associée d'une paire  $t\bar{t}$  et de bosons de jauge**. Les diagrammes correspondant à ces bruits de

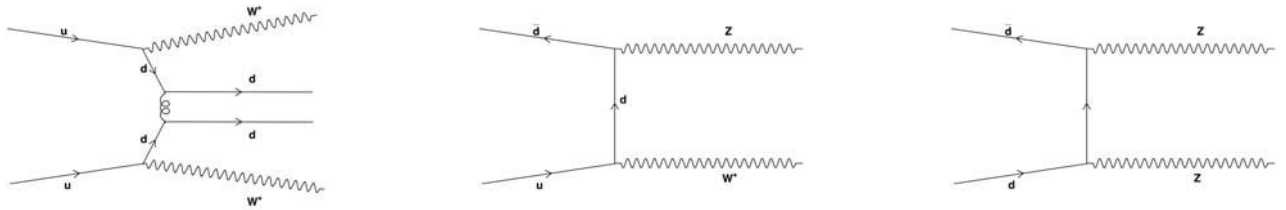


FIGURE 1.7 – Diagrammes de production associée d’une paire de bosons W (a), d’une paire de boson W et Z (b) ou d’une paire de bosons Z (c).

fond sont présentés sur la Figure 1.8. Ces deux processus sont très proches du processus  $t\bar{t}H$  par la présence de deux quarks  $top$  mais aussi par la présence d’un boson vecteur supplémentaire.

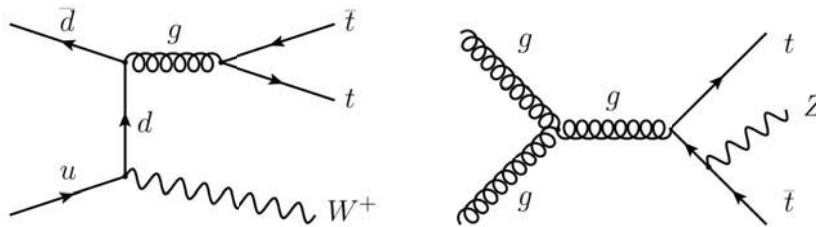


FIGURE 1.8 – Diagrammes de production associée d’une paire  $t\bar{t}$  et d’un boson W (a) ou Z (b).

# Outils expérimentaux : Le LHC et le détecteur ATLAS

## 2.1 Introduction

Depuis les années 1960 de nombreuses preuves de la justesse du modèle standard ont été apportées par la découverte de particules précédemment prédites telles que les bosons  $W^\pm$ ,  $Z$  ou le boson de Higgs. La poursuite des tests du modèle standard et la recherche de Nouvelle Physique nécessite aujourd'hui un accélérateur-collisionneur capable d'atteindre de très hautes énergies ainsi que des détecteurs assez précis pour étudier les particules produites lors de ces collisions. A l'heure actuelle l'accélérateur-collisionneur utilisé pour cette recherche est le LHC [23] associé aux détecteurs ATLAS, CMS et LHCb.

Dans une première partie le LHC est présenté avec ses caractéristiques spécifiques à la période de collision 2015-2016. Un aperçu de la phénoménologie associée aux collisions proton-proton est également présenté. Le détecteur ATLAS avec ses différents sous-détecteurs sont décrits dans la section 2.3.

## 2.2 Le LHC et les collisions proton-proton

### 2.2.1 Le grand collisionneur de hadrons (LHC)

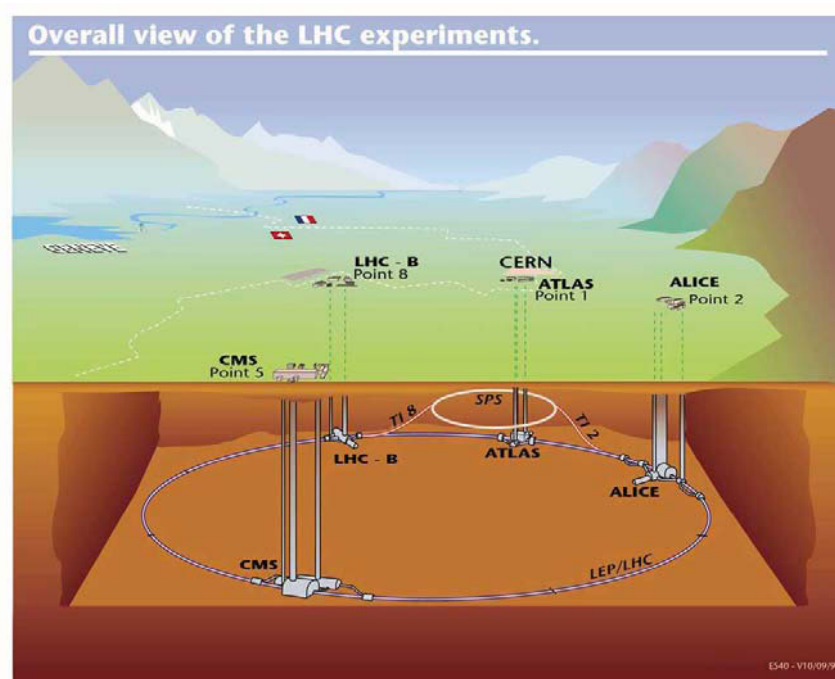


FIGURE 2.1 – Le LHC et les expériences associées.

Le LHC (Large Hadron Collider) est un collisionneur proton-proton (pp) installé environ 100 m sous terre au CERN près de Genève en fonctionnement depuis 2008. Il a remplacé le LEP [24], collisionneur électron-positron, et utilise certaines des infrastructures de ce dernier, tel le tunnel où il est installé. L'énergie dans le centre de masse maximale prévue pour le LHC est de 14 TeV pour une luminosité supérieure à  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ .

Le LHC est constitué de deux faisceaux de protons circulant dans des directions opposées qui se rejoignent en quatre points de collision le long des 27 km de tunnel tel que décrit sur la Figure 2.1. Un détecteur est installé à chacun de ces points :

- ALICE (A LHC Ion Collider Experiment) qui a pour but d'étudier les collisions avec des ions lourds au LHC (1 mois par an) [25].
- ATLAS (A Toroidal LHC ApparatuS) qui est un détecteur polyvalent dont les buts premiers sont l'étude du boson de Higgs et la recherche de Nouvelle Physique [26].
- CMS (Compact Muon Solenoid) qui est comme ATLAS un détecteur polyvalent [27].
- LHCb qui travaille sur la violation de matière/antimatière et cherche la Nouvelle Physique par des mesures de précision du modèle standard [28].

## 2.2.2 Le circuit d'accélération et les propriétés du LHC

Le LHC est le dernier élément d'une chaîne d'accélération reposant entre autres sur de précédents accélérateurs construits au CERN. La chaîne complète d'accélérateurs linéaire et circulaires est présentée sur la Figure 2.2.

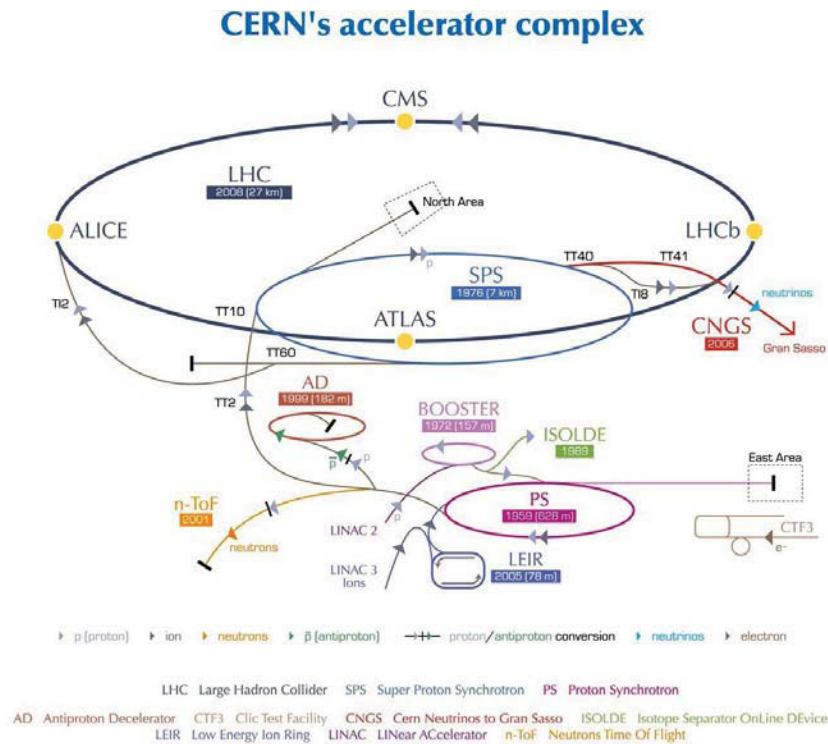


FIGURE 2.2 – Complexe des accélérateurs du CERN.[29]

Les protons utilisés dans les collisions sont obtenus par ionisation d'atomes d'hydrogène. La première phase d'accélération est réalisée grâce au LINAC 2, un accélérateur linéaire, qui amène les protons à une énergie de 50 MeV. Les protons sont ensuite accélérés par différents accélérateurs circulaires, le Proton Synchrotron Booster (PSB), le Proton Synchrotron (PS) et enfin le Super Proton Synchrotron (SPS) qui les amènent respectivement à des énergies de 14 GeV, 25 GeV et finalement 450 GeV. Les protons sont ensuite séparés en deux faisceaux circulant dans deux tubes à vide dans

des directions opposées dans le LHC et accélérés jusqu'à une énergie de 6.5 TeV en 2015-2016. Pour atteindre une telle énergie, les protons sont accélérés grâce un système de 8 cavités radio-fréquence par faisceau. La trajectoire des faisceaux est assurée par un système d'aimants supraconducteurs émettant un champ magnétique intégré de 8.6 T : 1232 aimants dipolaires pour imposer la courbure de la trajectoire et 392 aimants quadrupolaires pour conserver la focalisation des faisceaux. Les protons atteignent ainsi une vitesse proche de celle de la lumière en une vingtaine de minutes. En moyenne, un faisceau au LHC en 2016 était composé d'environ 2200 paquets contenant chacun  $1.2 \times 10^{11}$  protons, chaque paquet étant séparé du suivant de 25 ns.

### 2.2.3 Les collisions protons-protons au LHC

#### 2.2.3.1 Luminosité

Un collisionneur de particules est caractérisé par le nombre de collisions par seconde, grandeur appelée luminosité instantanée  $L$  définie en  $\text{cm}^{-2}\text{s}^{-1}$ . Elle peut être reliée à la section efficace de production d'un processus donné par :

$$N_{events} = L\sigma_{events}, \quad (2.1)$$

où  $N_{events}$  est le nombre d'événements produits par seconde et  $\sigma_{events}$  la section efficace du processus considéré. L'expression de  $L$  ne dépend quant à elle que des propriétés intrinsèques des faisceaux de protons :

$$L = \frac{N_b^2 n_b f_{rev} \gamma}{4\pi \varepsilon_n \beta^*} F, \quad (2.2)$$

où  $N_b$  est le nombre de protons par paquet,  $n_b$  le nombre de paquets par faisceau,  $f_{rev}$  la fréquence de révolution,  $\gamma$  le facteur de Lorentz,  $\varepsilon_n$  l'émittance transverse du faisceau,  $\beta$  la fonction beta au point de collision qui quantifie l'amplitude des oscillations des protons et  $F$  un facteur géométrique relié à l'angle de croisements des faisceaux au point de collision. La luminosité instantanée, pour un nombre de paquet et une fréquence de révolution constants, varie en fonction des collisions principalement à cause de la variation du nombre de protons par paquets et la valeur maximale de cette grandeur pour les collisions 2015-2016 est reportée en fonction du temps sur la Figure 2.3.

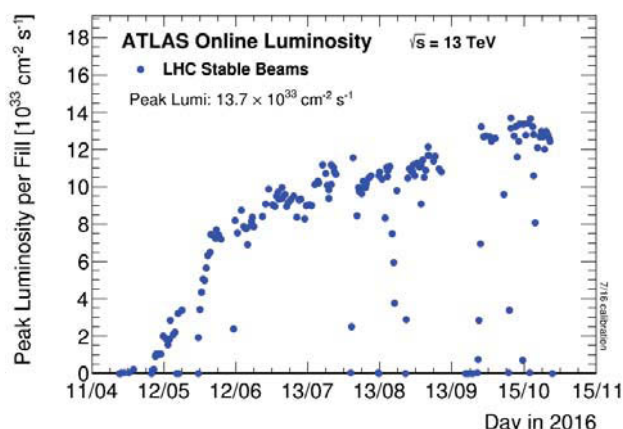


FIGURE 2.3 – Pic de luminosité instantanée par jour en 2016.[30]

Il est possible de définir la luminosité intégrée  $\mathcal{L}$  comme étant le résultat de l'intégration en temps de la luminosité instantanée et est notée en femtobarns inverse ( $fb^{-1}$ ). Cette grandeur permet de quantifier la quantité totale de données délivrées par le LHC ou la quantité de données enregistrées par un détecteur tel que ATLAS. Ces deux éléments sont représentés sur la figure 2.4.

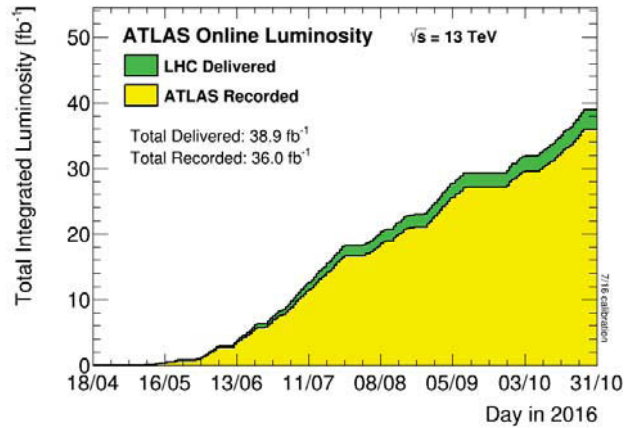


FIGURE 2.4 – Luminosité intégrée délivrée lors de collisions pp au LHC (vert) et enregistrée par ATLAS (jaune) en 2016.[30]

### 2.2.3.2 Les événements de biais minimum

Parmi les très nombreux croisements de paquets de protons au LHC, les interactions proton-proton sont dominées par les collisions inélastiques. Les événements produits sont caractérisés par une basse impulsion transverse et sont appelés événements de biais minimum. Ces événements sont utilisés pour mesurer la luminosité intégrée enregistrée dans ATLAS.

### 2.2.3.3 Effet d'empilement d'événements

Le grand nombre de particules par paquet signifie que plusieurs collisions par croisements peuvent avoir lieu. Ce phénomène est appelé effet d'empilement et augmente avec la luminosité instantanée. Le nombre moyen d'interactions par croisement de faisceaux pour 2016 est présenté sur la figure 2.5.

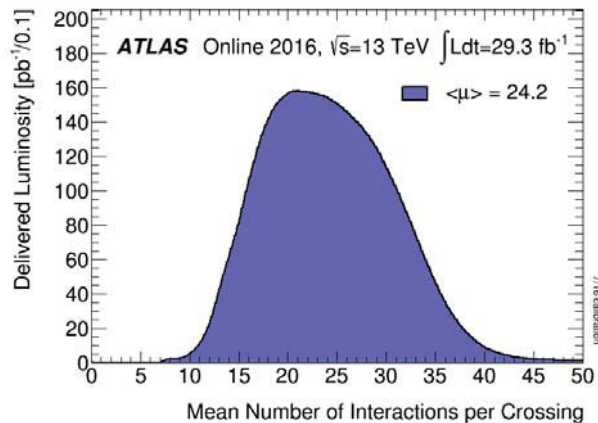


FIGURE 2.5 – Distribution du nombre moyen d'interactions par croisement de faisceau,  $\mu$ , en 2016.[30]

## 2.3 Le détecteur ATLAS (A Toroidal LHC ApparatuS)

ATLAS est l'un des quatre détecteurs situés sur le LHC et le plus imposant. Il mesure 45 m de long, 25 m de haut pour un poids total de 7000 tonnes. Pour pouvoir observer les différentes

particules qui le traversent, c'est un détecteur couvrant presque  $4\pi$  sr formé d'une partie tonneau (central) et deux bouchons. A l'instar d'autres détecteurs de physique des particules il se compose de différents sous-détecteurs spécialisés superposés autour du point d'interaction qui seront décrits par la suite. Le détecteur interne, partie la plus proche du faisceau, a pour but de mesurer la trajectoire des particules chargées. Les calorimètres permettent la mesure de l'énergie de toutes les particules (sauf muons et neutrinos). Finalement les muons, qui traversent les sous détecteurs précédents, vont être détectés dans les spectromètres à muons qui complètent la mesure de leur impulsion.

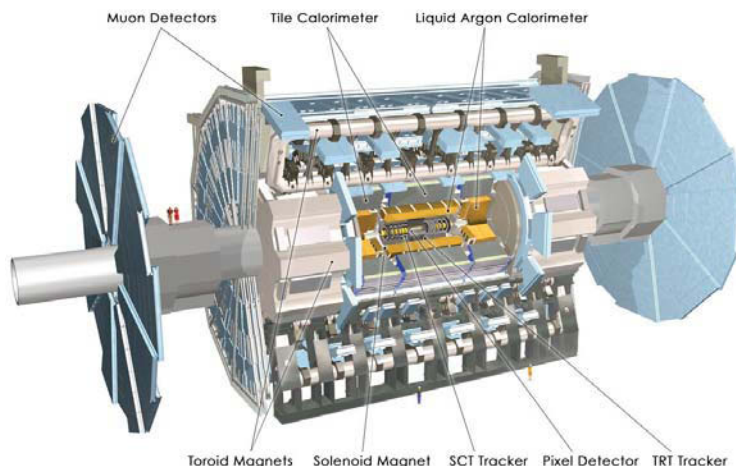


FIGURE 2.6 – Schéma du détecteur ATLAS.

### 2.3.1 Conventions géométriques

Des conventions spécifiques sont utilisées pour décrire les quantités géométriques à travers le détecteur et les différentes sous-couches. L'axe  $z$  est placé suivant l'axe du faisceau et les coordonnées  $x$  et  $y$  sont décidées de façon à avoir un repère direct. L'angle  $\phi$  est défini dans le plan dit transverse  $(O,x,y)$  comme étant l'angle par rapport à l'axe  $y$ . L'angle  $\theta$  est l'angle par rapport à l'axe  $z$ . Toutes ces grandeurs sont indiquées sur la Figure 2.7. L'utilisation de la pseudorapidité  $\eta$  est préférée à celle de l'angle  $\theta$ . Cette grandeur est calculée d'après l'équation suivante :

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right). \quad (2.3)$$

Un intervalle en  $\theta$  de  $\pi/16$  à  $\pi/2$  correspond à un intervalle en  $\eta$  de  $-2.5$  à  $0$ . A partir de la définition de ces deux angles,  $\eta$  et  $\phi$ , on définit la distance entre deux particules comme étant :

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.4)$$

où  $\Delta\eta$  est la différence entre les pseudorapidités et  $\Delta\phi$  la différence entre les angles  $\phi$ .

On utilisera par la suite la notion de grandeurs transverses, telle que l'impulsion transverse qui est la projection du vecteur impulsion dans le plan transverse. En effet, les protons n'étant pas des particules élémentaires seule l'impulsion totale initiale dans le plan transverse est connue (elle est nulle). Une impulsion transverse finale non-nulle est un indice de la présence d'un neutrino produit lors de la collision puisque la valeur de l'impulsion transverse est conservée lors de la collision.

La notion de plan  $(R-\phi)$  est aussi utilisée et un objet dans ce plan a des coordonnées en fonction de sa distance au faisceau ( $R$ ) et de l'angle  $\phi$ .

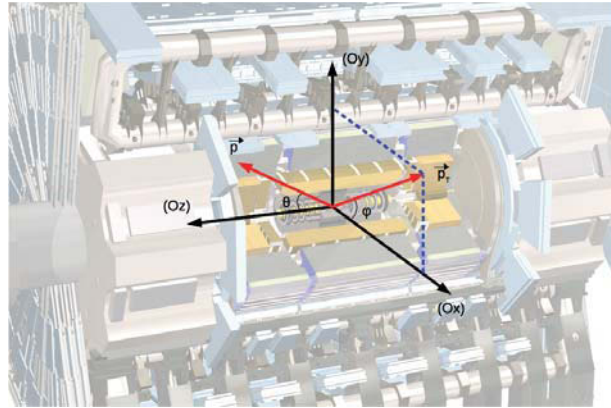


FIGURE 2.7 – Système de coordonnées utilisé dans ATLAS. Le vecteur  $\vec{p}$  représente l'impulsion de la particule alors que le vecteur  $\vec{p}_T$  représente son impulsion transverse.

### 2.3.2 Le détecteur interne

Le détecteur interne [31][31] est un trajectographe, il a pour but de reconstruire la trajectoire des particules chargées. La présence d'un champ magnétique va permettre de courber la trajectoire de la particule observée (si cette particule est chargée). Son impulsion et sa charge pourront être obtenues à partir du rayon de courbure. De plus ce sous-détecteur permet de reconstruire les vertex primaires et secondaires de la collision. Le vertex primaire est l'endroit de l'interaction entre les constituants des deux protons des faisceaux qui ont initialement interagi. Les vertex secondaires sont les endroits où des particules ayant des temps de vie assez longs pour qu'elles puissent se déplacer sur des distances observables dans le détecteur se sont désintégrées. Ce dernier point concerne particulièrement les particules belles<sup>1</sup> et l'existence de ces vertex secondaires joue un rôle important dans le *b-tagging*, c'est-à-dire l'étiquetage d'une particule belle dans l'événement.

Ce détecteur interne est constitué de trois sous-détecteurs qui apparaissent sur la Figure 2.8 : un détecteur à pixels, un détecteur à micropistes de silicium (SCT) et un trajectographe à rayonnement de transition (TRT).

Le premier est constitué de plus de 90 millions de pixels disposés sur quatre couches cylindriques autour de l'axe du faisceau. La première de ces couches en est située à 3.3 cm. Ces quatre couches permettent d'obtenir au plus quatre points pour chaque particule chargée venant du point d'impact. Le SCT est quant à lui constitué de quatre couches de micropistes de silicium. Grâce à celles-ci, quatre points supplémentaires par particule chargée devraient être visibles.

Enfin le TRT est constitué à la fois de brins de polypropylène et de gaz (mélange de Xénon, dioxygène et dioxyde de carbone) et son but est de différencier les traces venant d'électrons et celles venant de pions chargés<sup>2</sup>, cela en détectant les photons émis par les électrons et les pions lors de leur changement de milieu.

#### 2.3.2.1 Détecteurs à pixels

Le détecteur à pixels [33] est le sous-détecteur le plus proche du faisceau situé à 3.3 cm de son axe au point le plus proche. Il est constitué de quatre couches de jonctions PN polarisées pour un total de plus de 90 millions de pixels. Il s'étend jusqu'à 12 cm du faisceau. Ainsi lors du passage d'une particule chargée à travers les pixels une paire électron-trou est créée et la présence d'un champ électrique permet la production d'un signal mesurable. La résolution sur la position de la

1. Les particules belles sont des particules provenant de l'hadronisation de quarks *b*.

2. Les pions sont des mésons constitués d'un quark *u* et d'un antiquark  $\bar{d}$  ( $\pi^+$ ) ou d'un antiquark  $\bar{u}$  et d'un quark *d* ( $\pi^-$ ).



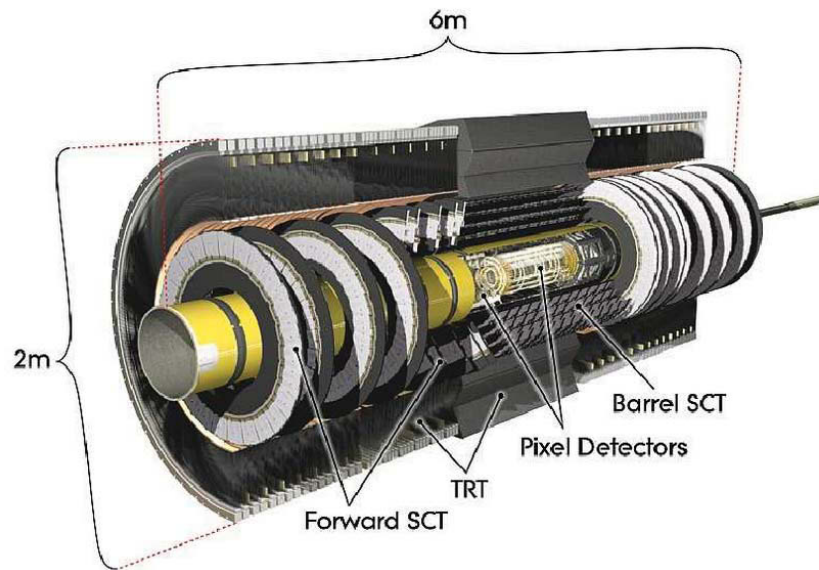


FIGURE 2.8 – Les différents composants du détecteur interne d'ATLAS [32].

particule est alors de  $8\mu\text{m}$  dans le plan ( $R-\phi$ ) et  $40\mu\text{m}$  dans la direction longitudinale. La présence des quatre couches permet une reconstruction précise de la trajectoire des particules et joue donc un rôle prépondérant dans la reconstruction des vertex et l'étiquetage des jets de saveurs lourdes.

### 2.3.2.2 Détecteur à micropistes de silicium (SCT)

La seconde partie du détecteur interne est le détecteur à micropistes de silicium qui tout comme le détecteur à pixels repose sur la technologie des semi-conducteurs. Il est formé d'une partie tonneau composée de quatre couches cylindriques situées à une distance comprise entre 30 et 52 cm du faisceau et de deux parties bouchons composées de neuf disques orthogonaux à l'axe du faisceau. La résolution sur la position d'une particule dans ce détecteur est de  $17\mu\text{m}$  selon ( $R-\phi$ ) et  $580\mu\text{m}$  dans la direction longitudinale.

### 2.3.2.3 Trajectographe à rayonnement de transition

La couche la plus externe du détecteur interne est le trajectographe à rayonnement de transition. Il est composé de tubes à dérive de 4mm de diamètre remplis d'un mélange gazeux ( $70\% \text{Xe} + 27\% \text{CO}_2 + 3\% \text{O}_2$ ). Un fil plaqué en or-tungstène est placé sur l'axe de révolution de chacun de ces tubes. Le passage d'une particule chargée va ioniser le gaz et les électrons induits vont dériver vers le fil central et ainsi créer un signal électrique. Les tubes sont entourés d'un radiateur composé de fibres de polypropylène/polyéthylène qui produit un rayonnement de transition au passage d'une particule chargée. L'énergie rayonnée est dépendante du facteur de Lorentz de la particule (inversement proportionnel à sa masse) ainsi cette énergie sera beaucoup plus grande pour un électron qu'un pion chargé  $\pi^\pm$  ce qui permet une bonne discrimination entre ces deux particules. Le TRT est composé d'une partie tonneau, qui couvre la région  $|\eta| < 0.7$  avec des tubes longitudinaux placés sur trois couches entre 56 et 107 cm, et de parties bouchons comptant 20 disques de tubes orthogonaux au faisceau. La région ainsi couverte va jusqu'à  $|\eta| < 2.0$ . La résolution intrinsèque du TRT sur la position des particules chargées dans le plan ( $R-\phi$ ) est de  $160\mu\text{m}$ .

### 2.3.2.4 L'aimant solénoïdal

L'aimant solénoïdal [34] est situé entre le détecteur interne et le système calorimétrique, à une distance comprise entre 2.46 m et 2.56 m du faisceau. Il est refroidi à l'hélium à une température de 4.2 K et génère un champ magnétique de 2 T. Ce champ magnétique permet de courber la trajectoire des particules chargées et plus particulièrement des électrons pour obtenir leur charge et impulsion.

### 2.3.3 Les calorimètres

Le système calorimétrique d'ATLAS, présenté sur la Figure 2.9, est situé au delà de l'aimant solénoïdal et couvre la région  $|\eta| < 4.9$ . Ce système absorbe la totalité de l'énergie des particules venant du point d'interaction exception faite des neutrinos et muons et permet la mesure de leur énergie. Ce système calorimétrique peut être divisé en trois parties : le calorimètre électromagnétique, le calorimètre hadronique et le calorimètre vers l'avant.

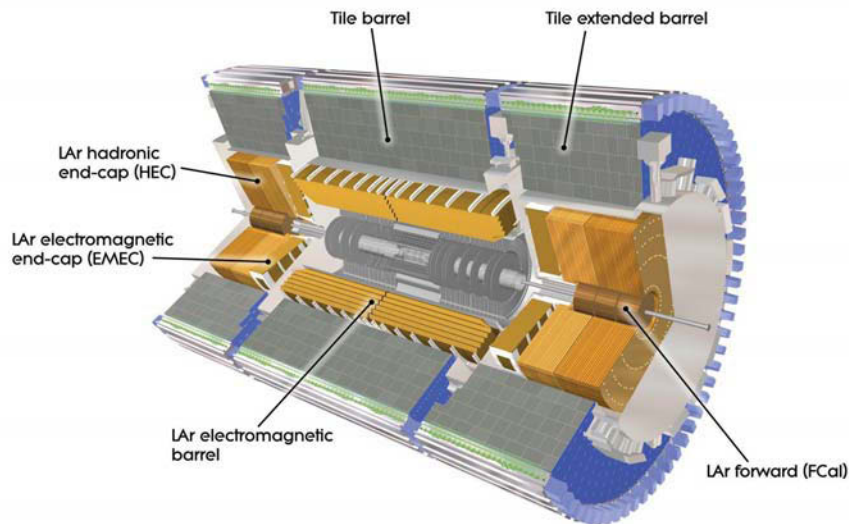


FIGURE 2.9 – Système calorimétrique d'ATLAS [35].

#### 2.3.3.1 Le calorimètre électromagnétique

Le calorimètre électromagnétique (*Ecal*) [36] est la première partie du système calorimétrique. Le but premier de ce système est la mesure de l'énergie des électrons et photons. Il permet en outre la mesure de la composante électromagnétique des jets. C'est un calorimètre à échantillonnage composé d'une partie tonneau centrale et de deux parties bouchons portant l'acceptance jusqu'à  $|\eta| < 3.2$ .

Du plomb est utilisé comme matériau absorbeur alors que la partie active est composée d'argon liquide ce qui impose que le système reste baigné à une température de 88 K. Une particularité de ce sous-détecteur est sa géométrie en accordéon telle que visible sur la Figure 2.10 qui induit une couverture complète en  $\phi$  sans espace mort. Ce calorimètre est composé de trois couches dans la région  $|\eta| < 2.5$  puis de deux dans les régions plus proches du faisceau. Une couche supplémentaire, appelée pré-échantillonneur, est placée avant le calorimètre dans la région  $|\eta| < 1.8$  afin d'estimer l'énergie perdue par les particules dans les matériaux précédents le calorimètre.

La première couche a pour but de reconstruire la direction de la particule entrante et elle a donc une segmentation particulièrement fine en  $\eta$  :  $\eta \times \phi = 0.0031 \times 0.1$ . La seconde couche où la majorité

de l'énergie est déposée a une segmentation de  $\eta \times \phi = 0.025 \times 0.25$ . La dernière couche possède une segmentation plus grossière de  $\eta \times \phi = 0.050 \times 0.25$  et permet d'estimer l'énergie des gerbes non contenues dans les couches précédentes. Ces différents éléments confèrent au calorimètre une résolution en énergie de  $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{(E)}} \oplus 0.7\%$  [26].

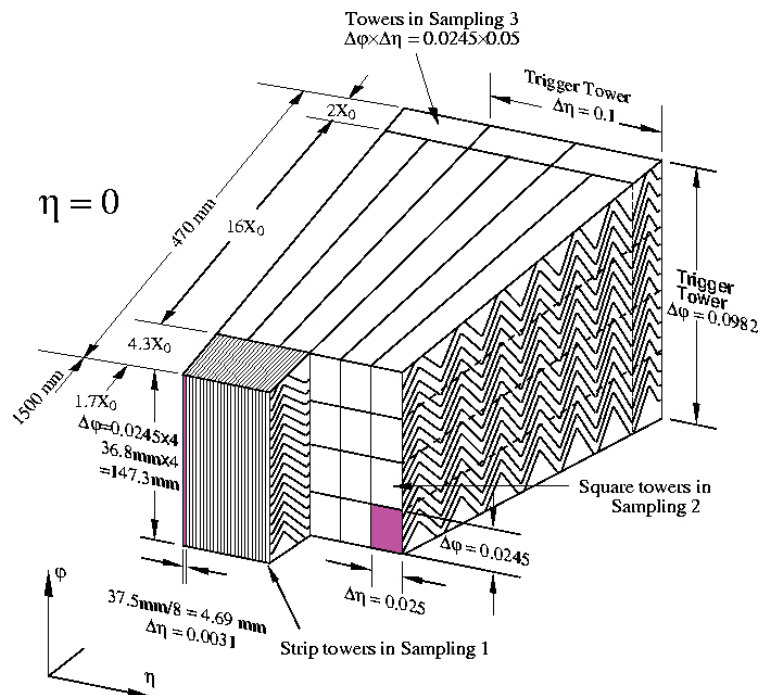


FIGURE 2.10 – Schéma d'un module de la partie tonneau du calorimètre électromagnétique.[26]

### 2.3.3.2 Les calorimètres hadroniques

Le calorimètre hadronique est la couche extérieure au *Ecal* du détecteur ATLAS. Son but est de contenir les gerbes hadroniques et de fournir une mesure de leur énergie. Il participe donc à la reconstruction de l'énergie des jets et de leur position ainsi qu'à la mesure de l'énergie transverse manquante. Tout comme le calorimètre électromagnétique, il est composé d'un tonneau central et de deux bouchons. Toutefois les technologies utilisées dans les deux cas diffèrent.

La partie tonneau est appelé calorimètre hadronique à tuiles scintillantes [37] (ou *TileCal* pour *Tile Calorimeter*). C'est un calorimètre à échantillonnage utilisant du fer comme absorbeur et des tuiles scintillantes en polystyrène comme matériau actif. L'interaction d'une particule dans le fer va produire des particules chargées qui vont ioniser les tuiles. Celles-ci vont alors émettre des photons qui sont acheminés par des fibres optiques jusqu'aux blocs photomultiplicateurs (PMT). Ces derniers sont au nombre de 9852 dans l'ensemble du calorimètre.

Selon l'axe  $z$ , on peut distinguer une partie centrale pour  $|\eta| < 1$  composée de deux tonneaux longs (LBA et LBC) et deux tonneaux étendus (EBA et EBC) qui couvrent la région  $0.7 < |\eta| < 1.7$ . Cette géométrie est présentée sur la figure 2.11. En outre le *TileCal* est découpé en 64 modules disposés autour de l'axe du faisceau. Le schéma d'un module est donné sur la Figure 2.12. Dans chacun de ces modules, des cellules lisent les informations venant de différentes tuiles scintillantes correspondant à des torons de fibres optiques. Il existe trois couches de tuiles scintillantes A, BC et D tel que présenté sur la figure 2.13. Ces trois couches correspondent respectivement à 1.4, 3.9 et 1.8 longueurs d'interaction. Selon ces trois couches la granularité  $\Delta_\eta \times \Delta_\phi$  est de  $0.1 \times 0.1$  pour les deux premières et  $0.2 \times 0.1$  pour la couche D. Finalement des cellules supplémentaires, appelées cellules E, sont placées entre les tonneaux longs et étendus pour évaluer l'énergie perdue par les

particules dans les parties non instrumentées du détecteur.

Il y a 5182 cellules dans le *TileCal* qui récupèrent les informations provenant de 460 000 tuiles scintillantes. Deux PMTs sont associées à chacune de ces cellules<sup>3</sup> excepté pour certaines cellules des tonneaux étendus qui n'ont qu'un PMT. Outre les cellules et les PMTs, le terme de canal sera aussi employé. Un canal est pour le *TileCal* l'ensemble formé par la partie optique avec les tuiles scintillantes et les fibres, une seconde partie correspondant aux PMTs et enfin une partie liée à l'électronique de lecture. Concernant cette électronique de lecture qui permet de traduire le signal électrique en énergie, un élément important est la présence de deux gains (un haut gain avec une amplification d'un facteur 32 du signal et un bas gain avec une amplification d'un facteur 0.5 du signal). Ces deux gains imposent que le nombre de canaux dans le *TileCal* est le double du nombre de PMTs.

Les bouchons du calorimètre hadronique sont construits avec la même technologie d'argon liquide que le calorimètre électromagnétique excepté que l'absorbeur utilisé est du cuivre. Il couvre la région  $|\eta| < 3.2$  et sa granularité est  $\eta \times \phi = 0.1 \times 0.1$  dans la région  $1.5 < |\eta| < 2.5$  et  $\eta \times \phi = 0.2 \times 0.2$  dans la région  $2.5 < |\eta| < 3.2$ .

La résolution sur l'énergie des pions pour l'ensemble du calorimètre hadronique est de :  $\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{(E)}} \oplus 3\%$  [26].

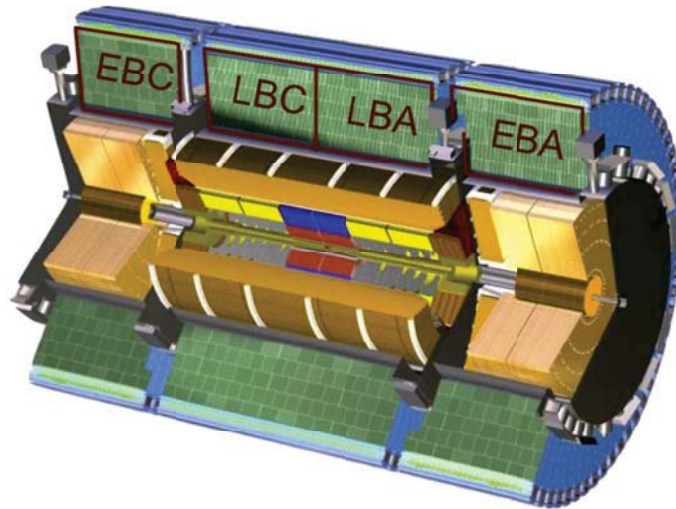


FIGURE 2.11 – Schéma du calorimètre hadronique à tuiles d'ATLAS (en vert) avec les séparations en tonneaux long et étendus indiqués.

### 2.3.3.3 Le calorimètre vers l'avant

La dernière partie du système calorimétrique d'ATLAS est un calorimètre vers l'avant qui couvre la région  $3.2 < |\eta| < 4.9$ . Ce calorimètre assure à la fois le rôle de calorimètre hadronique et électromagnétique et est composé de quatre roues. Il repose sur l'utilisation d'argon liquide comme matériau actif, comme le calorimètre électromagnétique, mais utilise différents absorbeurs et géométries pour faire face aux radiations importantes présents dans cette région du détecteur. Ainsi la première roue qui joue le rôle de calorimètre électromagnétique contient du cuivre tandis que les deux roues suivantes qui sont utilisées comme calorimètre hadronique sont constituées de tungstène comme absorbeur. Enfin la dernière roue n'est pas instrumentée et fait office de blindage pour protéger les chambres à muons qui viennent ensuite. Chaque roue instrumentée est

3. Toutes les fibres des tuiles d'une même cellule et d'un même côté sont reliées à un unique PMT

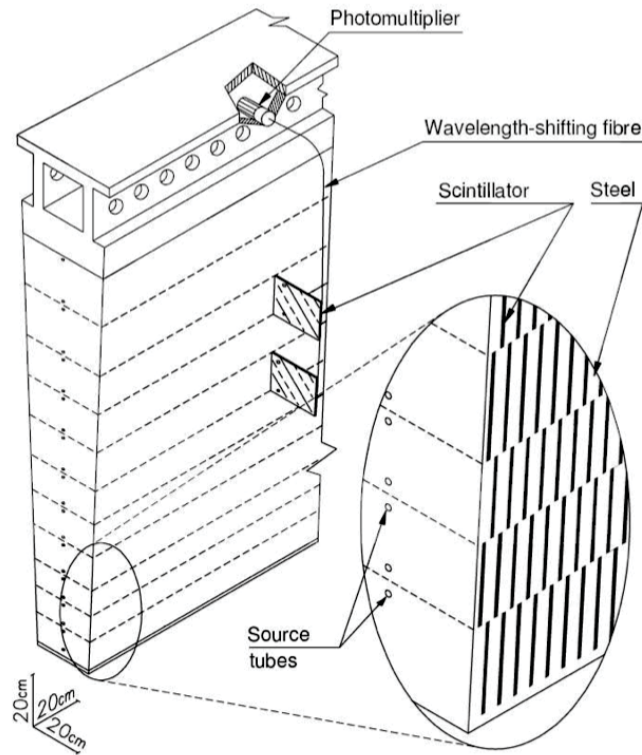


FIGURE 2.12 – Schéma d'un module de la partie tonneau du calorimètre hadronique à tuile. [38]

construite d'une plaque de matériau absorbeur percée de trous de 5mm de diamètre contenant un tube métallique de 4.5 mm de diamètre utilisé comme électrode. L'espace restant est rempli d'argon liquide, cet espace particulièrement réduit favorisant la réduction du temps de dérive des charges et accélérant le temps de réponse du détecteur. La résolution attendue sur l'énergie des pions est de  $\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$  [26].

### 2.3.4 Le spectromètre à muons

Ce sous-détecteur complète le système de détection utilisé dans l'expérience ATLAS en permettant une mesure précise de l'impulsion et de la charge des muons, seules particules (avec les neutrinos) qui vont traverser le système calorimétrique. La courbure de la trajectoire des muons est mesurée grâce à un système d'aimants toroidaux qui délivrent un champ magnétique de 2.6 T.m dans la région centrale et 6 T.m dans les bouchons. Le spectromètre à muons est composé de quatre sous-détecteurs tel que présenté sur la Figure 2.14. Les MDT (*Monitored Drift Tube*) et les CSC (*Cathode Strip Chambers*) sont utilisés comme détecteur de précision tandis que les RPC (*Resistive Plate Chamber*) et TGC (*Thin Gap Chamber*) participent aux systèmes de déclenchement d'ATLAS.

Les MDT couvrent la région  $|\eta| < 2.7$  et sont formés de trois couches sauf dans la partie bouchon où la première couche est remplacée par des CSC plus à même de résister au flux important de particules reçu dans cette région. Le principe utilisé est celui du tube à dérive avec un mélange de gaz formé à 93% d'argon et 7% de CO<sub>2</sub>. L'anode centrale est composée de tungstène et rhénium pour un diamètre de 50  $\mu\text{m}$ . Les CSC sont segmentés en  $\phi$  sur 2 roues avec 8 chambres multifils chacune. Le gaz utilisé est un mélange d'argon et de dioxyde de carbone.

Les RPC sont utilisées dans la partie tonneau tandis que les TGC vont être utilisées dans les bouchons. Dans les deux cas, un mélange gazeux (C<sub>2</sub>H<sub>2</sub>F<sub>4</sub>/Iso-C<sub>4</sub>H<sub>10</sub>/SF<sub>6</sub> pour les RPC et CO<sub>2</sub>/n-

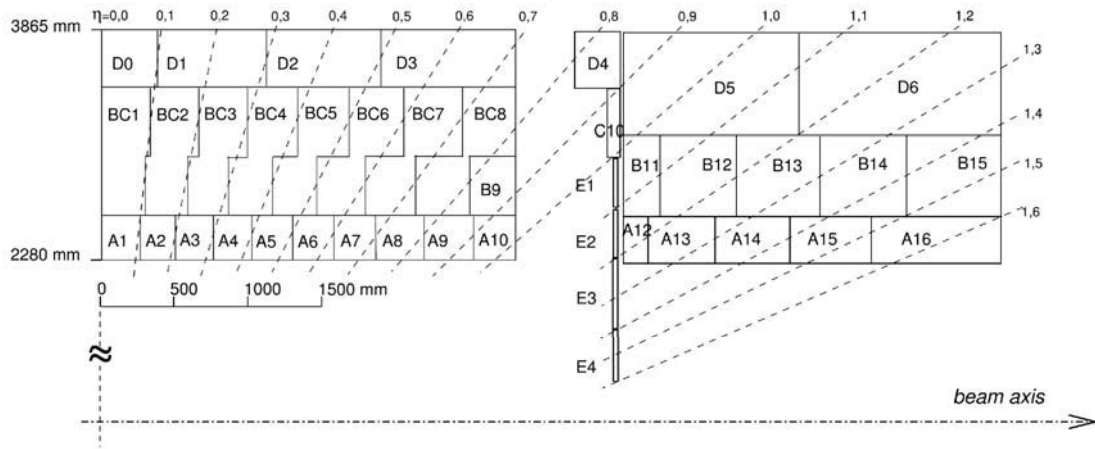


FIGURE 2.13 – Schéma de la distribution en couche des cellules du calorimètre hadronique à tuiles. [38]

pentane pour les TGC) est contenu entre deux plaques résistives et va se ioniser au passage d'un muon. L'intérêt de ces technologies est le temps de réponse très rapide, caractéristique utilisée par la suite pour le système de déclenchement.

### 2.3.5 Le système de déclenchement

La fréquence de croisement des faisceaux au LHC est de l'ordre de 40 MHz. La quantité de données extraites des collisions qui se déroulent est alors beaucoup trop importante pour que toutes soient stockées. Pour pouvoir faire face à ce problème un système de déclenchement (ou *trigger*) a été mis en place de façon à ne conserver que les événements potentiellement intéressants pour la physique<sup>4</sup>. Le principe du système de déclenchement d'ATLAS est présenté sur la Figure 2.15 et est constitué de deux niveaux.

#### 2.3.5.1 Le système de déclenchement de niveau 1

Le premier niveau du système de déclenchement fonctionne à la vitesse de croisement des faisceaux et permet une réduction du taux d'événements de 40MHz à 100kHz. Pour cela les informations calorimétriques et celles des chambres à muons sont utilisées, et plus particulièrement la multiplicité ou l'énergie des objets passant dans chacun de ces systèmes. Ainsi des régions d'intérêts en  $\eta$  et  $\phi$  sont définies selon l'énergie des objets présents dans ces régions.

#### 2.3.5.2 Le système de déclenchement de haut niveau

Pour les événements ayant passés le niveau de déclenchement 1, les différents objets vont être reconstruits (jets, leptons) en utilisant toute la granularité du détecteur. Des algorithmes de sélection complexes sont appliqués par le système de déclenchement de haut niveau pour réduire le taux d'événements à 1kHz.

4. Cela peut être des événements contenant des leptons ou des jets d'impulsion transverse suffisamment grande.

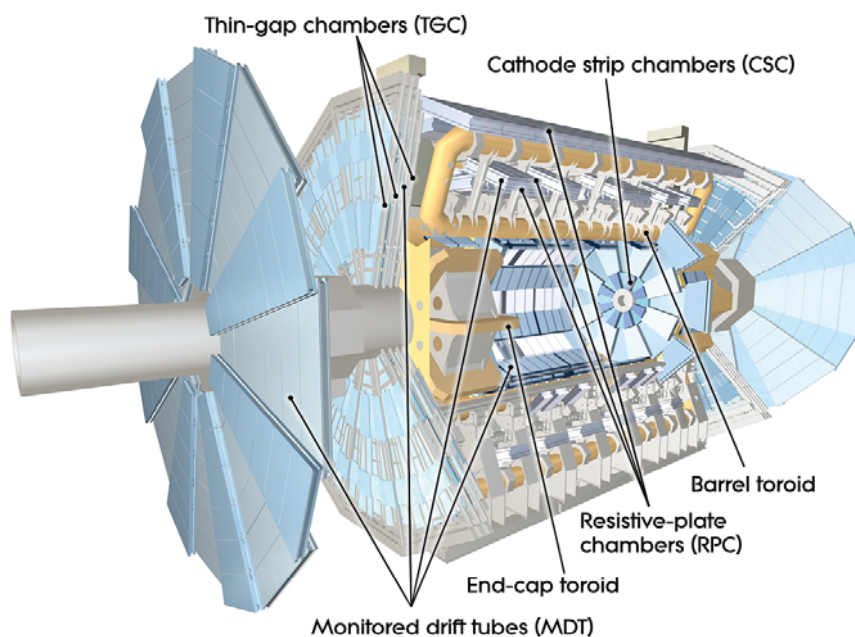


FIGURE 2.14 – Schéma du spectromètre à muon d'ATLAS.[26]

Les algorithmes utilisés lors du déclenchement sont regroupés selon une nomenclature précise qui se présente sous la forme :

[LEVEL] \_ [N TYPE] [THRESHOLD] \_ [QUALITY] \_ [ISOLATION]

avec :

- LEVEL : correspond au niveau du système de déclenchement utilisé (L1 pour niveau 1, HL pour haut niveau).
- N TYPE : donne le nombre de particules et le type de l'objet considéré pour le déclenchement.
- THRESHOLD : seuil requis en impulsion/énergie sur l'objet.
- QUALITY : qualité de l'identification des objets.
- ISOLATION : qualité de l'isolation des objets.

Par exemple, un filtre d'événements utilisé par la suite est HLT\_e120\_lhloose qui correspond à un électron qui passe le critère de qualité LHloose, ayant une impulsion transverse supérieure à 120GeV.

## Conclusion

Le LHC et le détecteur ATLAS forment un couple performant pour la production et la détection de particules de hautes énergies et haute masse. Tout d'abord le LHC offre une échelle en énergie nouvelle à 13 TeV tandis que le détecteur ATLAS qui couvre presque un angle  $4\pi$ sr constitue un excellent outil pour la détection des particules produites. Les différents sous-détecteurs dont il est composé permettent une reconstruction précise des vertex primaires et secondaires des collisions mais aussi de l'énergie et impulsion des particules venant de ces points de collisions. La quantité de données disponible atteint  $36 \text{ fb}^{-1}$  à 13 TeV à la fin de l'année 2016.

Outre ces excellents résultats de nombreuses améliorations sont encore prévues sur ATLAS jusqu'en 2035 pour répondre par exemple à l'augmentation de la luminosité prévue au LHC. Par ailleurs plusieurs projets d'accélérateurs linéaire (ILC : *International Linear Collider*, CLIC : *Compact*

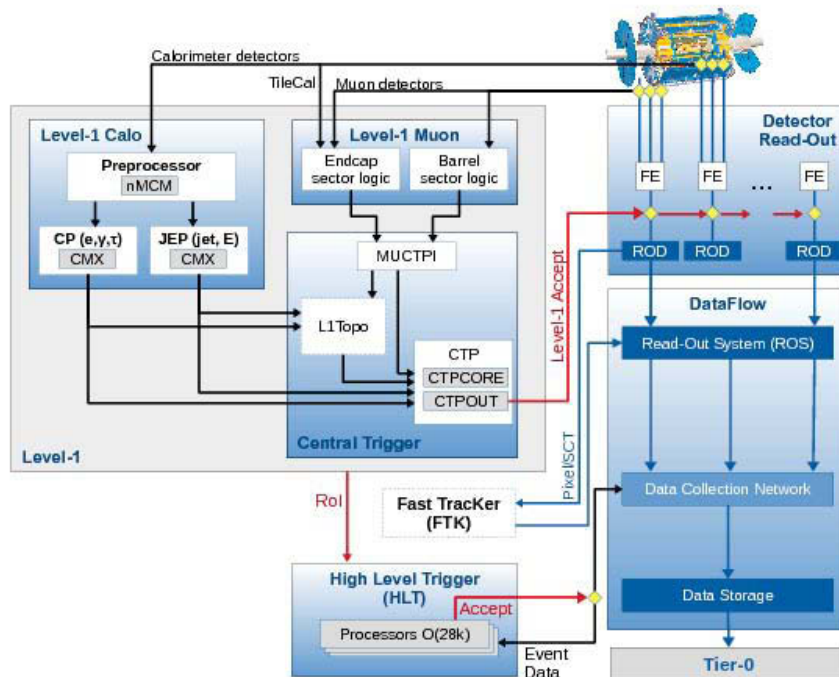


FIGURE 2.15 – Le système de déclenchement d'ATLAS.[39]

*Linear Collider*) ou circulaire (FCC : *Futur Circular Collider*) sont examinés pour repousser les limites atteintes par le LHC en précision (ILC, FCC, CLIC) ou énergie (FCC).



# Étalonnage du calorimètre hadronique à tuiles d'ATLAS avec un système laser

---

## 3.1 Introduction

Le calorimètre à tuiles d'ATLAS participe à la mesure des énergies des particules produites lors des collisions. Un étalonnage est nécessaire pour établir le lien entre la réponse du détecteur et l'énergie déposée par les particules. Dans le cadre du *TileCal*, plusieurs systèmes sont utilisés pour réaliser un étalonnage des différentes composantes que sont la partie optique (tuiles scintillantes et fibres optiques), les photomultiplicateurs et l'électronique de lecture. Dans un premier temps, une description de chacune de ces composantes sera réalisée. Une description rapide des systèmes d'étalonnage du détecteur sera ensuite faite avant une section plus détaillée sur l'étalonnage du *TileCal* avec un système Laser.

## 3.2 Fonctionnement du calorimètre à tuiles d'ATLAS

Une description de la structure du *TileCal* a été donnée dans la section 2.3.3.2 et une description plus détaillée des composantes est donnée dans cette section.

### 3.2.1 Composante optique

La partie optique du *TileCal* est composée des tuiles scintillantes et des fibres optiques. Une caractéristique des tuiles scintillantes est la présence de deux dopants pour éviter la réabsorption de la lumière par la tuile. Les tuiles sont toutes trapézoïdales mais de tailles différentes selon leur emplacement. Toutefois l'épaisseur est la même (3 mm) pour toutes les tuiles. Deux trous ont aussi été réalisés dans chaque tuile pour permettre le passage d'une source de Césium utilisée pour l'étalonnage du *TileCal*. Enfin les tuiles sont entourées dans un matériau réfléchissant pour limiter la perte de lumière.

Les fibres optiques, dites à décalage de longueur d'onde, transportent la lumière depuis les tuiles jusqu'aux photomultiplicateurs tout en modifiant la longueur d'onde de cette lumière pour atteindre 480 nm. Cette longueur d'onde finale est celle correspondant à la meilleure efficacité des photomultiplicateurs. Ces fibres sont aussi dopées, cette fois pour limiter l'effet Cherenkov d'une particule traversant la fibre.

### 3.2.2 Les blocs photomultiplicateurs (PMT)

Les PMT ont pour rôle de convertir la lumière reçue en signal électrique. Il est formé d'un bloc de plexiglas pour mélanger la lumière venant des différentes fibres d'un toron, du photomultiplicateur, d'un diviseur de haute tension et d'une carte "3-en-1". Ces PMT sont regroupés avec l'électronique de lecture dans des tiroirs au nombre de deux par modules. Le pont diviseur haute tension est utilisé pour répartir la haute tension entre les dynodes et pour connecter le photomultiplicateur à la carte 3-en-1.

Chaque photomultiplicateur est formé d'une photocathode et de dynodes. Les photons interagissent avec la photocathode par effet photoélectrique en arrachant un électron. Le nombre de charges sera ensuite multiplié à travers les dynodes. Finalement le signal est collecté par l'anode placée en sortie de la dernière dynode. Les photomultiplicateurs utilisés dans ATLAS ont un gain de l'ordre de  $10^5$  et l'effet de la luminosité sur le gain doit être inférieur à 2%. L'égalisation des gains des photomultiplicateurs est faite grâce à la haute tension qui leur est appliquée. Les tiroirs sont alimentés par une tension qui est comprise entre -830 V et -950 V (notée  $HV_{in}$ ). L'ajustement des tensions des photomultiplicateurs se fait alors entre  $HV_{in} - 360$  et  $HV_{in}$ . Un changement de la valeur de cette haute tension va affecter la réponse du photomultiplicateur. En effet la valeur de la tension et le gain des photomultiplicateurs  $G$  sont reliés par :

$$G = \alpha(HV)^\beta \quad (3.1)$$

avec  $\alpha$  et  $\beta$  des paramètres intrinsèques des PMT, respectivement valant environ  $10^{-14}$  et 7 en moyenne. À partir de l'équation 3.1 il est possible de calculer la variation relative de gain due à une variation de la haute tension entre le temps de référence  $t_r$  et un temps donné  $t$  :

$$\frac{\Delta G_{HV}}{G_{HV}} = \frac{G_{HV}(t) - G_{HV}(t_r)}{G_{HV}(t_r)} = \frac{HV^\beta(t)}{HV^\beta(t_r)} - 1. \quad (3.2)$$

Ainsi une variation de l'ordre du Volt de la haute tension entre les temps  $t$  et  $t_r$  conduit à une variation du gain du photomultiplicateur d'environ 1%.

Un autre effet à l'origine de la variation de gain des photomultiplicateurs est lié à la luminosité. En effet, plus les photomultiplicateurs reçoivent de la lumière, plus leurs gains diminuent. Cet effet est réversible et en l'absence de collisions le gain des photomultiplicateurs revient à sa valeur initiale.

### 3.2.3 L'électronique de lecture

La carte 3-en-1 a trois sorties analogiques : deux sorties pour l'amplification du signal selon un gain bas de 0.5 et un gain haut de 32 pour l'électronique de numérisation et une sortie pour le système de déclenchement de niveau 1. Elle contient aussi un intégrateur de signal dont le temps d'intégration est situé entre 10 et 20 ms. Cet intégrateur est utilisé pour mesurer le taux d'événements à biais minimum et donc la luminosité. L'intégrateur est aussi utilisé pour l'étalonnage du calorimètre par une source Cesium. Finalement la carte 3-en-1 permet l'étalonnage de l'ensemble de l'électronique par injection de charge.

### 3.2.4 La reconstruction du signal

Le signal en sortie de PMT est amplifié et mis en forme pour les gains haut et bas. Les deux signaux sont alors convertis simultanément en signal numérique par deux ADC (*Analog to Digital Converter*) et stockés dans les DMU (*Data Management Unit*) afin de procéder à un premier traitement. Ensuite si l'événement correspondant au signal passe le système de déclenchement de niveau 1, le signal est envoyé aux ROD (*ReadOut Drivers*) qui le mettent en forme et le stockent en attendant la décision du système de déclenchement de niveau 2. Pour cela des processeurs DSP (*Digital Signal Processors*) appliquent des algorithmes de reconstruction et fournissent une amplitude en coups ADC et un temps. Ces différents éléments sont schématisés sur la Figure 3.1.

Ainsi l'énergie reconstruite,  $E$  en GeV, à partir du nombre de coups ADC obtenus,  $A$ , est donnée par :

$$E(\text{GeV}) = A(\text{ADC}) \cdot C_{\text{ADC} \rightarrow pC} \cdot C_{pC \rightarrow \text{GeV}} \cdot C_{\text{Cesium}} \cdot C_{\text{Laser}} \quad (3.3)$$

avec  $C_{\text{ADC} \rightarrow pC}$ ,  $C_{\text{Cesium}}$  et  $C_{\text{Laser}}$  des constantes obtenues lors de l'étalonnage du TileCal à l'aide respectivement d'un système d'injection de charge [40], d'un système Césium [41] et d'un

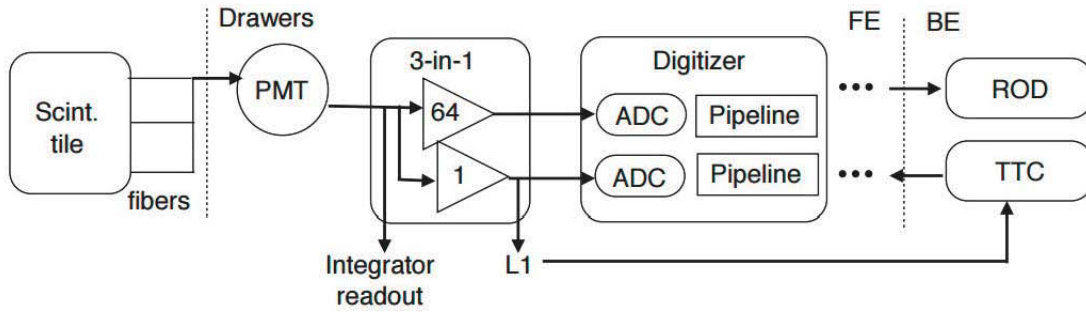


FIGURE 3.1 – Schéma décrivant les différentes composantes de l'électronique du *TileCal* utilisées pour la reconstruction du signal.

système laser [42].  $C_{ADC \rightarrow pC}$  correspond à la relation entre le nombre de coups ADC et la charge. La constante  $C_{Cesium}$  maintient l'échelle absolue tandis que  $C_{Laser}$  fournit une correction relative. Les systèmes d'étalonnage sont décrits plus en détail dans les sections suivantes. Le dernier facteur  $C_{pC \rightarrow GeV}$  a été extrait pour des électrons lors de faisceaux tests dédiés [43] et fait le lien entre la charge et l'énergie déposée. Le *TileCal* est ainsi étalonné à l'échelle électromagnétique.

### 3.3 Les systèmes d'étalonnage du TileCal

L'étalonnage du calorimètre à tuiles est utilisé pour connaître précisément la relation entre l'énergie déposée par les particules dans le détecteur et sa réponse. Or la réponse du calorimètre varie dans le temps il est donc nécessaire de faire un étalonnage du détecteur régulièrement.

Il existe trois systèmes d'étalonnage du calorimètre qui ont chacun un rôle précis : le système Césium est utilisé pour l'étalonnage de la partie optique du détecteur, des PMTs et de l'électronique de lecture par intégration, le système Laser pour l'étalonnage des PMTs et de l'électronique de lecture et le système par injection de charge qui contrôle l'électronique de lecture. Ainsi le couple formé par le système Césium et le système d'injection de charge permet l'étalonnage de l'ensemble du *TileCal*. En outre le système Laser maintient l'échelle d'énergie entre deux étalonnages par le système Césium. L'ensemble des systèmes de calibration et leurs zones d'application sont illustrés sur la Figure 3.2.

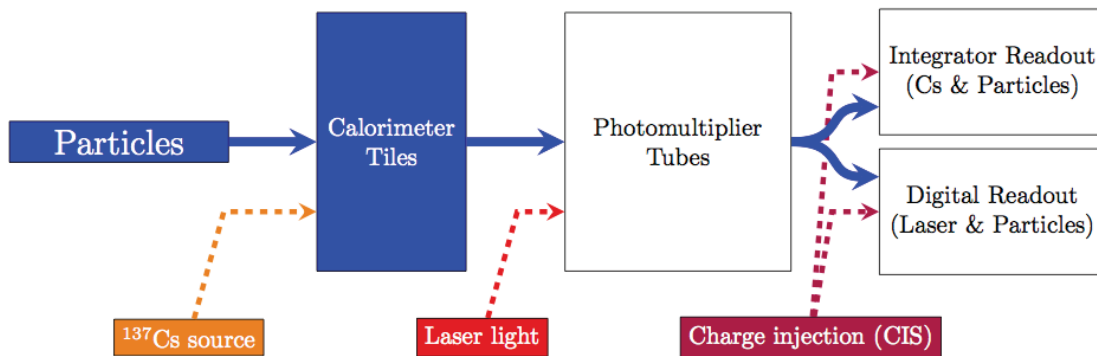


FIGURE 3.2 – Schéma décrivant quelles parties du détecteur sont étalonnés par chaque système d'étalonnage.

### 3.3.1 Le système Césium

Le système Césium est le seul système qui contrôle la partie optique du détecteur. Il permet aussi l'étalonnage des PMTs et de la partie intégrateur de l'électronique de lecture. Pour cela, trois sources radioactives de Césium (une dans les tonneaux longs et une chaque tonneau étendu) circulent grâce à un système hydraulique à travers le détecteur par des tuyaux passant à travers les tuiles scintillantes tel que présenté sur la Figure 3.3. Seules les cellules E ne sont pas calibrées par cette méthode. En outre, faire passer la source dans toutes les cellules pour l'étalonnage du calorimètre est long (plusieurs heures) et la mise en oeuvre de cette méthode se fait au mieux une fois par mois pour une précision d'environ 0.3%.

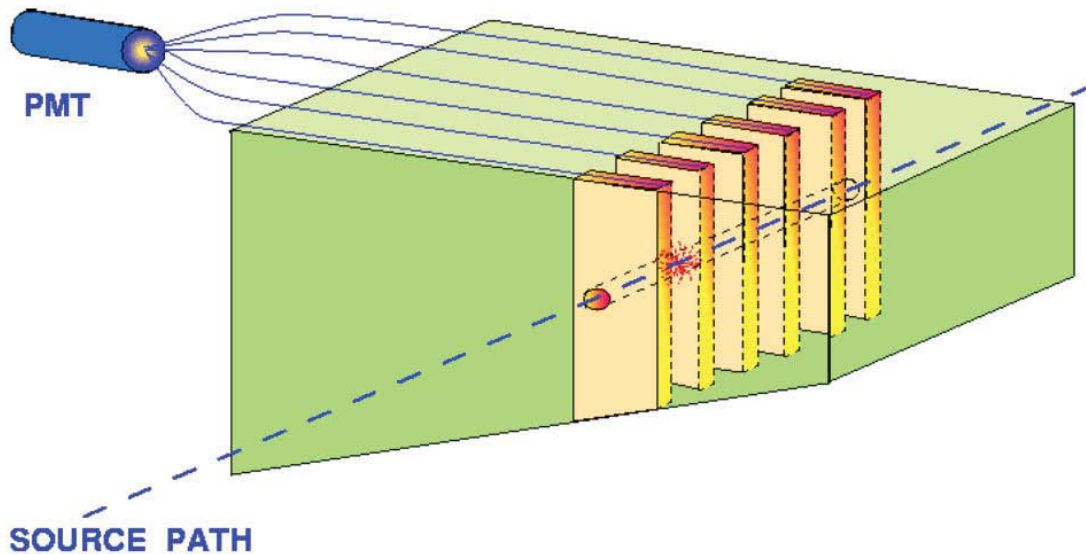


FIGURE 3.3 – Schéma montrant le passage de la source de Césium à travers les tuiles scintillantes du calorimètre.

### 3.3.2 Le système d'intégration de courants (*Minimum Bias*)

Le système *Minimum Bias* a pour but premier de mesurer la luminosité instantanée dans ATLAS à travers le comptage du nombre d'événements à biais minimum. Tout comme le système Césium, il mesure le signal provenant des scintillateurs et la variation de gain des PMTs. Il utilise aussi la partie intégrateur de l'électronique de lecture. Il n'a pas la précision pour une calibration canal par canal et n'est disponible que lors des collisions. Il est utilisé pour suivre l'évolution des cellules E et certaines cellules A (A13 et A14) quand il n'y a pas d'étalonnage avec le système Césium.

### 3.3.3 Le système d'étalonnage par injection de charge

Le système d'étalonnage par injection de charge consiste à envoyer une charge donnée pour mesurer la réponse de l'électronique du détecteur. Le facteur de conversion entre pC et coups ADC est extrait à partir de cette méthode à la fois pour le gain haut et le gain bas de l'électronique de lecture. La précision sur la mesure de ces constantes avec ce système est de l'ordre de 0.2% et en moyenne moins de 1% des canaux voient leurs constantes évoluer de façon significative en un mois. L'envoi de charges pour tester la mesure de l'électronique est rapide et cette procédure est réalisée trois à quatre fois par semaine environ.

### 3.3.4 Le système Laser

Le système Laser envoie de la lumière avec une longueur d'onde de 532 nm, proche de la longueur d'onde de la lumière provenant de l'interaction des particules avec les tuiles scintillantes lors des collisions (480 nm). Son but est l'étalonnage des PMTs et de l'électronique de lecture du calorimètre. Il recouvre donc ce qui est fait par le Césium et par le système d'injection de charge. L'utilisation de ce système est très rapide et deux à trois prises de données par semaine sont faites. Une description plus détaillée de l'étalonnage par le système Laser est donnée dans la suite de ce chapitre.

## 3.4 Description du système Laser

Le système Laser utilisé par ATLAS lors de la prise de données en 2015 et 2016 est composé de plusieurs composants qui sont décrits dans les sections suivantes et présentées sur la Figure 3.4.

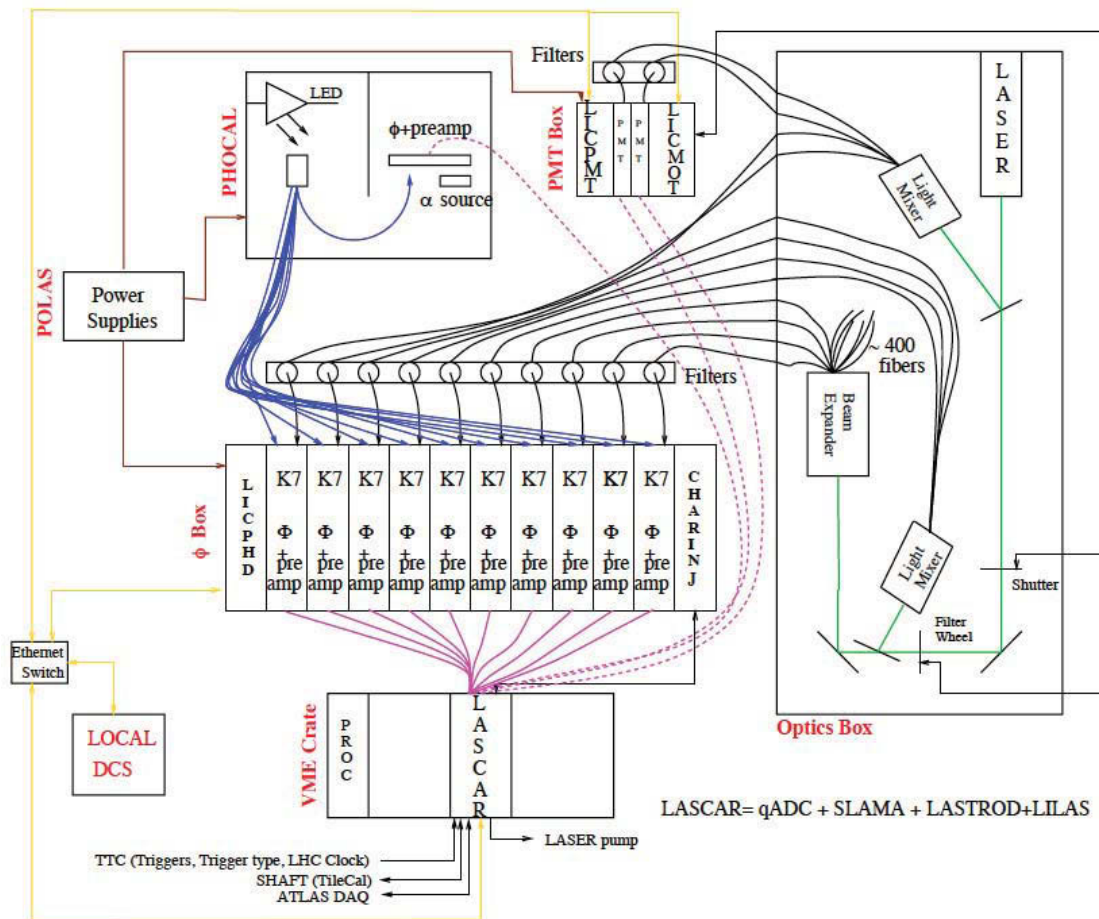


FIGURE 3.4 – Schéma du système Laser.

### 3.4.1 La partie optique

La lumière est produite par un Laser Q-switched DPSS (*Diode-Pumped Solid State*) de SPECTRA-PHYSICS [44].

À la sortie de la tête Laser, le faisceau de lumière est divisé en deux sous-faisceaux. Le premier est envoyé à un mélangeur de lumière qui renvoie la lumière vers cinq fibres optiques qui l'achemine

vers trois diodes qui contrôlent l'amplitude de la lumière à la sortie de la tête laser et vers deux PMT qui sont utilisés pour le système de déclenchement. Le reste de la lumière est envoyé vers une roue à filtres qui supporte huit emplacements. L'un de ces emplacements est vide tandis que les autres portent des filtres permettant une atténuation de la lumière entre 3 et 1000. Lors de l'étalonnage du TileCal, les filtres correspondant à des facteurs d'atténuation 3 et 330 sont utilisés correspondants aux voies bas gain et haut gain du *TileCal*.

Une seconde fois la lumière est divisée en deux faisceaux, avec une partie de la lumière envoyée vers trois nouvelles diodes pour en contrôler l'amplitude. Le reste de la lumière est envoyé vers le système de distribution de lumière qui doit répartir uniformément la lumière à travers 400 fibres pour illuminer les 9852 PMTs du TileCal. Avant ce système de distribution de lumière, un obturateur est placé pour éviter toute illumination des PMTs pendant les collisions ou lors de l'utilisation des autres systèmes d'étalonnage.

Finalement quatre autres diodes sont placées après le système de distribution de lumière pour contrôler une dernière fois la lumière. Ainsi plusieurs diodes sont présentes à chaque étape et il y a donc redondance entre celles-ci. Les 10 diodes présentes dans le système sont aussi étalonnées grâce à la lumière d'une LED. La lumière de la LED est aussi envoyée à une diode de référence appelée *Phocal* qui est elle-même étalonnée par une source  $\alpha$  de  $^{241}\text{Am}$ .

### 3.4.2 La partie électronique

Tout d'abord une boîte PMT contient les deux PMTs utilisés pour le système de déclenchement. Ensuite la partie la plus importante de l'électronique du système Laser est la carte LASCAR (*LASer CALibration Rod*) qui a pour but de contrôler le système Laser grâce à différents composants :

- la digitisation du signal des PMTs du système Laser,
- un système d'injection de charge (LILAS) pour surveiller la stabilité de l'électronique,
- un TDC (*Time-to-Digital converter*) pour estimer la réponse en temps du Laser en fonction de son intensité,
- une puce pour récupérer le signal du LHC,
- une carte HOLA pour envoyer des informations au système de contrôle de qualité des données d'ATLAS,
- une interface pour contrôler la puissance et le système de déclenchement du Laser.

## 3.5 Stabilité du système Laser

L'étalonnage du TileCal par le système Laser consiste à envoyer de la lumière dans les PMTs du calorimètre hadronique et de mesurer l'amplitude de la lumière envoyée ainsi que la réponse des photomultiplicateurs. Le rapport entre ces deux grandeurs sert à mesurer les variations de gain d'un PMT. Seule une diode est utilisée comme référence. Dans les résultats présentés dans la suite du chapitre, la diode 7 est utilisée car son placement après le système de distribution de lumière permet de contrôler tous les effets possibles, de la production de la lumière à la répartition de la lumière dans les fibres.

Il est nécessaire d'avoir des diodes correctement étalonnées et stables. Trois paramètres sont utilisés pour obtenir la stabilité des diodes :

- la valeur des piédestaux : valeurs obtenues dans le cas où il n'y pas de prise de données. Cette valeur est soustraite systématiquement.
- la réponse à la source  $\alpha$  pour l'étalonnage interne.
- la linéarité de l'électronique des diodes.

### 3.5.1 Piédestaux

Les piédestaux sont des séquences d'étalonnage pour lesquels aucune lumière n'est envoyée aux diodes, le but étant de mesurer le bruit électronique. La Figure 3.5 présente les variations relatives de la réponse des diodes pour les piédestaux pour quatre mois en 2015. La stabilité des piédestaux pour les différentes diodes est ainsi meilleure que 0.5% sur les quatre mois présentés.

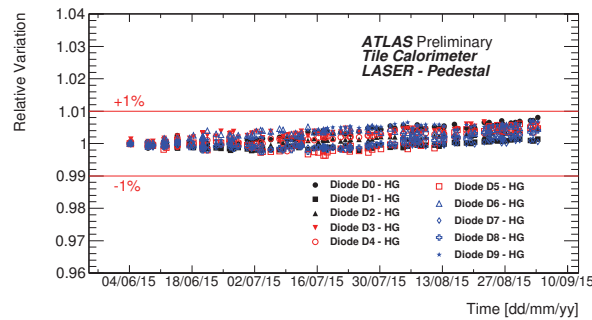


FIGURE 3.5 – Variations relatives de la réponse des piédestaux des diodes par rapport au premier piédestal entre Mai et Septembre 2015.

### 3.5.2 Étalonnage interne des diodes

Pour l'étalonnage des diodes, seule la diode Phocal est illuminée par la source  $\alpha$  et ensuite utilisée pour l'étalonnage des autres diodes. La variation de la réponse de la diode Phocal à la source  $\alpha$  sur quatre mois en 2015 est présenté sur la Figure 3.6.

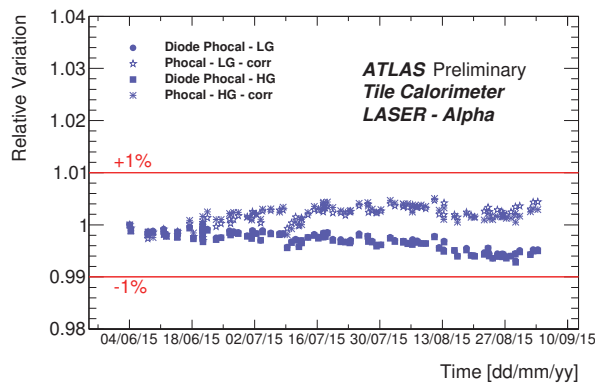


FIGURE 3.6 – Stabilité de la réponse de la diode Phocal entre Mai et Septembre 2015. Les valeurs sont normalisées à la valeur de la première mesure de la période.

En outre les diodes sont illuminées par une LED et les réponses sont comparées à la réponse de la diode Phocal à la lumière de cette même LED. La mesure de la lumière par les diodes est présentée sur la Figure 3.7. La stabilité de la mesure est meilleure que 0.1% pour toutes les diodes excepté la diode 6 qui a une stabilité de l'ordre de 0.7%. Cette stabilité est suffisante pour qu'aucune correction ne soit nécessaire.

### 3.5.3 Linéarité

La linéarité de l'électronique associée aux diodes est testée par l'injection d'une charge connue. La variation relative du signal dû à l'injection de charge pour les différentes diodes est présentée sur la Figure 3.8. La linéarité est meilleure que 0.3% pour l'ensemble des diodes.

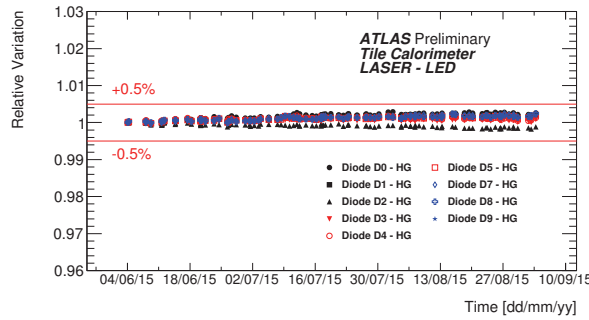


FIGURE 3.7 – Stabilité de la de la réponse des diodes au signal de la LED entre Mai et Septembre 2015 par les différentes diodes. Les valeurs sont normalisées à la valeur de la première mesure de la période et au signal de la diode **Phocal**.

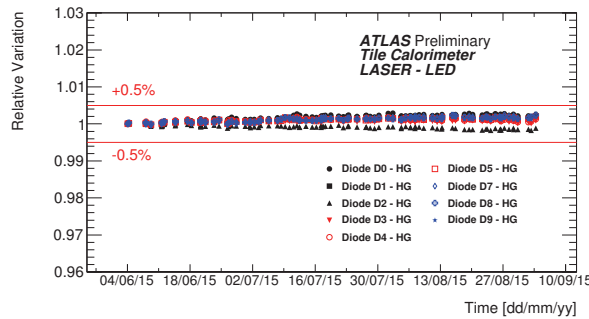


FIGURE 3.8 – Stabilité de la réponse des diodes à l'injection de charge entre Mai et Septembre 2015. Les valeurs sont normalisées à la valeur de la première mesure de la période.

## 3.6 Calcul des constantes d'étalonnages avec le système Laser

### 3.6.1 Données Laser

Un étalonnage Laser est composée de deux séquences d'étalonnage :

- une à bas gain (*LG* pour *Low Gain*) avec 10000 pulses, une amplitude constante et un facteur d'atténuation de 3.
- une à haut gain (*HG* pour *High Gain*) avec 100000 pulses, une amplitude constante et un facteur d'atténuation de 330.

Pour chaque pulse, les réponses des PMT et des diodes sont enregistrées. Ainsi, la réponse  $R_{i,p}$  d'un canal est normalisée et définie par :

$$R_{i,p} = \frac{E_{i,p}}{D_{ref,p}} \quad (3.4)$$

où  $E_{i,p}$  est le signal mesuré par le PMT et  $D_{ref,p}$  le signal mesuré par la diode de référence pour le pulse  $p$ . L'étalonnage est au final réalisé à partir de la valeur moyenne de ce rapport sur tous les pulses pour chaque canal :  $R_i = \langle R_{i,p} \rangle$ . Cette moyenne est proportionnelle au gain  $G_i$  d'un PMT.

### 3.6.2 Analyse des données Laser

L'étalonnage Laser est utilisé entre deux séquences d'étalonnage Césium afin de maintenir l'échelle d'énergie définie par ce dernier. Une séquence de référence est prise lors du premier éta-



lonnage Césium. La variation de gain est alors définie pour un canal  $i$  comme :

$$\Delta G_i = \frac{R_i - R_i^{ref}}{R_i^{ref}} \quad (3.5)$$

avec  $R_i^{ref}$  la réponse normalisée du canal  $i$  pendant la séquence d'étalonnage Laser de référence. Un autre élément à prendre en compte est le fait que la lumière transmise à chaque fibre n'est pas identique et peut varier en fonction du temps. Une fibre alimente deux moitiés de deux modules dans le tonneau long (par exemple les canaux pairs d'un module du tonneau A et les canaux impairs d'un module du tonneau B) et la moitié des canaux d'un tonneau étendu. Une correction,  $\Delta G_i^l$ , est donc appliquée au calcul de variation de gain. L'hypothèse utilisée pour extraire cette correction est que tous les gains des PMTs ne varient pas simultanément et que la correction apportée impose la moyenne des variations à 0. La variation de gain corrigée  $\Delta G_i^{corr}$  peut donc s'écrire :

$$\Delta G_i^{corr} = \Delta G_i - \Delta G_i^l. \quad (3.6)$$

Une méthode itérative est utilisée pour extraire ce terme correctif. La moyenne des gains mesurée pour les canaux liés à la même fibre permet l'estimation des variations de lumière d'une fibre à l'autre et d'y associer une correction. Comme certains PMT peuvent être instables, seuls les canaux d'une fibre ayant une variation de gain inférieure à deux fois la valeur de la RMS sont pris en compte.

Un autre effet à prendre en compte lors de cette correction est la variation de gain due à la luminosité. Pour limiter l'impact de cet effet seules les cellules les plus éloignées du faisceau sont considérées comme assez stables pour le calcul de la correction. Ce sont les cellules D pour les canaux longs et les cellules B13, B14, B15, D5, D6 pour les tonneaux étendus.

Finalement la constante d'étalonnage  $f_{las}^i$  pour un canal  $i$  utilisée pour la reconstruction du signal est donnée par :

$$f_{las}^i = \frac{1}{1 + \Delta G_i^{corr}}. \quad (3.7)$$

### 3.6.3 Précision statistique du système Laser

Pour extraire la précision du système Laser, des données de séquence Laser sont considérées pendant une période où aucune collision n'a eu lieu pour limiter la variation de gain des PMT due à la luminosité. La RMS de la distribution des variations de gain des PMT dans cette période est alors représentative de la précision statistique du système. La distribution de la variation de gain est présentée sur la Figure 3.9 et la précision est de l'ordre de 0.2%.

### 3.6.4 Résultat sur la variation de gain des canaux

L'origine principale des variations de gain des PMT est la luminosité ce qui se traduit par le fait que les canaux qui ont des variations de gain les plus importantes sont associés aux cellules les plus proches du point de collision. Ainsi, tel qu'illustré sur la Figure 3.10, les cellules A13, A14 et les cellules E sont les cellules qui ont les plus grandes variations de gain sur une période de 5 mois de prise de données.

La variation moyenne des gains par cellule peut aussi être représentée par en fonction de l'angle  $\phi$  et la distance au faisceau. Ces variations sont représentées sur la Figure 3.11. La présence de corrélation sur ces figures permettraient de repérer un problème dans ces cellules ce qui n'est pas le cas sur ces figures où la plus grande variation des cellules plus proches du faisceau sont bien visibles.

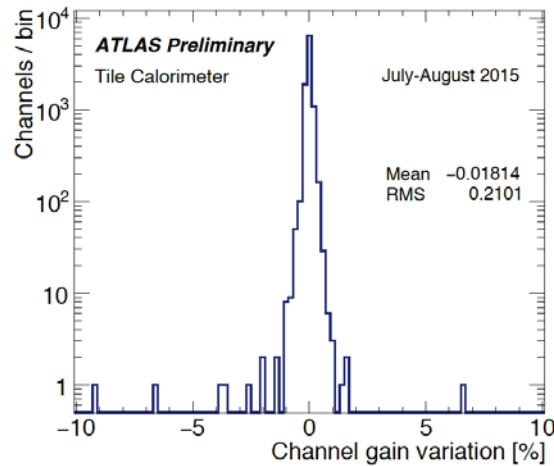


FIGURE 3.9 – Distribution de la variation de gain des 10000 PMT du TileCal pendant une période avec peu de prise de données. La RMS donne une idée de la précision du système Laser.

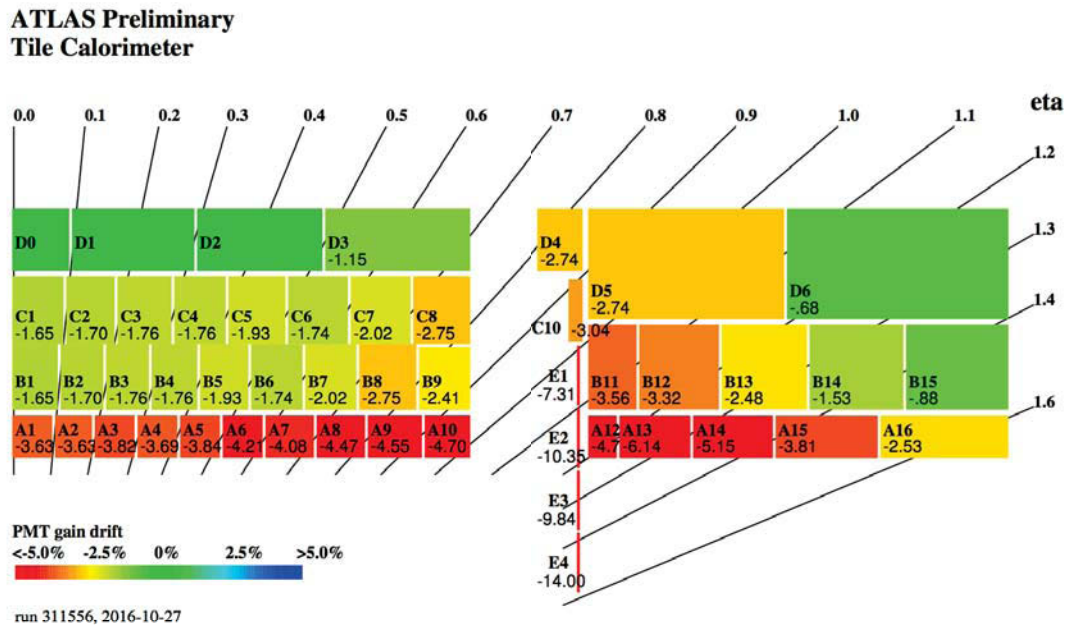


FIGURE 3.10 – Variation moyenne de gain par type de cellule en fonction de la pseudo-rapacité et de la distance au faisceau entre le 24 Mai 2016 et le 27 Octobre 2016 (en %).

### 3.6.5 Comparaison avec les variations de gain mesurées par le système Césium

Pour tester la précision du système Laser mais aussi la validité des variations de gains mesurées par le Laser, une comparaison est réalisée entre les variations de gains mesurées par le système Laser et le système Césium. Pour faire cette comparaison, les séquences d'étalonnages Laser les plus proches des séquences Césium (moins fréquentes) sont utilisées. Cette comparaison est présentée sur la Figure 3.12 pour la variation des gains des PMT entre Avril et Mai 2016. Tandis que la linéarité est illustrée sur la Figure 3.12a, la largeur du rapport des constantes Laser et Césium est de 0.1% comme illustrée sur la Figure 3.12b. Cette largeur combine la précision sur les systèmes Laser et Césium et les variations dues au système optique.

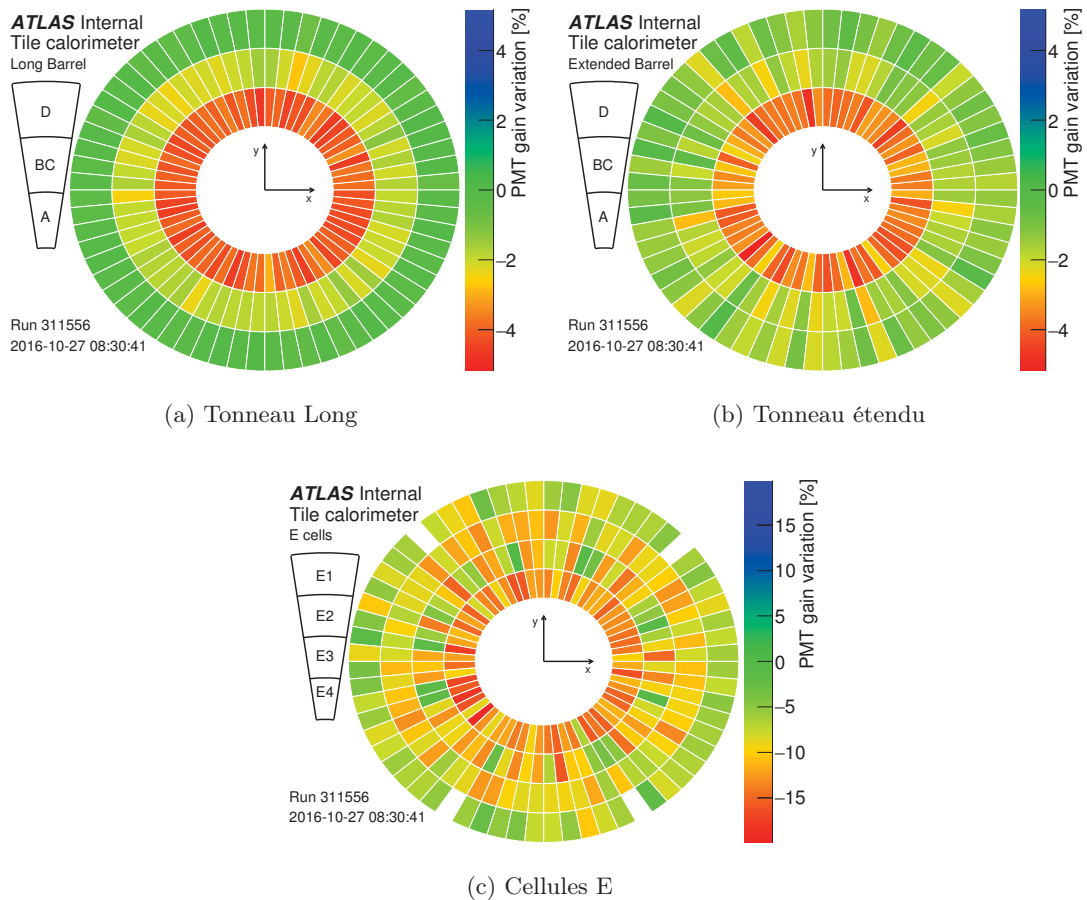


FIGURE 3.11 – Variation moyenne de gain par cellule en fonction de l'angle  $\phi$  et de la distance au faisceau entre le 24 Mai 2016 et le 27 Octobre 2016 pour les cellules des modules des tonneaux longs (a), pour les cellules des modules des tonneaux étendus (b) et pour les cellules E.

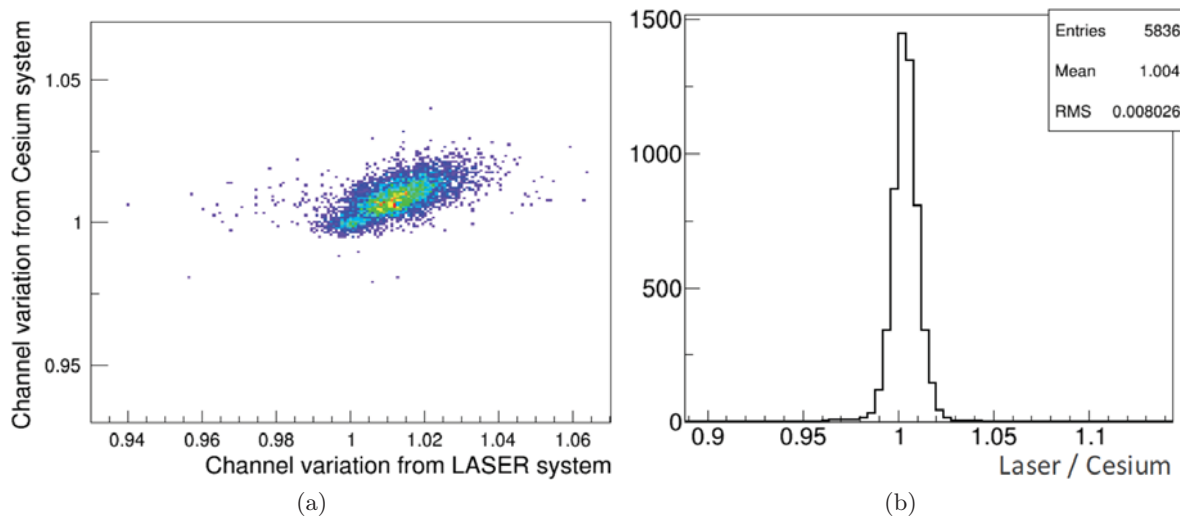


FIGURE 3.12 – (a) Variation de gain par canal vue par le système Césium en fonction de la variation vue par le Laser. (b) Rapport entre les variations de gain vues par les systèmes Césium et Laser. Dans les deux cas la période considérée est d'un mois entre Avril et Mai 2016.

À partir des variations de gains entre les données Laser et le Césium, il est aussi possible de tester l'effet du vieillissement des tuiles scintillantes. En effet le système Laser ne teste que les PMTs du *TileCal* tandis que le système Césium teste les PMTs et l'optique. Ainsi pour les cellules les plus exposées aux collisions que sont les cellules A13 et A14, une comparaison entre les variations observées par le Laser et le Césium peut être réalisée. Pour compléter cette analyse les résultats venant d'événements à biais minimum (*Minimum Bias*) sont ajoutés. Les variations des gains mesurées pour les cellules A13 pour la période correspondant à la prise de données en 2016 sont présentées pour les système Laser et *Minimum Bias* sur la Figure 3.13. La courbe en fond de l'augmentation de la luminosité intégrée permet de mettre en évidence la baisse du gain des PMTs avec l'augmentation de la luminosité puis le rétablissement du gain en absence de collisions. Enfin l'écart entre les gains donnés par les mesures Laser et *Minimum Bias* est interprété comme étant l'effet de l'irradiation des scintillateurs qui atteint une valeur de 3% à la fin de la prise de données en 2016.

### 3.7 Algorithme d'étalonnage des canaux du TileCal

Lors de l'étalonnage du TileCal par le système Laser, la variation de gain de chaque canal est calculée en utilisant des séquences d'étalonnage à bas gain, toutefois la constante d'étalonnage Laser n'est pas toujours calculée et sera mise dans ce cas à 1. Un algorithme a été mis en place pour construire de manière automatisée la liste des canaux à étalonner et pour extraire les canaux qui démontrent un comportement suspect. La Figure 3.14 montre un schéma des étapes du fonctionnement de l'algorithme. Trois catégories de canaux sont extraites à la suite de cet algorithme :

- Les canaux qui seront étalonnés dont la variation de gain est comprise ;
- Les canaux qui ne seront pas étalonnés et ne sont pas problématique ;
- Les canaux qui ne seront pas étalonnés et sont problématiques.

Les différents tests réalisés par cet algorithme sont appliqués aux canaux individuellement :

- si la variation de gain du canal est inférieure à 0.5%, le canal voit sa constante d'étalonnage gardée à 1.
- si le canal est masqué et non utilisé pour la prise de donnée lors de collisions alors la constante d'étalonnage est gardée à 1.

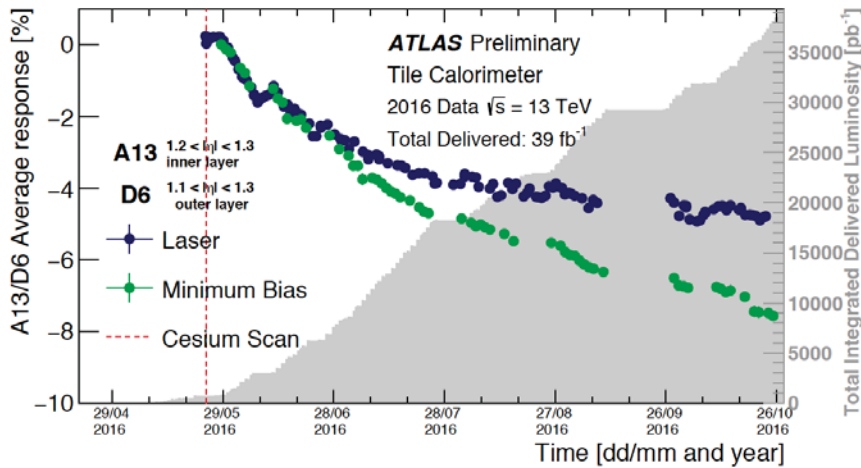


FIGURE 3.13 – Variation des gains mesurés pour les cellules A13 durant la prise de données en 2016 avec les systèmes Laser et *Minimum Bias*. En fond, la courbe de la luminosité instégrée est ajoutée.

- la variation de gain d'un canal est comparée à la variation moyenne des canaux appartenant au même type de cellules (A, BC, D, E). Si la variation du canal individuel est compatible avec la moyenne alors le canal peut être étalonné. Si ce n'est pas le cas, les étapes suivantes sont appliquées. Le critère jugeant de la compatibilité est dépendant du type de cellule avec le seuil de compatibilité étant de 1.5% pour les cellules B et de 6% pour les cellules E.
- la variation de gain obtenue à partir de la variation de haute-tension est calculée, si cette variation est compatible avec celle obtenue par le système Laser alors le canal est étalonné car la variation de gain du PMT est alors imputée à une variation dans la valeur de la haute-tension. Si ce n'est pas le cas les tests suivants sont réalisés. La Figure 3.15 présente le cas d'un canal dont la variations de gain correspond à une variation de haute-tension.
- si plus de la moitié des canaux reliés à une même fibre ont une variation de gain du même ordre de grandeur, la comparaison est faite avec les canaux de la deuxième fibre du même module. Si seuls les gains des canaux de l'une des fibres varient alors cette variation est interprétée comme un problème avec la fibre optique et ces canaux ne sont pas étalonnés.
- si le canal a eu une variation de gain non linéaire sur les 15 derniers jours alors le canal n'est pas calibré et est signalé comme étant problématique. La linéarité est testée en regardant si le signe de la variation de gain est constant (ou ne change qu'une fois), il n'y a pas de saut dans la variation de gain entre deux séquences Laser successives.
- Finalement si les canaux ont passé tous les autres tests et ont une variation de gain inférieure à 10% alors ils sont étalonnés. Dans le cas contraire ils sont signalés comme étant problématiques et ne sont pas étalonnés.

Au final le nombre de canaux problématiques est en moyenne inférieure à 10 et pour ces canaux des tests supplémentaires sont nécessaires pour expliquer les variations de gains. La variation de gain de deux canaux problématiques est présenté entre le 12 et le 27 Juillet 2017 sur la Figure 3.16. Dans ces cas, la variation du gain n'est pas linéaire au cours du temps.

L'étalonnage du *TileCal* grâce aux systèmes Laser peut être illustré par la Figure 3.17 qui présente la variation de l'échelle d'énergie des jets en fonction du temps. Le saut observé vers le *run* 305671 qui ramène la variable vers une valeur moyenne est lié à la mise à jour des constantes d'étalonnages Laser. L'effet vient principalement de la correction des gains des PMTs des cellules A qui sont les plus exposées. L'effet de la mise à jour des constantes Laser sur la reconstruction de l'énergie des objets physiques est aussi visible sur la Figure 3.18 qui superpose la variation de la réponse en énergie d'électrons provenant de bosons Z dans les cellules E4 à la variation de gain de

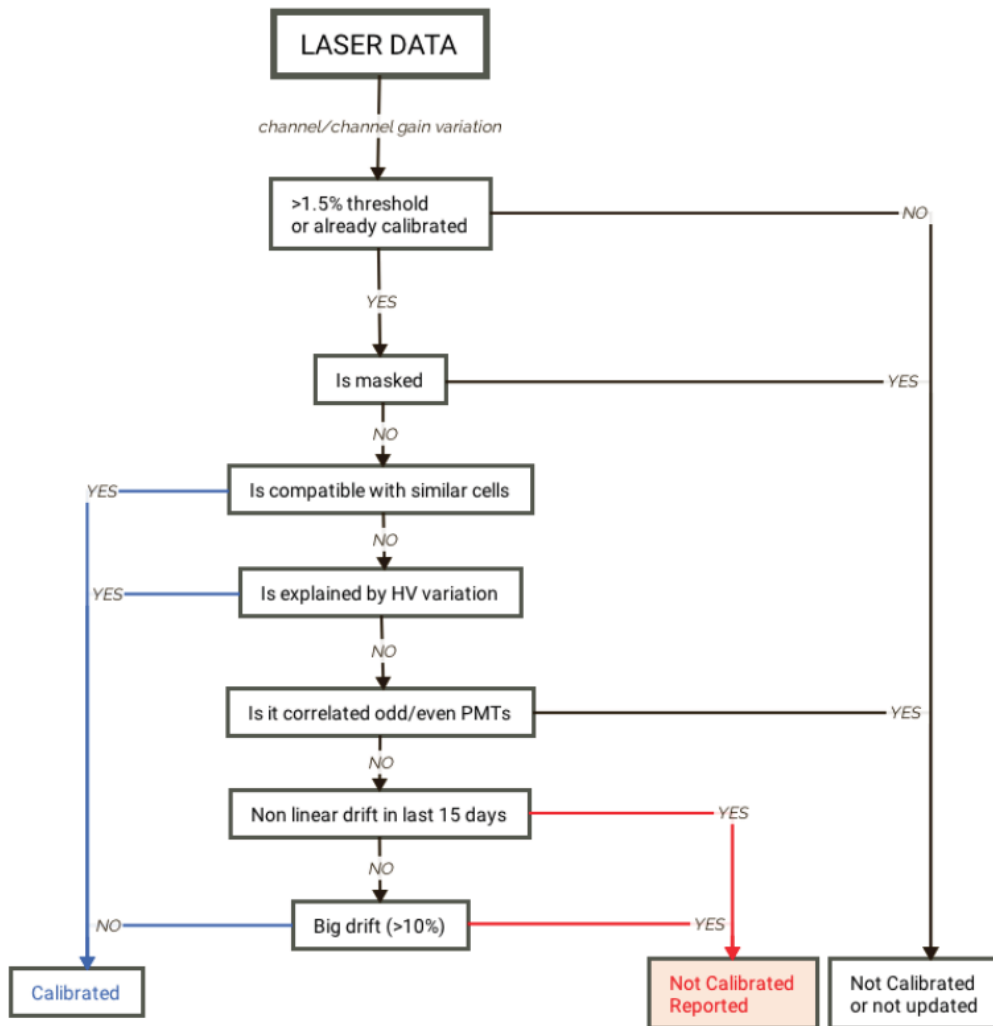


FIGURE 3.14 – Schéma du fonctionnement de l’algorithme utilisé pour extraire la liste de canaux du TileCal dont la constante d’étalonnage doit être mise à jour.

ces cellules obtenus par le système Laser. Lors de la mise à jour des constantes d’étalonnage Laser, la variation de la réponse en énergie des électrons revient à une valeur proche de 0 ce qui est désiré.

### 3.8 Conclusion

Une présentation rapide des systèmes de calibration du TileCal a été faite dans ce chapitre. Une présentation plus avancée de l’étalonnage du TileCal par le système Laser a été effectuée avec la description du calcul de la variation des gains des PMT ainsi que la variation des gains observée en 2015 et 2016. Un algorithme a été développé afin de calculer et produire de façon automatisée les constantes d’étalonnage Laser. Un ensemble de diagnostics a été implémenté afin d’identifier l’origine des variations de gains pour émettre une décision d’étalonnage canal par canal et identifier des canaux problématiques. Cet algorithme est maintenant utilisé dans ATLAS et a permis d’améliorer des données dès 2016.

Le Run 2 du LHC se caractérise par une forte augmentation de la luminosité délivrée par le LHC. Contrairement au Run 1, l’effet de la luminosité sur les PMTs nécessite un étalonnage plus régulier, au minimum hebdomadaire en période de prise de données. Le système d’étalonnage Laser et son automatiser répond à ce besoin. L’importance de son utilisation a été accentuée par l’arrêt du

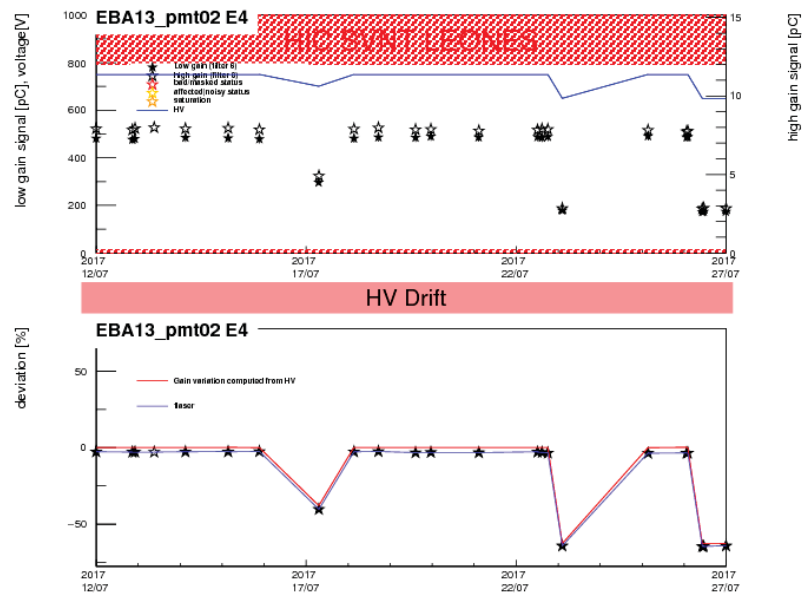
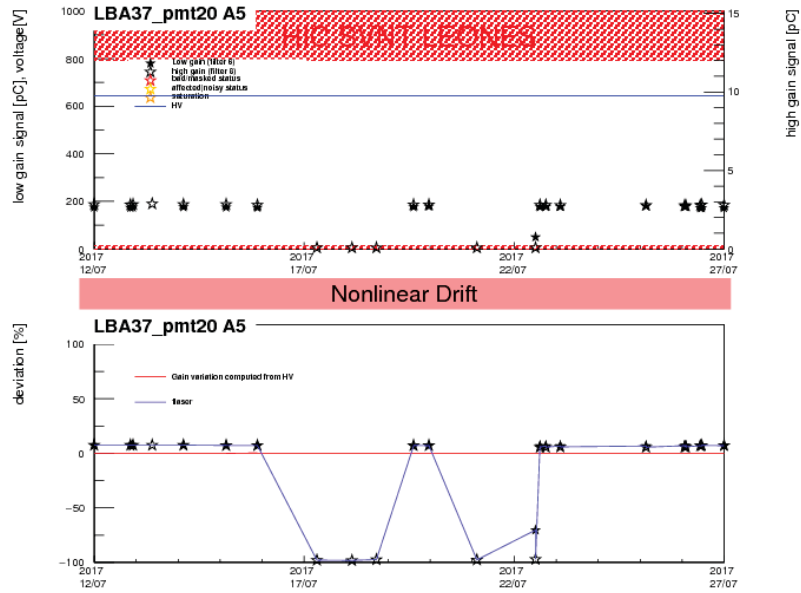


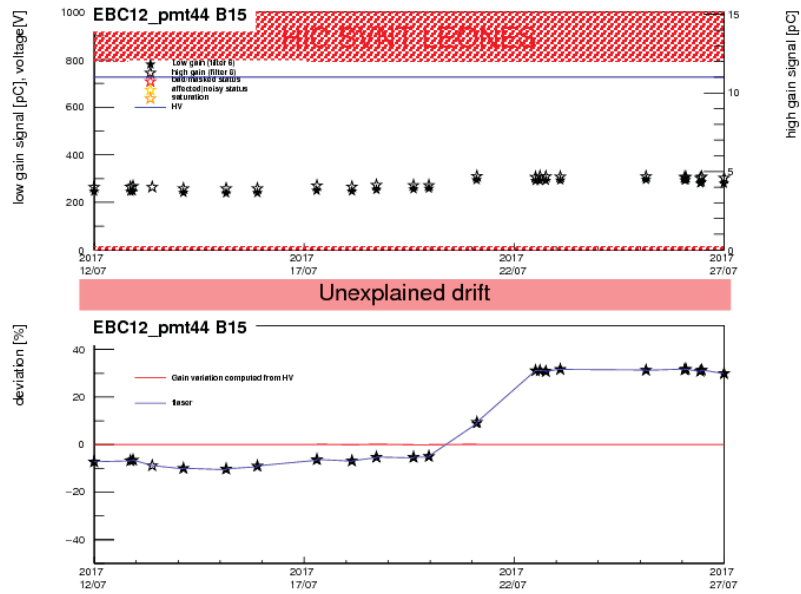
FIGURE 3.15 – Variation de gain d’un canal du tonneau étendu sur une période de 15 jours entre le 12 et le 27 Juillet 2017. La variation de gain telle qu’obtenue par le système Laser (étoile noire) et celle calculée à partir de la variation de la haute-tension (ligne rouge sur la partie basse) sont compatibles. Cette variation de gain est donc imputée à une variation de haute-tension.

système d’étalonnage Césium pendant plusieurs mois en 2016, suite à une fuite de son dispositif hydraulique. Le système Césium n’a été utilisé qu’en mai 2016 et le système Laser a permis de maintenir l’échelle d’énergie du *TileCal* jusqu’à la fin des prises de données en décembre 2016. Cet étalonnage a néanmoins dû être complété par des corrections moyennes issues de l’intégrateur (*Minimum Bias*) afin de prendre en compte l’impact des radiations sur les systèmes optiques les plus exposés (cellules E, certaines cellules A).

Au delà de l’étalonnage, l’utilisation combinée du système Laser et de l’ensemble des systèmes d’étalonnage fournit des diagnostics complets sur les éléments du calorimètre. Ces diagnostics guident les interventions en période de maintenance. Ce sont également des données, telle que l’identification des éléments robustes ou des éléments fragilisés, qui servent à préparer l’*upgrade* du calorimètre pour les prochaines phases du LHC.



(a)



(b)

FIGURE 3.16 – Variation de gain pour un canal du tonneau long (a) et d'un tonneau étendu (b) obtenu par le système Laser (étoile noire). Dans les deux cas cette variation n'est pas expliquée et les canaux sont considérés comme problématiques.



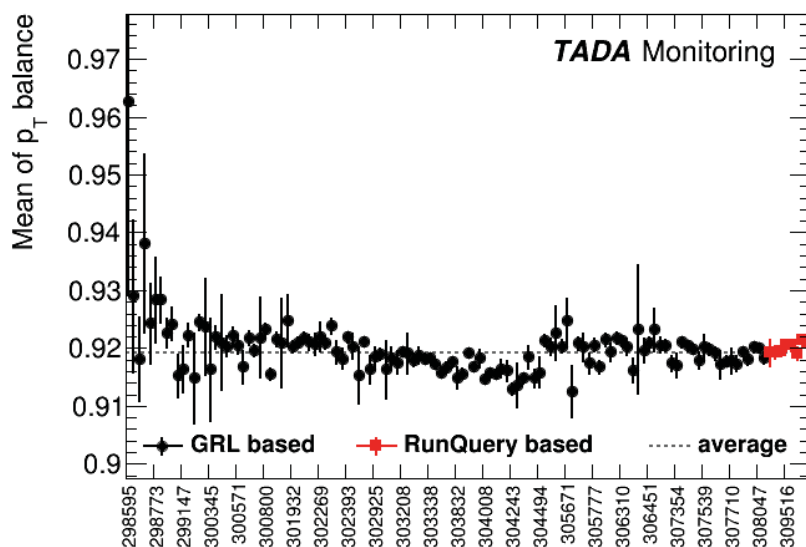


FIGURE 3.17 – Variation de l'échelle d'énergie des jets au cours du temps. Cette figure a été obtenue à partir d'événements  $\gamma$ +jets par le système TADA [45] qui a pour but de fournir un retour rapide sur les données récupérés lors des collisions par la représentation de nombreuses variables dans la simulation et les données. L'étalonnage par système Laser a été appliqué de façon régulière à partir du run 304494.

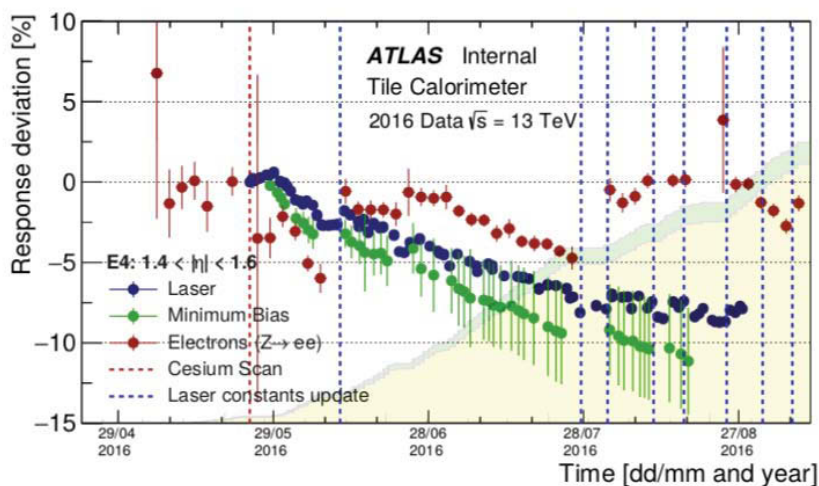


FIGURE 3.18 – Variation de la réponse en énergie des cellules E4 du calorimètre hadronique pour des électrons issus de désintégrations de boson Z. La variation de gain des mêmes cellules observées par les systèmes Laser et *Minimum Bias* est superposée ainsi que les dates correspondant à la mise à jour des constantes Laser pour ces cellules (lignes verticales).



# Simulation des événements dans ATLAS et définition des objets

## 4.1 Introduction

Le but des analyses présentées dans la suite du document est de tester le modèle standard par la comparaison entre ce qui est prédit par la théorie et ce qui est observé dans les données. Ainsi une simulation précise des processus physiques du modèle standard et des interactions des particules dans le détecteur ATLAS est nécessaire.

Dans un premier temps, le processus de simulation Monte-Carlo des collisions protons-protons est présenté dans la Section 4.2. Une deuxième étape, que ce soit pour les données simulées ou réelles, est la reconstruction des objets physiques utilisés dans les analyses présentées dans la suite du document. Cette reconstruction dans ATLAS est réalisée au sein de l'environnement ATHENA [46]. Les objets considérés et leur définition sont décrits dans la Section 4.3.

## 4.2 Simulation Monte-Carlo des événements dans ATLAS

Les simulations Monte-Carlo d'événements dans ATLAS doivent être les plus proches possibles de ce qui est observé dans les données. Ces simulations doivent donc prendre en compte les différents processus se déroulant dans le détecteur selon différentes étapes :

- **la génération de l'événement** : correspond à l'interaction entre les partons (quarks et gluons) provenant de deux protons lors de la collision ainsi que la désintégration et l'hadronisation des produits de cette collision tel que montré sur la Figure 4.1.
- **la simulation du détecteur** : correspond à l'ensemble des interactions des particules produites lors de l'interaction initiale dans le détecteur.
- **la digitisation** : correspond à la réponse du détecteur au passage des particules et la conversion sous le format des données réelles enregistrées par ATLAS.

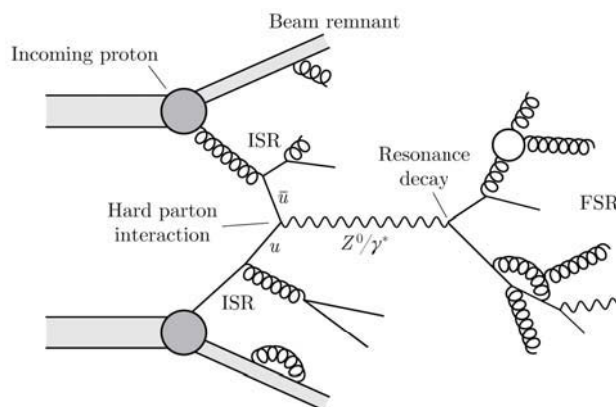


FIGURE 4.1 – Interaction entre deux partons provenant de deux protons et présentant des radiations dans l'état initial (ISR) ou final (FSR).

## 4.2.1 Génération des événements

### 4.2.1.1 Section efficace totale

Lors d'une collision proton-proton, la section efficace totale d'un processus  $pp \rightarrow X$  est donnée [47] par :

$$\begin{aligned}\sigma_{pp \rightarrow X} &= \sum_{a,b} \int_0^1 dx_a dx_b \int f_a(x_a, \mu_F) \cdot f_b(x_b, \mu_F) \cdot d\hat{\sigma}_{ab \rightarrow X}(\mu_F, \mu_R) \\ &= \sum_{a,b} \int_0^1 dx_a dx_b \int d\Phi_X \cdot f_a(x_a, \mu_F) \cdot f_b(x_b, \mu_F) \cdot \frac{1}{2x_a x_b s} |\mathcal{M}_{ab \rightarrow X}|^2(\Phi_X, \mu_F, \mu_R)\end{aligned}\quad (4.1)$$

pour laquelle :

- $x_a$  et  $x_b$  correspondent à la fraction d'impulsion des hadrons initiaux emportée par les partons  $a$  et  $b$ .
- $f_{a(b)}(x_{a(b)}, \mu_F)$  sont les fonctions de densité partonique (PDF) et décrivent la probabilité pour un parton  $a(b)$  d'emporter une fraction  $x_{a(b)}$  de l'impulsion du proton d'origine pour une échelle de factorisation  $\mu_F$  donnée.
- $\hat{\sigma}_{ab \rightarrow X}(\mu_F, \mu_R)$  est la section efficace de production au niveau partonique de l'état final  $X$  à partir des partons initiaux  $a$  et  $b$ . Cette section efficace dépend de l'espace des phases de l'état final  $\Phi_X$  et des échelles de factorisation  $\mu_F$  et de renormalisation  $\mu_R$ . Elle peut être exprimée en fonction de  $\mathcal{M}_{ab \rightarrow X}$ , l'élément de matrice moyenné sur les degrés de liberté de spin et de couleur de l'état initial, et du flux partonique  $2x_a x_b s$  avec  $s$  l'énergie dans le centre de masse au carré.

L'élément de matrice  $\mathcal{M}_{ab \rightarrow X}$  peut s'exprimer comme une somme sur tous les diagrammes de Feynman  $\mathcal{F}^{(i)}$  :

$$\mathcal{M}_{ab \rightarrow X} = \sum_i \mathcal{F}_{pp \rightarrow X}^{(i)} \quad (4.2)$$

### 4.2.1.2 Radiation dans les états initial et final

Lors des interactions entre protons, les partons peuvent émettre des gluons et des photons. Dans le cas d'une émission par l'un des partons initiaux on parle d'ISR (*Initial State Radiation*) alors que pour une émission venant d'une particule de l'état final on va parler de FSR (*Final State Radiation*). Ces émissions supplémentaires de gluons et photons vont conduire à des états finals plus complexes tel que  $pp \rightarrow Xg$  où  $g$  est un gluon supplémentaire émis lors de l'interaction.

### 4.2.1.3 Hadronisation

Plusieurs modèles sont utilisés pour décrire l'hadronisation des quarks, c'est-à-dire le passage d'un état final partonique à un état final hadronique. Les deux principaux modèles utilisés sont les modèles des cordes de couleur [48] et des amas [49]. Dans le premier modèle, l'interaction forte entre deux quarks est symbolisée par une corde. Un potentiel  $V$  est associé proportionnellement à l'éloignement des quarks et lorsque ce potentiel  $V$  a une valeur suffisante la corde casse et une nouvelle paire de quarks est formée, conduisant à deux paires de quarks. Dans le modèle des amas, les corrélations de couleurs existant entre les partons sont utilisées pour former des amas de couleurs qui vont se séparer en amas plus légers jusqu'à la création des hadrons.

### 4.2.1.4 Les générateurs Monte-Carlo

Au sein de la collaboration ATLAS, différents générateurs de simulation Monte-Carlo sont utilisés. Certains ne prennent en charge qu'une partie des processus présentés ci-dessus, tandis que

d'autres peuvent être utilisés pour décrire toutes les étapes de génération d'un événement. Il est en outre possible de combiner les deux types de générateurs.

**PYTHIA [50]** : générateur à l'ordre de l'arbre (LO) généraliste. Le modèle des cordes de couleur est utilisé pour l'hadronisation des quarks et les événements sous-jacents sont aussi simulés.

**HERWIG [51]** : générateur LO généraliste qui utilise le modèle des amas pour les processus d'hadronisation.

**SHERPA [52]** : générateur à l'ordre sous-dominant (NLO) généraliste qui utilise le modèle des amas pour les processus d'hadronisation.

**MADGRAPH [53]** : générateur LO permettant de calculer des éléments de matrice. L'hadronisation doit être réalisée à partir d'un autre générateur tel que PYTHIA.

**MC@NLO [54]** : générateur permettant de calculer des éléments de matrice à l'ordre sous-dominant (NLO). Les processus d'hadronisation sont réalisés par HERWIG.

**POWHEG [55]** : générateur permettant de calculer des éléments de matrice à l'ordre sous-dominant (NLO). L'hadronisation doit être réalisée à partir d'un autre générateur tel que PYTHIA ou HERWIG.

### 4.2.2 Simulation des interactions avec le détecteur ATLAS

Après la génération des événements, la simulation de l'interaction des particules avec le détecteur et de la digitisation doivent être réalisées. Deux options sont actuellement disponibles dans ATLAS pour réaliser cette simulation [56].

La première option consiste à réaliser une simulation complète du détecteur ATLAS. Cette simulation s'appuie sur GEANT4 [57] pour obtenir une simulation précise du passage des particules dans le détecteur et de leurs interactions au prix d'un temps de calcul important.

Une simulation rapide de l'interaction des particules avec le détecteur appelée ATLFASSTII [58] est aussi disponible basée sur la simplification de la géométrie du détecteur et de la réponse de certains sous-détecteurs. Par exemple, une bibliothèque de gerbes électromagnétiques et hadroniques est utilisée plutôt que la simulation par événements des interactions des particules avec les calorimètres. En effet cette étape est l'une des plus lourdes en terme de temps de calcul lors de la simulation complète.

Après avoir simulé le détecteur, les informations des coups simulés des particules dans le détecteur sont convertis en réponse digitale. C'est à cette étape que l'effet d'empilement d'événement est simulé en superposant plusieurs événements.

Suite à cette simulation du détecteur, les informations disponibles sont les traces et les dépôts d'énergie simulés dans le détecteur. Pour obtenir les objets utilisés dans les analyses (électrons, muons, jets ...), la reconstruction doit être faite à partir de ces informations. La définition de ces objets est détaillée dans la section suivante.

## 4.3 Définition des objets physiques

Les informations transmises par le détecteur lors des collisions ou générées lors des simulations ne permettent pas directement de définir quels sont les objets présents dans un événement.

Ainsi, des algorithmes de reconstruction, utilisant les données fournies par les différents sous-détecteurs d'ATLAS, sont utilisés pour reconstruire des objets physiques (électrons, muons, jets, énergie transverse manquante). Ce sont les objets qui se rapprochent le plus de particules stables dans le détecteur. La définition de ces objets et des algorithmes associés est présentée dans cette section.

### 4.3.1 Les électrons

#### 4.3.1.1 Reconstruction et identification

La reconstruction des électrons dans la région centrale ( $|\eta| < 2.5$ ) est obtenue en plusieurs étapes [59]. Dans un premier temps un algorithme de fenêtre glissante est utilisé pour repérer des zones de  $3 \times 5$  cellules élémentaires de la couche moyenne du calorimètre électromagnétique, chacune de ces cellules correspondant à une surface dans le plan  $(\eta, \phi)$  de  $0.025 \times 0.025$ , avec une énergie totale déposée supérieure à 2.5 GeV. Ensuite la reconstruction de la trace des électrons est effectuée dans le trajectographe à partir d'une ébauche de trace correspondant à trois dépôts dans différentes couches du détecteur à pixels ou du SCT (décrit dans la Section 2). Finalement une association entre les traces reconstruites et les zones reconstruites dans le calorimètre est effectuée en utilisant la distance en  $\eta$  et  $\phi$  entre les deux éléments. Si plusieurs traces sont candidates, la plus proche dans le plan  $(\eta, \phi)$  est sélectionnée. L'efficacité de reconstruction, définie comme la fraction d'électrons effectivement reconstruits par l'algorithme précédent est représentée sur la Figure 4.2 en fonction de l'énergie déposée dans le calorimètre électromagnétique.

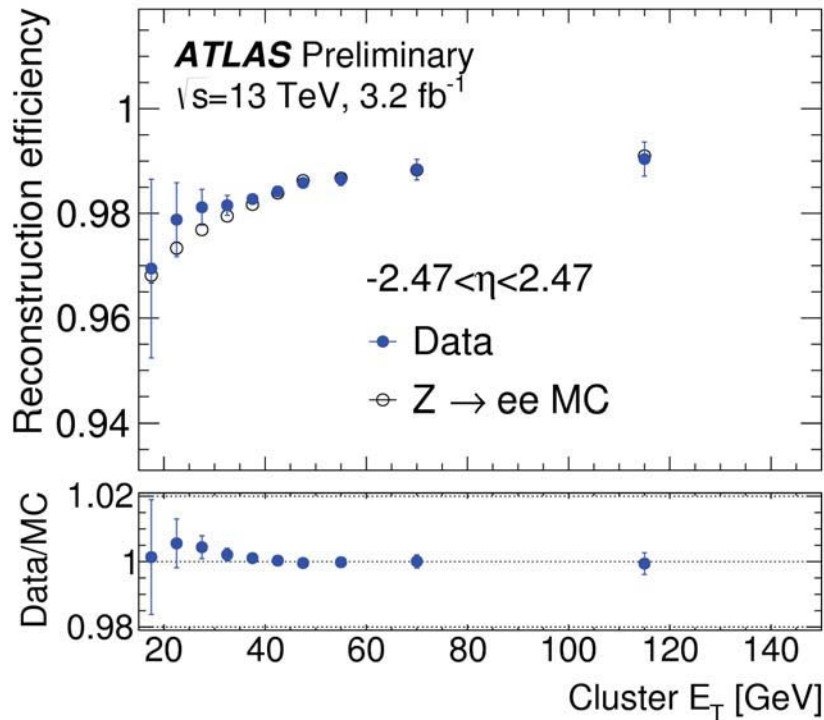


FIGURE 4.2 – Efficacité de reconstruction en fonction de l'énergie déposée dans le calorimètre électromagnétique pour les données 2015. Elle est obtenue à partir d'événements  $Z \rightarrow e^\pm e^\mp$ .

Il existe plusieurs critères pour identifier un objet comme un électron avec une probabilité plus ou moins grande. Des algorithmes d'identification des électrons sont utilisés pour distinguer les

candidats étant de vrais électrons des candidats étant par exemple des jets reconstruits comme des électrons. Ces algorithmes utilisent diverses informations telles que la forme des dépôts d'énergie dans le calorimètre, les propriétés des traces reconstruites ou encore la fraction d'impacts avec une haute énergie dans le TRT. L'ensemble des variables discriminantes considérées sont combinées pour construire une fonction de vraisemblance sur laquelle une sélection va être effectuée. Selon le critère utilisé, trois points de fonctionnement sont définis, *LooseLH*, *MediumLH* et *TightLH* qui sont inclusifs. Ainsi un électron avec une identification *TightLH* est aussi un électron *LooseLH*. L'efficacité de ces trois points de fonctionnement pour un électron avec une énergie transverse de 25 GeV est de 91% (*LooseLH*), de 87% (*MediumLH*) et 78% (*LHTight*). Ces efficacités augmentent avec l'énergie transverse de l'électron.

Outre l'identification des électrons, un autre critère portant sur l'isolation des électrons peut être appliqué afin de réduire la proportion d'électrons provenant des désintégrations de hadrons de saveurs lourdes. Plusieurs types de critères seront appliqués par la suite :

- **Loose isolation** : Elle est constituée de deux quantités fondées sur la somme des énergies transverses des dépôts calorimétriques dans un cône de  $\Delta R = 0.2$  autour de l'électron et sur la somme des impulsions transverses de toutes les traces dans un cône de  $\Delta R = \min(0.2, 10 \text{ GeV}/E_T)$ . Un seuil maximal est imposé à ces deux quantités de façon à obtenir une efficacité de sélection de 99 % dans les processus  $Z \rightarrow ee$ .
- **Gradient isolation** : ce point de fonctionnement est construit de façon à avoir une efficacité de 99% pour des électrons de 60GeV et une efficacité plus basse à plus basse énergie.
- **FixedCutTight isolation** : Dans ce cas, les sommes des énergies transverses des dépôts calorimétriques et des impulsions transverses des traces vont être utilisées mais la sélection sur les quantités précédentes va être fixée *a priori*. Cette isolation doit fournir un haut taux de réjection des leptons non isolés.

#### 4.3.1.2 Réjection des électrons avec une mauvaise reconstruction de la charge électrique

Dans l'analyse effectuée sur les données combinées de 2015 et 2016, un discriminant est considéré dans le but de rejeter les événements pour lesquels un électron a sa charge mal reconstruite. Il est construit à partir du résultat d'un arbre de décision boostée (BDT). Huit variables allant du  $p_T$  des électrons à la différence dans le plan transverse  $\Delta\Phi$  entre la position du dépôt d'énergie dans le calorimètre électromagnétique et la trace de l'électron dans le trajectographe, sont considérées. Une sélection est faite sur la variable de sortie du BDT  $> 0.0670415$  pour une efficacité sur le signal de 95% et une réduction du nombre des électrons avec une mauvaise charge reconstruite par un facteur 14. L'impact de ce discriminant sur des événements  $Z \rightarrow e^\pm e^\mp$  est présenté sur la Figure 4.3. Pour des événements avec deux leptons reconstruits de signes opposés, correspondant à des charges correctement reconstruites, l'impact de l'ajout de la sélection sur le BDT est faible. Inversement, l'impact de cette sélection va être important sur des événements avec deux leptons de même signe dont l'un a vu sa charge électrique mal reconstruite. Dans le cas du canal avec deux leptons de même charge pour le processus  $t\bar{t}H$ , l'impact de cette sélection est faible car les deux leptons sont prédits comme étant de même charge électrique. Il n'y a pas de mauvaise reconstruction de la charge dans ce cas.

#### 4.3.1.3 Facteurs d'échelles

Comme indiqué sur la Figure 4.4, l'efficacité combinée de reconstruction et d'identification obtenue dans la simulation est différente de celle obtenue dans les données. Pour corriger cette différence, un facteur d'échelle proche de 1 est utilisé pour pondérer l'événement. Un facteur similaire est extrait pour l'isolation des électrons mais aussi pour corriger la résolution en énergie des leptons reconstruits. Dans tous les cas, les désintégrations du boson Z en paire  $e^\pm e^\mp$  sont utilisées.

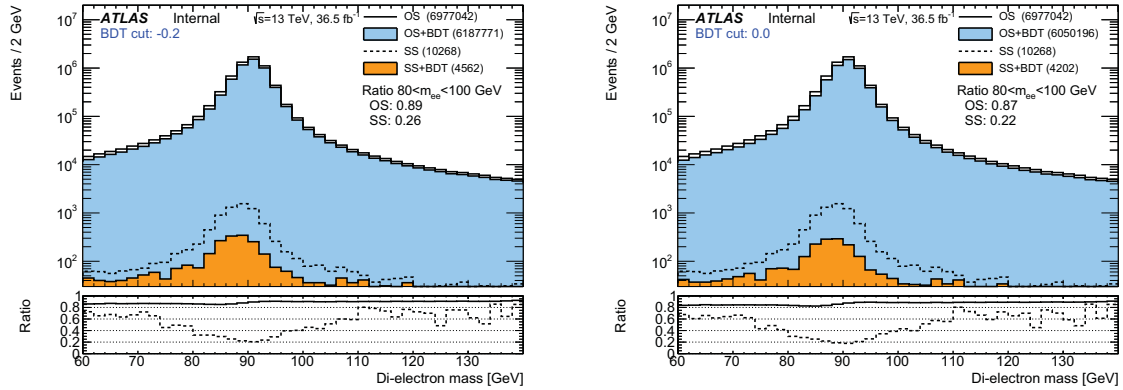


FIGURE 4.3 – Masse invariante diélectron pour événements avec deux leptons de même signe (SS) et deux événements avec deux leptons de charges opposées pour une sélection sur le discriminant  $> 0.2$  (a) et  $> 0$  (b). Les distributions sont données sans sélection sur le BDT (lignes pleine et pointillée) et avec sélection sur le BDT (histogrammes pleins). L’encadré du bas présente le rapport entre le nombre d’événements avec et sans sélection sur le BDT.

### 4.3.2 Les muons

#### 4.3.2.1 Reconstruction et identification

La reconstruction des muons [60] se fait à partir des traces laissées dans le détecteur interne et les chambres à muons. Dans la région centrale,  $|\eta| < 0.1$ , où la couverture par le spectromètre à muon est réduite, les informations des calorimètres sont aussi utilisées pour reconstruire les muons. L’efficacité de reconstruction des muons est présentée sur la Figure 4.5 en fonction de l’impulsion transverse des muons. Cette efficacité est mesurée pour des événements  $Z \rightarrow \mu\mu$ .

#### 4.3.2.2 Isolation

Les critères d’isolation des muons sont similaires à ceux considérés pour les électrons excepté que les informations des dépôts calorimétriques ne sont pas utilisés. Ainsi le point de fonctionnement `FixedCutTightTrackOnly` est utilisé. Celui-ci a la même définition que le point `FixedCutTight` à l’exception de la partie calorimétrique qui n’est pas utilisée.

#### 4.3.2.3 Facteurs d’échelle et correction en impulsion

Similairement aux cas des électrons, les différences pour les efficacités de reconstruction, d’identification et d’isolation obtenues dans la simulation et les données sont corrigées par des facteurs d’échelle. Ceux-ci sont très proches de 1 dans le cas des muons. Les événements  $Z \rightarrow \mu\mu$  sont utilisés pour extraire ces facteurs d’échelle.

La Figure 4.6 présente la masse invariante dimuon pour des événements  $Z \rightarrow \mu\mu$  pour les données et la simulation. Le décalage entre les deux spectres et la différence de largeur du pic sont corrigés par une correction de l’impulsion et de la résolution des muons dans la simulation. Le résultat de cette correction est aussi présenté sur la Figure 4.6.

### 4.3.3 Les jets

Les jets peuvent être reconstruits [61] de différentes manières.

Tout d’abord les jets **calorimétriques** sont reconstruits à partir de dépôts dans les calorimètres



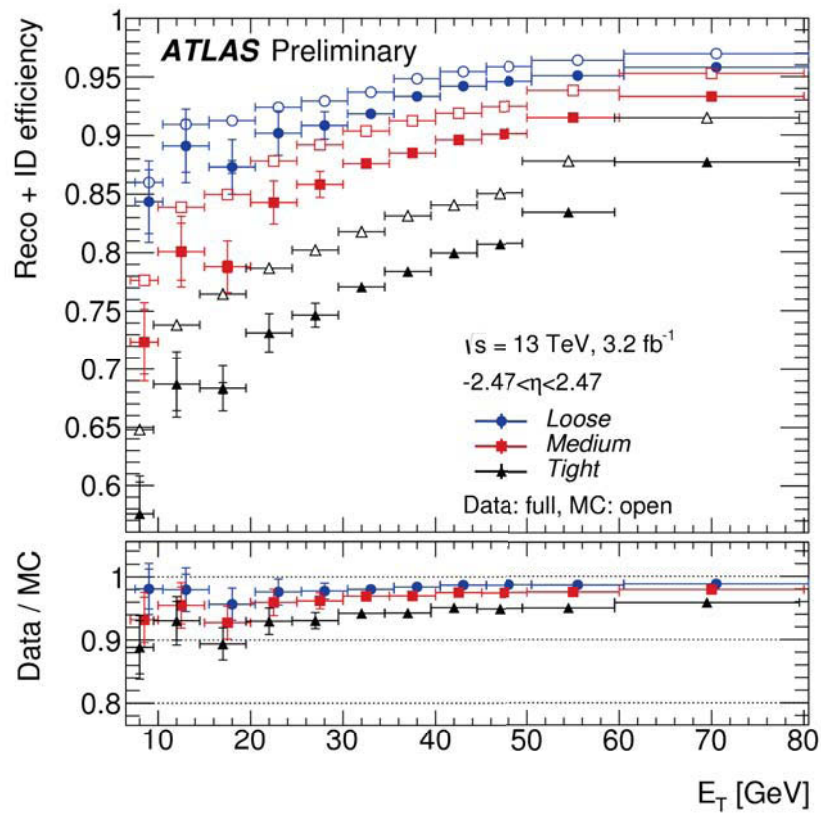


FIGURE 4.4 – Efficacité combinée de reconstruction et d’identification pour des événements  $Z \rightarrow e^\pm e^\mp$  en fonction de l’énergie transverse dans les données (symboles pleins) et dans la simulation MC (symboles vides).

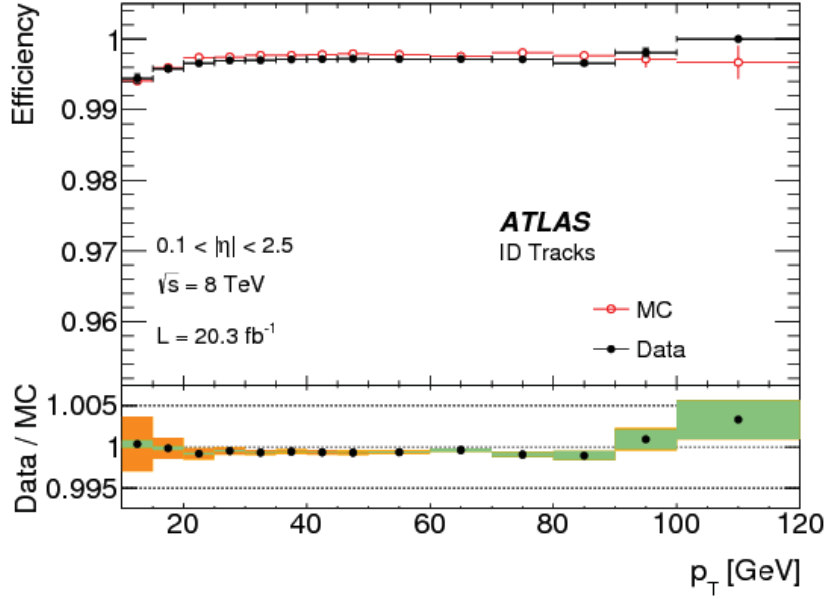


FIGURE 4.5 – Efficacité de reconstruction en fonction de l’impulsion transverse des muons pour les données 2012. L’encadré du bas montre le rapport entre données et prédictions de la simulation.

regroupés dans des amas topologiques de cellules calorimétriques (*topo-clusters*). Ces amas sont construits en sommant l’énergie dans ces cellules si celle-ci est supérieure à un seuil définie par le bruit de fond électronique et l’effet d’empilement. La direction de l’amas ainsi reconstruit est prise comme étant celle correspondant à la moyenne pondérée en  $\eta$  et  $\phi$  des cellules considérées. La pondération se fait en fonction de l’énergie présente dans chacune des cellules utilisées lors de la reconstruction de l’amas. La reconstruction des jets se fait par combinaison des amas au moyen d’algorithme de reconstruction basé sur les distances :

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \times \left( \frac{(\Delta R_{ij})^2}{R^2} \right) \quad (4.3)$$

$$d_i = p_{T,i}^{2p} \quad (4.4)$$

où  $p_{T,i(j)}$  représente l’impulsion transverse de l’amas  $i$  ( $j$ ),  $R$  la distance entre deux amas  $i$  et  $j$  dans le plan  $(y, \phi)$  avec  $y$  la rapidité<sup>1</sup>.  $p$  est un facteur qui peut être égal à 1, 0 ou -1. L’algorithme anti- $k_t$  [62] est utilisé pour cette reconstruction dans la suite du document avec un rayon  $R=0.4$  et  $p = -1$ .

Si  $d_i < d_{ij}$  alors l’amas  $i$  forme un jet et est retiré de la liste des amas utilisés par l’algorithme. Dans le cas contraire, les amas  $i$  et  $j$  sont regroupés au sein d’un nouvel amas  $k$ . Cet algorithme s’arrête quand il n’y a plus d’amas à combiner.

Les jets peuvent aussi être reconstruits à partir de **traces chargées** uniquement (ou *track-jet*). Dans ce cas, seules les traces et les impulsions sont utilisées. L’algorithme anti- $k_t$  permet l’association des traces pour la définition d’un jet. Ces traces doivent avoir une impulsion transverse supérieure à 500 MeV, être proche du vertex primaire et correspondre à au moins un impact dans la couche la plus interne du détecteur à pixels, un impact dans la couche la plus externe du détecteur à pixel ou la plus interne du SCT et 6 impacts dans le SCT.

Enfin à partir de la simulation Monte-Carlo, des **jets vrais** peuvent être reconstruits. Ces jets sont

1. La rapidité est définie comme  $y = \frac{1}{2} \log \left( \frac{E+p_z}{E-p_z} \right)$  avec  $E$  l’énergie de l’amas et  $p_z$  son impulsion selon l’axe du faisceau.

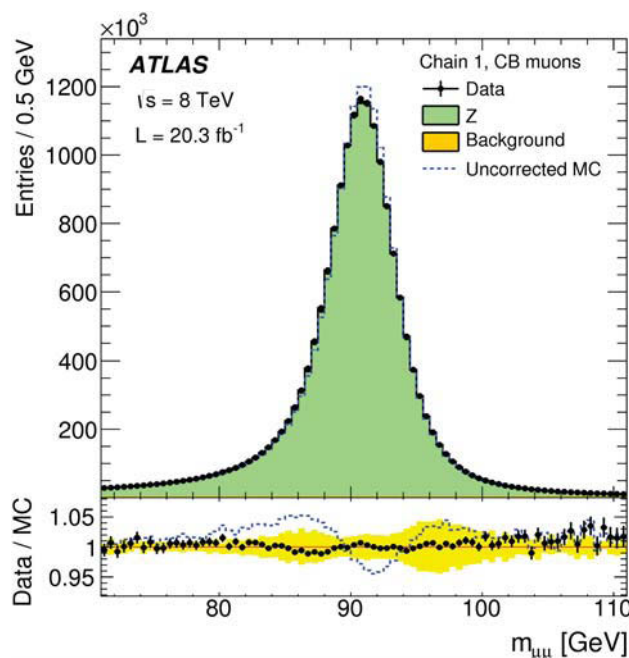


FIGURE 4.6 – Distributions de la masse invariante dimuon pour des événements  $Z \rightarrow \mu\mu$  obtenus dans les données et à partir de simulation MC. Les traits pointillés bleus représentent la simulation non corrigée tandis que les histogrammes pleins correspondent à la simulation avec correction de l’impulsion des muons. L’encadré du bas montre le rapport entre les données et les prédictions de la simulation.

définis en utilisant comme données d’entrée la simulation de particules avec un temps de vie  $\tau$  tel que  $c\tau > 10$  mm et qui ne sont ni des muons, ni des neutrinos.

#### 4.3.3.1 Etalonnage de l’énergie des jets

L’échelle d’énergie de référence des calorimètres d’ATLAS est une échelle électromagnétique ce qui signifie que les électrons verront leur énergie déposée bien reproduite. Toutefois les jets contenant des particules hadroniques, d’autres effets doivent être pris en compte ce qui amène à un étalonnage nécessaire de l’énergie des jets [63]. Ainsi l’étalonnage de l’énergie des jets va prendre en compte différents effets tels que la non-compensation du calorimètre pour corriger la différence entre les échelles d’énergie électromagnétique et hadronique, la présence de matière inactive dans le calorimètre ou encore la perte d’énergie par effet de bord dans le calorimètre (*Leakage*).

La première étape de cet étalonnage consiste à corriger la direction du jet pour pointer vers le vertex primaire ce qui améliore la résolution angulaire des jets. Ensuite l’énergie est corrigée de l’effet d’empilement en extrayant la densité d’énergie due à cet effet. L’étape suivante est l’extraction de l’échelle d’énergie des jets (*Jet Energy Scale* ou *JES*) à partir de simulation Monte-Carlo. Cette échelle d’énergie est extraite pour amener l’énergie du jet reconstruit à l’énergie des jets vrais. En outre une correction est aussi effectuée sur la distribution en  $\eta$  des jets lors de cette étape. À l’étape suivante la résolution en énergie des jets est améliorée selon la méthode GSC (*Global sequential calibration*). Les réponses en énergie des jets avant et après étalonnage sont présentées sur la Figure 4.7. Finalement un dernier étalonnage est fait en comparant le rapport des impulsions d’un jet et d’un autre objet (photon, boson Z) dans la simulation et dans les données. La différence entre les deux rapports est alors utilisée comme facteur d’étalonnage.

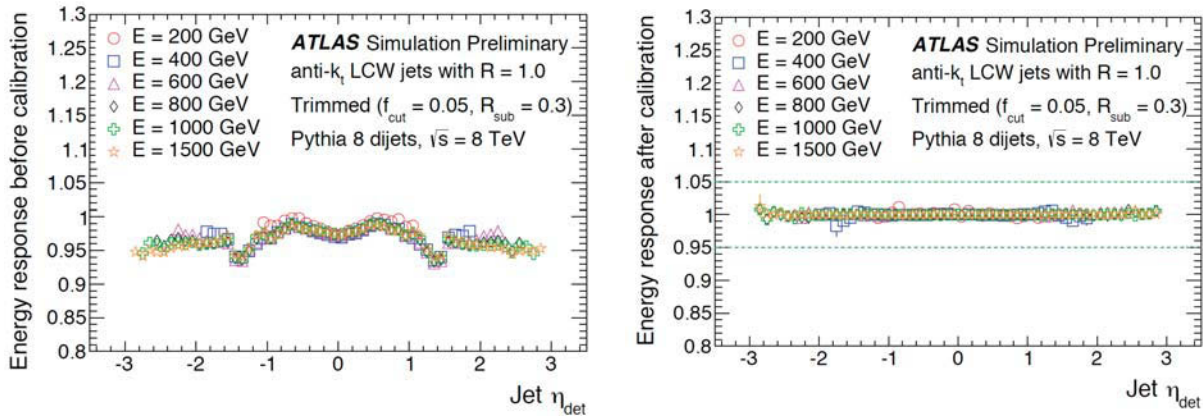


FIGURE 4.7 – Réponse en énergie des jets avant étalonnage (a) et après étalonnage (b) de l'échelle en énergie des jets.

#### 4.3.4 Les jets étiquetés $b$

Les jets provenant de hadrons beaux sont reconstruits à partir d'une analyse multivariée (MVA) qui utilise des informations telles que le paramètre d'impact des traces associées au jet et les propriétés topologiques du vertex secondaire du jet. En effet, les hadrons beaux ont un temps de vie suffisamment grand pour parcourir une distance mesurable avant leur désintégration. L'existence d'un vertex secondaire lié à un jet est donc un indice de la présence d'un quark  $b$ . Ces informations sont combinées selon trois variables IP3D, SV1 et JetFitterCombNN [64] dans l'algorithme MV1. Différents points de fonctionnement, correspondant à différentes valeurs de MV1, seront utilisés par la suite, un premier à 77% d'efficacité et un second à 70%.

#### 4.3.5 Les taus hadroniques

Les leptons  $\tau$  se désintégrant en hadrons vont être reconstruits [65] à partir des dépôts d'énergie dans les calorimètres et des traces dans le détecteur interne. Les candidats doivent avoir 1 ou 3 traces avec une charge totale unitaire. Seuls les candidats avec une impulsion transverse supérieure à 25 GeV et  $|\eta| < 2.5$  sont considérés. Un BDT est utilisé pour diminuer la contamination en jets. Les variables utilisées en entrée de ce BDT sont, par exemple, l'énergie et la forme du dépôt d'énergie dans les calorimètres, la masse invariante de toutes les traces dans l'objet ou encore la masse invariante de toutes les traces ainsi que des deux amas électromagnétiques les plus énergétiques dans l'objet reconstruit.

#### 4.3.6 L'énergie transverse manquante

Le vecteur impulsion transverse manquante est reconstruit [66] comme étant moins la somme vectorielle de l'impulsion transverse de tous les objets reconstruits et étalonnés et des autres dépôts d'énergie dans le détecteur. L'énergie transverse manquante correspondra à la norme de ce vecteur. Cette grandeur est le signe de la présence d'un neutrino dans le détecteur puisque celui-ci n'interagit pas avec la matière et ne laisse pas de traces dans le détecteur.

#### 4.3.7 BDT sur l'isolation des électrons et des muons

##### 4.3.7.1 Définition

Dans le cadre de l'analyse des données combinées de 2015 et 2016, un BDT (*Boosted Decision Tree*) est utilisé pour obtenir une meilleure isolation des leptons. Son but est d'identifier les leptons

provenant de désintégrations de saveurs lourdes. Dans le cadre de l'analyse et de la recherche du processus  $t\bar{t}H$  décrit dans la suite du document, les leptons provenant de hadrons beaux<sup>2</sup> découlent majoritairement de la désintégration du processus  $t\bar{t}$ . Pour limiter la présence de ce processus, le BDT a pour but d'identifier ces leptons non-prompts provenant de hadrons beaux à partir du temps de vie associé à un *track-jet* lui-même associé au lepton non-prompt reconstruit.

Ainsi un ensemble de variables est défini comme variable d'entrée du BDT. Ces variables sont :

- le nombre de traces dans le *track-jet* et l'isolation du lepton dans le calorimètre.
- trois variables liées à l'étiquetage d'un jet comme issu d'un hadron beau.
- le rapport entre l'impulsion transverse du lepton et celle du *track-jet*.
- la distance entre le jet et le lepton.

L'optimisation des variables a été réalisée sur une simulation du processus  $t\bar{t}$ . La Figure 4.8 présente la variable de sortie du BDT pour l'isolation des électrons. Les événements pour lesquels la valeur du BDT sera inférieure à -0.5 seront considérés.

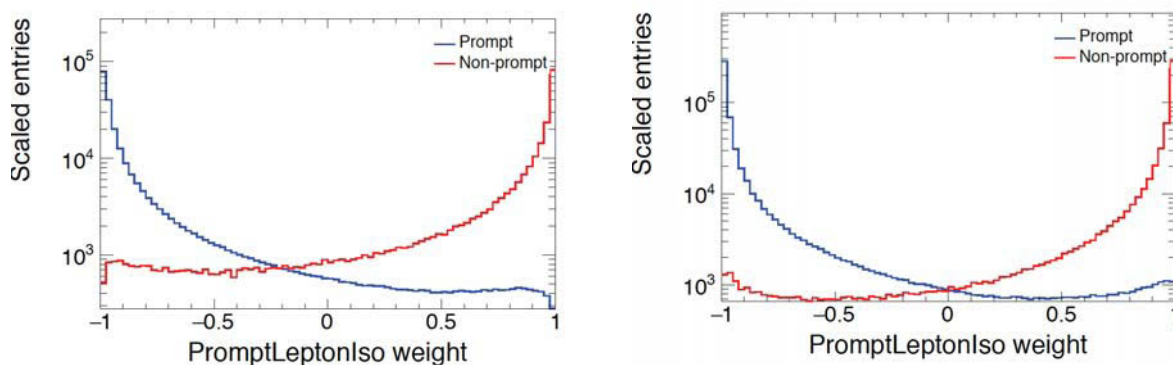


FIGURE 4.8 – Distributions de la variable de sortie du BDT pour l'isolation des électrons (a) et des muons (b) telles qu'attendues pour des processus contenant des leptons prompts (en bleu) et telles qu'attendues pour des électrons provenant de désintégrations de hadrons de saveurs lourdes (en rouge).

#### 4.3.7.2 Facteurs d'échelle

Des facteurs d'échelles sont extraits pour ce nouveau point de fonctionnement de l'isolation des leptons à partir d'événements  $Z + jets$ . Ces facteurs sont estimés en fonction de l'impulsion transverse et de la distance entre le lepton et le jet le plus proche pour les muons et en fonction de l'impulsion transverse et de la pseudorapidité  $\eta$  pour les électrons. Ainsi les facteurs d'échelles pour cette isolation sont compris entre 0.92 à bas  $p_T$  et 0.99 à haut  $p_T$  pour les muons et entre 0.9 et 0.98 pour les électrons.

La Fig. 4.9 présente l'efficacité de reconstruction en fonction de l'impulsion transverse des muons dans un bin en  $\eta$  donné.

La Fig. ?? présente l'efficacité de reconstruction en fonction de la pseudorapidité pour les électrons dans un bin spécifique en  $p_T$ .

2. Par la suite la dénomination de lepton non-prompt sera utilisée.

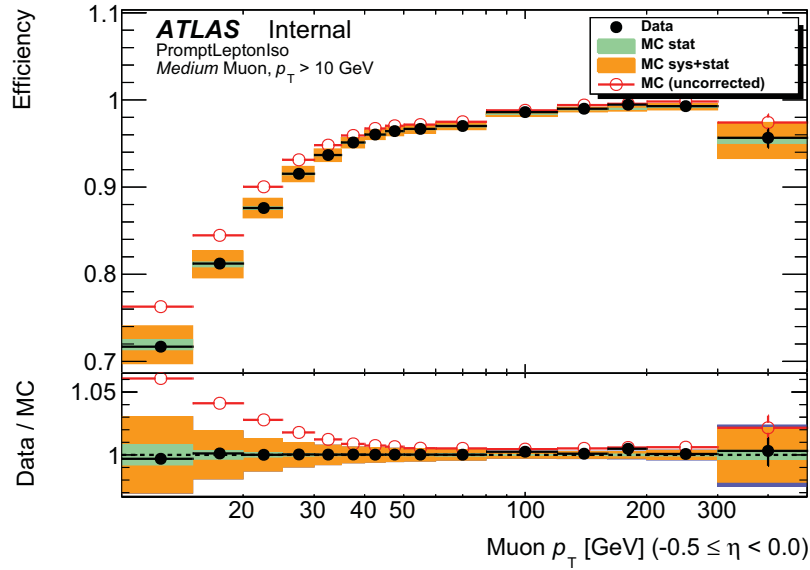


FIGURE 4.9 – Comparaison de l’efficacité de reconstruction pour les muons passant la sélection sur le BDT dans les données et la simulation MC corrigée en fonction de l’impulsion transverse pour les muons tel que ( $-0.5 < \eta < 0$ ).

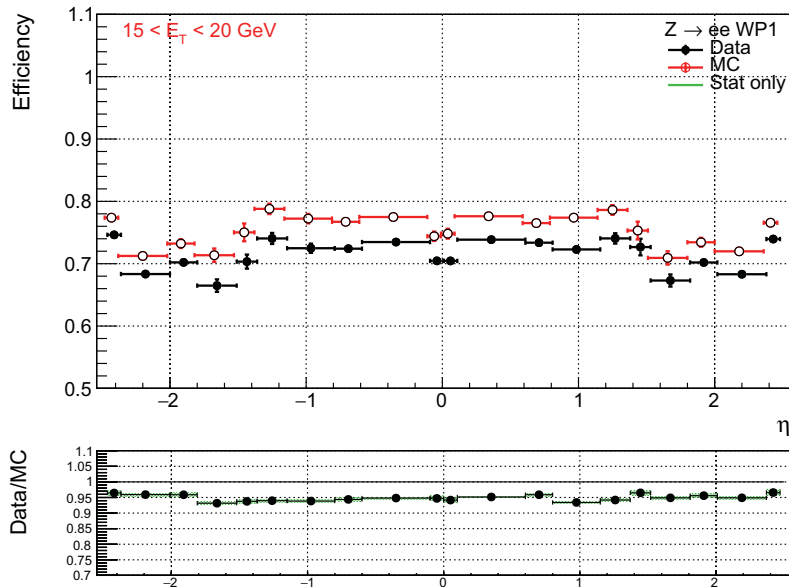


FIGURE 4.10 – Comparaison de l’efficacité de reconstruction pour les électrons passant la sélection sur le BDT dans les données et la simulation MC en fonction de la pseudo-rapidité pour les électrons tel que ( $15 < p_T < 20$  GeV).

# Mauvaise identification de la charge des électrons

## 5.1 Introduction

Un bruit de fond important pouvant contribuer à un état final avec deux leptons de même charge sont les processus du modèle standard qui produisent deux leptons de charges opposées tels que  $W^+W^-$ ,  $Z + jets$  ou  $t\bar{t}$  et dont un des leptons voit sa charge électrique mal reconstruite.

Pour les muons, ce bruit de fond est négligeable et ne sera pas considéré par la suite. En effet le taux de mauvaise identification de la charge est seulement affecté par la courbure de la trace qui est bien mesurée grâce au long bras de levier du système de détection des muons. Ainsi ce bruit de fond est estimé pour les états finals avec deux électrons ou avec un électron et un muon.

Cette mauvaise reconstruction de la charge électrique d'un électron peut avoir deux origines :

- Un rayonnement Bremsstrahlung qui produit un trident d'électrons ( $e^\pm \rightarrow e^\pm \gamma \rightarrow e^\pm e^+ e^-$ ) dont la charge finale reconstruite du trident est basée sur la trace du mauvais électron, d'où une mauvaise reconstruction de la charge. Ce processus est majoritaire, excepté pour des électrons de très hautes impulsions transverses ( $> 500$  GeV). La fraction d'électrons en trident dépend de la quantité de matière traversée par l'électron. Puisque dans le détecteur cette quantité varie en fonction de  $\eta$  (voir fig 5.1), une forte dépendance des taux de mauvaise identification de la charge en fonction de cette variable est attendue.
- Un trajectoire très peu courbée de la particule. Cet effet est important pour des électrons de hautes impulsions transverses. Ainsi une dépendance des taux de mauvaise identification de la charge est aussi attendue en fonction de l'impulsion transverse.

La modélisation de ces effets dépend fortement de la modélisation de l'alignement du détecteur, de celle de la quantité de matière présente ainsi que du bruit dans les détecteurs internes. Cette forte dépendance, combinée à l'imperfection de la simulation de ces effets, rend la précision de la simulation de la mauvaise estimation de la charge des électrons insuffisante<sup>1</sup>.

La mauvaise identification de la charge dans ATLAS a été estimée depuis les données dans de précédentes analyses [67]. J'ai utilisé la méthode dite de vraisemblance à 2 dimensions préexistante à cette thèse en l'appliquant aux données du Run 2. J'y ai apporté une évolution en utilisant une méthode à 3 dimensions, qui est également décrite dans ce chapitre. La stratégie d'estimation de ce bruit de fond est présentée dans les sections 5.2, 5.3 et 5.6. Puis les estimations des taux dans les données pour  $\sqrt{s} = 13$  TeV sont décrites dans les sections 5.8 et 5.9.

## 5.2 Stratégie d'estimation des taux de mauvaise identification de la charge des électrons

Soit un événement contenant deux électrons de charges opposées dans l'état final, posant  $\varepsilon$  le taux de mauvaise identification de la charge d'un électron, cet événement peut être reconstruit de trois manières différentes :

- $e^+e^-$  sans mauvaise reconstruction de la charge avec une probabilité de  $(1 - \varepsilon)^2$ ,

1. une incertitude empirique de 50% est généralement associée à la modélisation de la QMisID

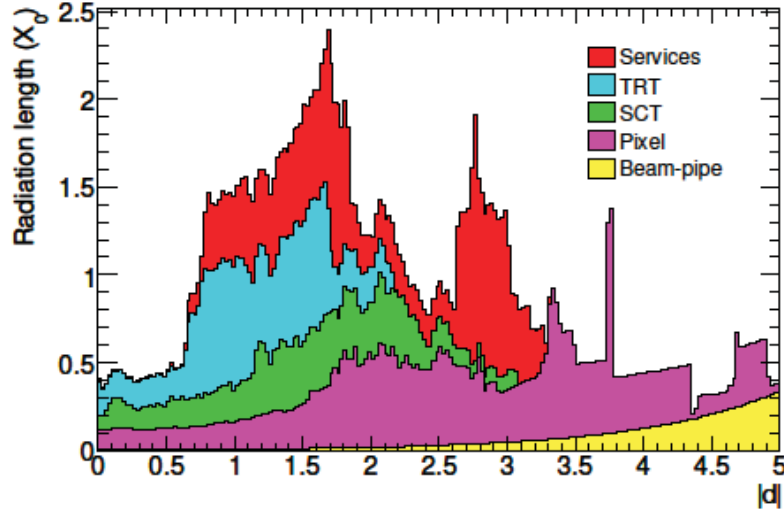


FIGURE 5.1 – Distribution de la quantité de matière en fonction de  $\eta$  après le détecteur interne

- $e^+e^-$  avec pour les deux électrons une mauvaise reconstruction de la charge avec une probabilité de  $\varepsilon^2$ ,
- $e^\pm e^\pm$  avec pour un des deux électrons une mauvaise reconstruction de la charge avec une probabilité de  $2\varepsilon(1 - \varepsilon)$ .

Ainsi pour  $N$  événements avec réellement deux électrons de charges opposées, les événements reconstruits seront :

- $N^{OS} = (1 - 2\varepsilon + 2\varepsilon^2)N$  événements de charges opposées
- $N^{SS} = 2\varepsilon(1 - \varepsilon)N$  événements de même charge

Connaissant le taux  $\varepsilon$ , il est possible de calculer le nombre d'événements reconstruits de même charge  $N^{SS}$  à partir du nombre d'événements reconstruits de charges opposées  $N^{OS}$  :

- $N^{ss} = \frac{\varepsilon_i + \varepsilon_j - 2\varepsilon_i \varepsilon_j}{1 - \varepsilon_i - \varepsilon_j + 2\varepsilon_i \varepsilon_j} N^{os}$  pour le canal  $ee$ ,
- $N^{ss} = \frac{\varepsilon}{1 - \varepsilon} N^{os}$  pour le canal  $e\mu$ ,

où  $\varepsilon_i$  et  $\varepsilon_j$  sont les taux de mauvaise identification de la charge pour chacun des électrons. Ces taux  $\varepsilon$  peuvent être estimés pour chaque électron selon plusieurs paramètres simultanément tels que :

- $\eta$  puisque la quantité de matière traversée a un impact direct sur la production de trident d'électron. Concernant cette grandeur, les taux de mauvaises estimations de la charge ne seront pas estimés dans la région  $|\eta| \in [1.37, 1.52]$  car cette région n'est pas instrumentée et très peu d'électrons y sont reconstruits.
- $p_T$  car la mauvaise identification de la charge d'un électron va augmenter avec des courbures faibles, proportionnelles à l'impulsion transverse.
- la qualité de l'électron, c'est-à-dire selon que l'électron passe ou non certains critères d'identification et d'isolation. Le critère choisi ici est binaire selon que l'électron passe ou non ces critères.

Deux options sont considérées pour l'estimation des taux de mauvaise reconstruction de la charge dans la suite de ce chapitre. Dans un premier temps une estimation en deux dimensions en  $\eta$  et  $p_T$  est réalisée sur  $13.2fb^{-1}$  de données. Puis une analyse à 3 dimensions est faite incluant la qualité de l'électron sur l'ensemble des données recueillies en 2015 et 2016. L'estimation selon les observables  $\eta$  et  $p_T$  est nécessaire car la distribution des événements Z+jets utilisés pour l'estimation des taux et des événements  $t\bar{t}$  présents dans l'analyse sont différentes pour ces deux observables.

L'ajout de la qualité de l'électron comme dernier paramètre a pour but d'extraire les taux de mauvaises identifications des électrons à la fois pour les leptons *Tight* et *anti-Tight*. Or le nombre



d'événements avec deux électrons *anti-Tight* n'est pas suffisant pour avoir un résultat cohérent avec deux paramètres  $p_T$  et  $\eta$ . Ainsi l'inclusion de la qualité des électrons comme paramètre permet d'utiliser aussi les événements avec un électron *Tight* et un électron *anti-Tight*. Ce critère supplémentaire pourrait aussi être séparé en plusieurs sous-régions pour lesquelles les taux de mauvaise reconstruction de la charge sont estimés. Le choix de n'avoir que deux sous-région est motivé par le fait que la définition des électrons *Tight* est suffisamment sévère pour qu'aucune dépendance ne soit observée sur les taux pour ces électrons.

### 5.3 Estimation des taux de mauvaise identification de la charge des électrons par maximum de vraisemblance

Les événements utilisés pour extraire les taux de mauvaise identification de la charge des électrons dans les données sont des événements ayant exactement deux électrons qui correspondent à des désintégrations d'un boson  $Z$ . Pour sélectionner ces événements une sélection est appliquée à la masse invariante des deux électrons. En outre, pour estimer la présence de processus donnant fortuitement deux leptons avec une masse invariante proche de celle du boson  $Z$ , une méthode de soustraction des événements est appliquée. Pour cela, le spectre de la masse invariante des deux leptons est divisé en trois régions de largeurs égales  $A$ ,  $B$  et  $C$  où la région  $B$  est centrée sur la masse reconstruite la plus probable du boson  $Z$ . Le nombre d'événements comptés dans les régions  $A$  et  $C$  est soustrait au nombre d'événements obtenu dans la région  $B$ . Au final le nombre d'événements considéré venant de désintégration d'un boson  $Z$  sera :

$$N_Z = n_B - \frac{n_A + n_C}{2}. \quad (5.1)$$

Par ailleurs, le pic de la masse du  $Z$  pour des électrons de même charge électrique est situé à plus basse valeur que dans le cas des événements avec deux électrons de charges opposées à cause de la perte d'énergie lors de la reconstruction des électrons en tridents. Pour prendre en compte cet effet, les trois régions  $A$ ,  $B$  et  $C$  définies pour la soustraction des bruits de fond sont différentes dans le cas d'événements avec deux électrons de même charge ou non. Une illustration de ces régions est donnée sur la Fig. 5.2 et la définition de ces régions est donnée dans la Table 5.1. Un ajustement gaussien dans une région proche de la masse attendue du boson  $Z$  est effectué. Dans les deux cas d'événements  $e^\pm e^\pm$  et  $e^\mp e^\mp$  les régions ont été définies à partir des valeurs centrales de la gaussienne  $\kappa$  et de quatre fois la déviation standard  $\sigma$ . Ainsi la région  $B$  est définie comme  $\kappa \pm 4\sigma$ . Les régions  $A$  et  $C$  sont définies de façon à avoir la même largeur que la région centrale.

Région	$A$	$B$	$C$
$e^\pm e^\pm$	[51.7,76.5]	[76.5,101.3]	[101.3,126.0]
$e^\mp e^\mp$	[54.7,78.5]	[78.5,102.3]	[102.3,126.0]

TABLE 5.1 – Définition des régions  $A$ ,  $B$  et  $C$  utilisées pour la soustraction des bruits de fond.

Après avoir soustrait le bruit de fond, la méthode utilisée pour extraire les taux de mauvaise identification des leptons est la méthode par maximum de vraisemblance. Ainsi en considérant un seul paramètre pour l'estimation des taux de mauvaise reconstruction de la charge des électrons tel que la pseudo-rapacité  $\eta$ , la probabilité d'avoir un nombre d'événements avec deux électrons de même charge électrique  $N_{ss}^{ij}$  avec des électrons dans les régions  $i$  et  $j$  en  $\eta$  peut être écrit en fonction du nombre total d'événement  $N^{ij}$  :

$$N_{ss}^{ij} = N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j). \quad (5.2)$$

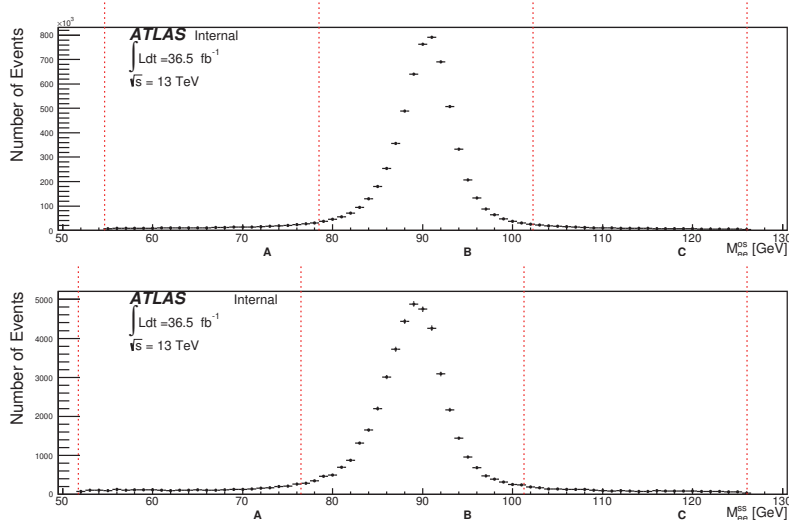


FIGURE 5.2 – Distribution de la masse invariante diélectron pour des événements avec deux électrons de charges opposées (haut) ou de même charge (bas) dans les données. Les traits verticaux représentent les limites des régions utilisées pour la soustraction du bruit de fond.

Si tous les événements avec deux électrons de même charge dans la région de signal considérée sont produits par mauvaise reconstruction de la charge, alors  $N_{ss}^{ij}$  est décrit par une distribution de Poisson

$$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad (5.3)$$

où  $k$  est le nombre d'événements observés, soit  $k = N_{ss}^{ij}$ , et  $\lambda$  est le nombre attendu, soit  $\lambda = N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)$ . Ainsi la probabilité totale d'avoir  $N_{ss}^{ij}$  événements avec deux leptons de même signe est :

$$P(\varepsilon_i, \varepsilon_j | N_{ss}^{ij}, N^{ij}) = \frac{[N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)}}{N_{ss}^{ij}!}. \quad (5.4)$$

La fonction de vraisemblance  $L$  pour tous les événements est obtenue par la combinaison de toutes les configurations des paires d'électrons en  $\eta$  :

$$L(\varepsilon | N_{ss}, N) = \prod_{i,j} \frac{[N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)]^{N_{ss}^{ij}} e^{-N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)}}{N_{ss}^{ij}!}, \quad (5.5)$$

où les taux  $\varepsilon_i$  et  $\varepsilon_j$  sont obtenus par maximisation de la fonction de vraisemblance. La fonction  $-\ln L$  est utilisée pour simplifier la minimisation. Les termes qui ne dépendent pas des taux  $\varepsilon_i$  et  $\varepsilon_j$  sont enlevés à cette étape. Ainsi, la fonction qui est finalement minimisée est donnée par :

$$-\ln L(\varepsilon | N_{ss}, N) \approx \sum_{i,j} \ln[N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j)] N_{ss}^{ij} - N^{ij}(\varepsilon_i + \varepsilon_j - 2\varepsilon_i\varepsilon_j). \quad (5.6)$$

La méthode par maximum de vraisemblance peut être étendue pour mesurer les taux de mauvaise identification de la charge en fonction de deux ou trois paramètres. Dans le cas de trois paramètres, la probabilité d'avoir un événement avec deux électrons de même charge est  $\varepsilon_{i,k,m} + \varepsilon_{j,l,n} - 2\varepsilon_{i,k,m}\varepsilon_{j,l,n}$ , avec les trois indices représentant les régions de chacun des trois paramètres considérés où sont situés les électrons. Ainsi, l'équation 5.6 devient :

$$-\ln L(\varepsilon | N_{ss}, N) \approx \sum_{i,j,k,l,m,n} \ln[N^{ij,k,l,m,n}(\varepsilon_{i,k,m} + \varepsilon_{j,l,n})] N_{ss}^{ij,k,l,m,n} - N^{ij,k,l,m,n}(\varepsilon_{i,k,m} + \varepsilon_{j,l,n}). \quad (5.7)$$

## 5.4 Choix des intervalles en $\eta$ et $p_T$ pour l'estimation des taux de mauvaise identification de la charge des électrons

Les taux de mauvaise estimation de la charge des électrons sont extraits a minima en deux dimensions selon  $\eta$  et  $p_T$  des électrons. Le nombre d'événements de données disponible étant limité, un choix est fait dans la définition des régions selon ces deux variables. Une étude est effectuée en utilisant des échantillons Monte-Carlo de processus Z+jets avec un nombre de région important en  $p_T$  et  $\eta$ . Les taux de mauvaise identification de la charge sont extraits à partir des informations Monte-Carlo vraies<sup>2</sup> dans ces régions et seulement en 1 dimension. Les intervalles adjacents qui ont des taux de mauvaise identification avec des valeurs proches vont être regroupés. La figure 5.3 donne les taux de mauvaise identification de la charge dans des intervalles en  $\eta$  et  $p_T$  fins.

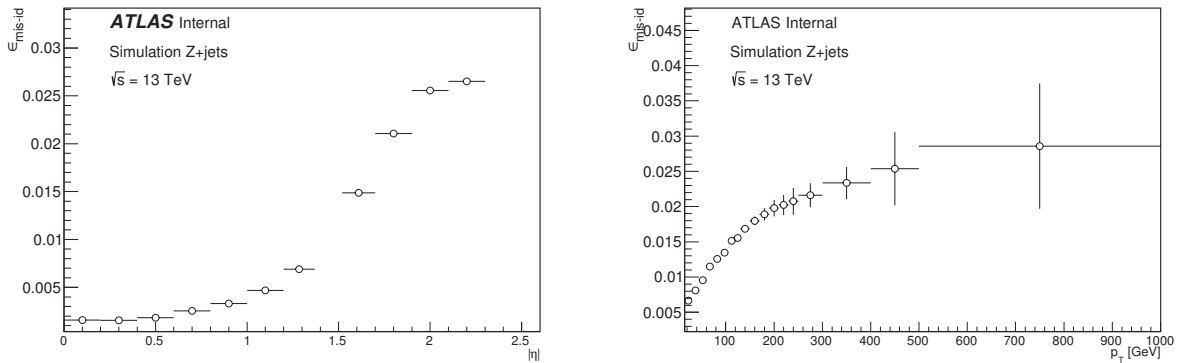


FIGURE 5.3 – Taux de mauvaise identification de la charge des électrons en fonctions de  $\eta$  et  $p_T$  des électrons pour des intervalles fins obtenus à partir de l'information vraies contenues dans des échantillons Monte-Carlo Z+jets.

Les limites des régions choisies sont :

- en  $p_T$  : [10,60,90,130,1000] GeV
- en  $|\eta|$  : [0, 0.6, 1.1, 1.37, 1.52, 1.7, 2.3, 2.6] pour l'analyse avec  $13.2fb^{-1}$  de données ou [0, 0.6, 1.1, 1.37, 1.52, 2.0, 2.6] pour l'analyse avec  $36.1fb^{-1}$  de données. Dans ce deuxième cas la définition des régions à haute valeur en  $|\eta|$  a dû être revue pour avoir un nombre d'événements suffisant dans chaque intervalle pour une estimation cohérente.

En outre pour l'analyse avec  $36.1fb^{-1}$  de données, deux régions sont définies selon la qualité des électrons : *tight* ou *anti-tight*. Un électron *tight* passe les critères d'identification TightLH, et la sélection sur les BDT pour les leptons prompts et la mauvaise identification de la charge des électrons définis dans la section 4. Un électron *anti-tight* ne passe pas l'un de ces critères.

## 5.5 Estimation des taux de mauvaise identification de la charge à haute impulsion transverse

La méthode par maximum de vraisemblance est robuste tant que le nombre d'événement dans les différentes régions, définies en  $\eta$  et  $p_T$  par exemple, est suffisamment grand. Dans le cadre de l'analyse réalisée avec les données 2015 et les premières données 2016 (pour une luminosité intégrée de  $13.2fb^{-1}$ ), le nombre d'événements était faible pour la région correspondant aux électrons à

2. Les informations Monte-Carlo vraies contiennent des information sur les objets générés avant simulation de la reconstruction par le détecteur, ce qui permet de savoir si un électron a vu sa charge mal reconstruite

haut  $p_T$ . Pour pallier ce problème, une autre méthode a été utilisée pour extraire les taux dans cette région.

Dans cette méthode, les taux de mauvaise identification de la charge des électrons dans la région à plus haut  $p_T$  sont obtenus par extrapolation de ceux mesurés dans une région plus faible en  $p_T$  avec un facteur correctif dérivé à partir de la simulation MC du processus  $t\bar{t}$ . Ainsi le taux de mauvaise identification de la charge des électrons sera donné dans la région à haut  $p_T$  par :

$$\varepsilon(|\eta|, p_T > 130\text{GeV}) = \varepsilon(|\eta|, p_T \in [90, 130]\text{GeV}) \times \alpha_{t\bar{t}}(|\eta|) \quad (5.8)$$

avec le facteur  $\alpha_{t\bar{t}}$  défini comme :

$$\alpha_{t\bar{t}} = \frac{\varepsilon(|\eta|, p_T > 130\text{GeV})_{t\bar{t}}}{\varepsilon(|\eta|, p_T \in [90, 130]\text{GeV})_{t\bar{t}}} \quad (5.9)$$

où les différents taux sont extraits à partir des informations vraies stockées dans la simulation Monte-Carlo.

## 5.6 Incertitudes associées à l'estimation par maximum de vraisemblance

Lors de l'estimation de la mauvaise reconstruction de la charge des électrons, plusieurs sources d'incertitudes doivent être prises en compte. L'incertitude finale sur les taux sera la somme quadratique de ces différentes contributions :

- Une incertitude statistique venant du nombre d'événements limité dans chacune des régions des paramètres considérés  $\sigma_\varepsilon(\alpha, \beta, \gamma)$  où  $\alpha, \beta, \gamma$  sont les 3 paramètres considérés pour l'extraction des taux en trois dimensions par exemple.
- La différence entre les taux obtenus par méthode de vraisemblance appliqués sur des événements  $Z \rightarrow e^+e^-$  simulés et les taux obtenus à partir de l'information vraie présente dans ces simulations. Cette incertitude n'est prise en compte que dans les régions des paramètres  $(\alpha, \beta, \gamma)$  où les taux extraits par les deux méthodes ne sont pas compatibles dans les incertitudes. Cette incertitude permet de tester la validité de la méthode du maximum de vraisemblance utilisée sur les données.
- La stabilité des taux obtenus en variant la définition de la région central B quand le bruit de fond est soustrait. Dans ce cas, les trois régions A, B et C sont modifiées d'un  $\sigma$ . Cette variation de la définition de la région B teste la forme du bruit de fond présent et la validité de la soustraction de ce bruit de fond.
- Dans le cas de l'analyse effectuée pour  $13.2\text{fb}^{-1}$  de données, une incertitude supplémentaire est ajoutée pour prendre en compte le nombre d'événements disponibles dans la simulation  $t\bar{t}$  utilisée pour l'extrapolation des taux de mauvaise reconstruction de la charge dans la région à plus haut  $p_T$ .

## 5.7 Tests de consistances

Des études supplémentaires sont effectuées pour tester la dépendance des taux de mauvaise identification de la charge des électrons en fonction de variables supplémentaires.

Dans un premier temps, en utilisant une simulation Monte-Carlo du processus  $t\bar{t}$  et l'information vraie disponible, la dépendance des taux  $\varepsilon$  en fonction du nombre de jets est testée. En effet, les événements utilisés pour mesurer les taux  $\varepsilon$  se caractérisent par une faible multiplicité en jets. La Figure 5.4 montre l'absence de dépendance significative entre ces deux grandeurs qui n'est donc pas prise en compte dans l'analyse. Un test similaire est effectué pour observer si une corrélation existe entre les taux  $\varepsilon$  et la distance d'un lepton au jet le plus proche. Ce test mesure également la

sensibilité des taux à la présence de leptons non-prompts. La sensibilité à l'effet d'empilement est également testée via la dépendance des taux au nombre de vertex par croisement de faisceau. Les Figures 5.5 et 5.6 démontrent qu'ils n'existent pas de dépendance significative entre ces grandeurs.

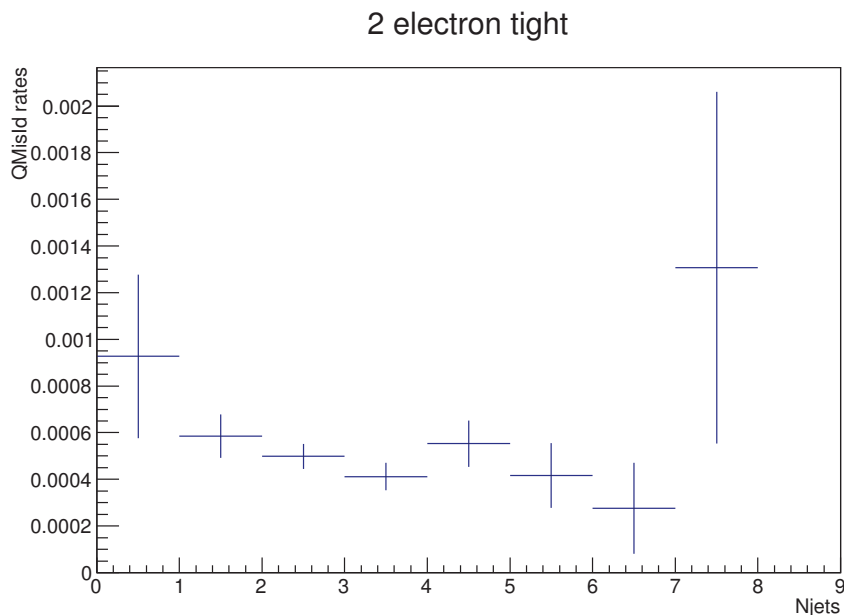


FIGURE 5.4 – Distribution du taux de mauvaise identification de la charge des électrons en fonction du nombre de jets. Une simulation Monte-Carlo du processus  $t\bar{t}$  est utilisée en ne sélectionnant que les événements avec deux électrons *tight*.

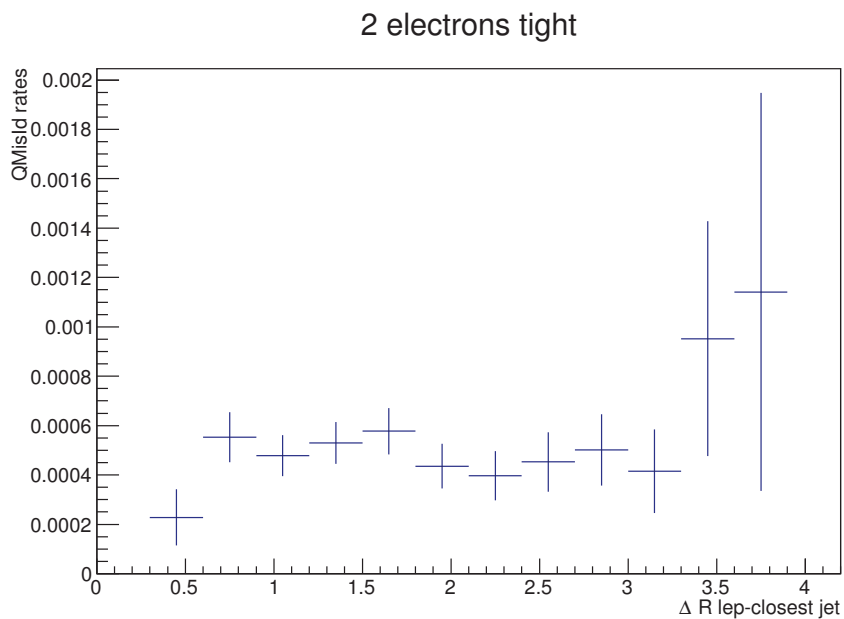


FIGURE 5.5 – Distribution du taux de mauvaise identification de la charge des électrons en fonction de la distance entre les leptons et le jet le plus proche. Une simulation Monte-Carlo du processus  $t\bar{t}$  est utilisée en ne sélectionnant que les événements avec deux électrons *tight*.

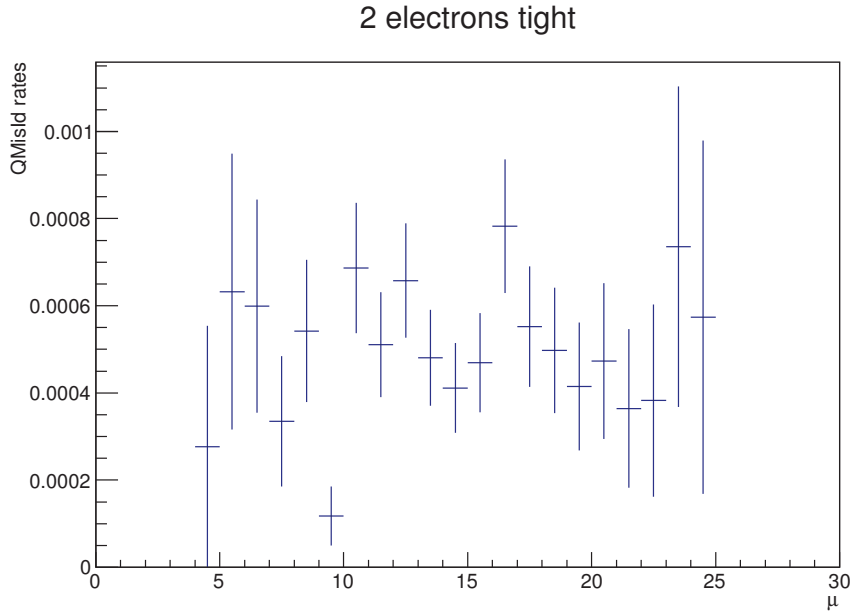


FIGURE 5.6 – Distribution du taux de mauvaise identification de la charge des électrons en fonction du nombre de vertex par interaction. Une simulation Monte-Carlo du processus  $t\bar{t}$  est utilisée en ne sélectionnant que les événements avec deux électrons *tight*.

## 5.8 Estimation des taux de mauvaise estimation de la charge des électrons à partir des données pour $13.2fb^{-1}$ de données à 13TeV

Une première estimation des taux de mauvaise identification de la charge des électrons a été effectuée dans le cadre de l’analyse des données à 13 TeV pour une luminosité intégrée de  $13.2fb^{-1}$  ce qui correspond aux données de l’année 2015 et aux six premiers mois de l’année 2016. Dans cette analyse, les taux de mauvaise identification de la charge des électrons ont été obtenus en fonction des paramètres  $\eta$  et  $p_T$ .

### 5.8.1 Validation de la méthode par maximum de vraisemblance

Pour valider la méthode par maximum de vraisemblance, une comparaison est effectuée entre les taux mesurés par la méthode de maximisation de vraisemblance et par l’utilisation des informations vraies en utilisant des événements  $Z$  simulés. Le rapport des taux obtenus par les deux méthodes est présenté sur la Fig. 5.7. La différence entre les deux méthodes est prise comme incertitude systématique si le rapport n’est pas compatible avec 1 ce qui n’est pas le cas dans le premier bin en  $p_T$  dans la Fig. 5.7.

### 5.8.2 Incertitudes systématiques

La composante principale des incertitudes est l’incertitude statistique liée au nombre d’événements disponible dans les différentes régions en  $\eta$  et  $p_T$ . La contribution de chacune des incertitudes est décrites sur la Fig. 5.8 pour les différentes régions en  $\eta$  et  $p_T$ .

Les valeurs des incertitudes varient selon les bins de 10% à 30% pour les bins à plus hauts  $p_T$ . La composante principale de cette incertitude est le nombre d’événements disponible pour l’extraction des taux et l’incertitude sur le facteur correctif appliqué dans le dernier bin en  $p_T$ .

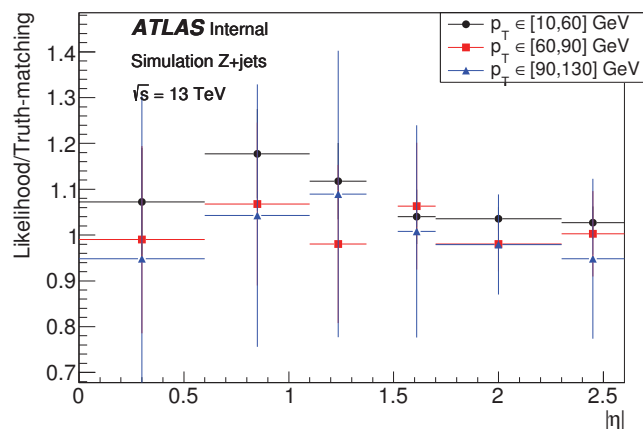


FIGURE 5.7 – Rapport entre les taux de mauvaise identification de la charge des électrons obtenus par méthode de maximum de vraisemblance et à partir des informations Monte-Carlo vraies.

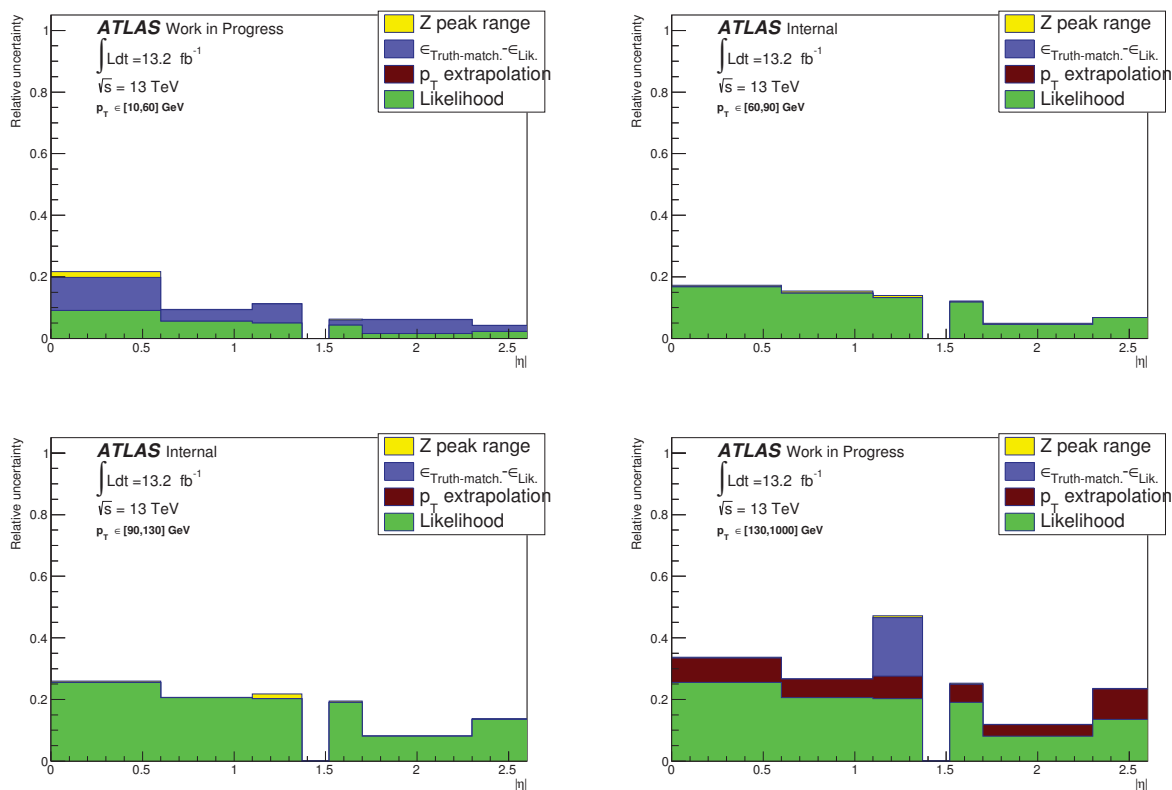


FIGURE 5.8 – Somme quadratique des contributions à l'incertitude systématique relative sur les taux de mauvaises identifications de la charge, pour les différentes régions en  $\eta$  et  $p_T$ .

### 5.8.3 Taux observés

Les taux de mauvaise identification de la charge mesurés dans les données sont présentés sur la Fig. 5.9, où toutes les incertitudes sont présentes.

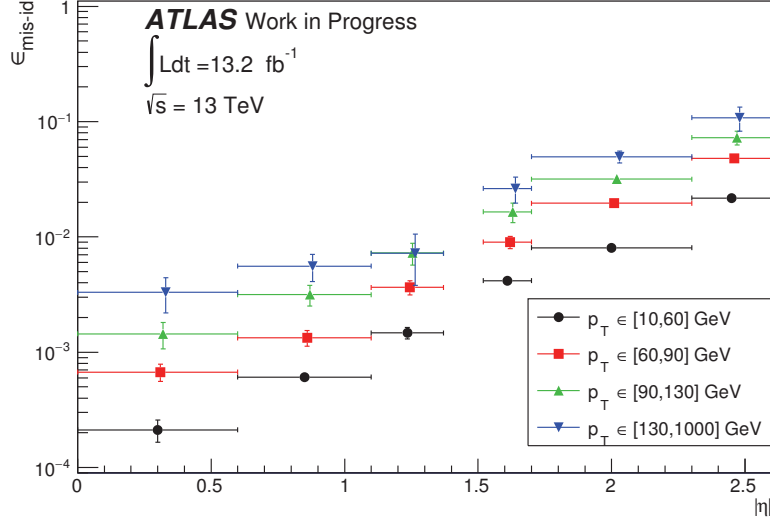


FIGURE 5.9 – Taux de mauvaise identification de la charge des électrons mesurés dans les données par la méthode du maximum de vraisemblance (points noirs, carrés rouges et triangles verts) en fonction des paramètres  $\eta$  et  $p_T$ . Pour les électrons avec un  $p_T > 130 \text{ GeV}$ , le taux de mauvaise identification de la charge a été obtenu par extrapolation des taux dans la région  $p_T \in [90, 130] \text{ GeV}$  avec un facteur obtenu à partir de simulation d'événement  $t\bar{t}$  (triangles bleus).

Les nombres d'événements dans les régions de signal (présentées dans le Chapitre 7) correspondant à ces taux sont :

- $N_{Q_{\text{MisId}}}^{ee}(5j) = 6.9 \pm 1.3.$
- $N_{Q_{\text{MisId}}}^{e\mu}(5j) = 7.1 \pm 1.7.$
- $N_{Q_{\text{MisId}}}^{2\ell SS1\tau}(4j) = 0.24 \pm 0.03.$

Les événements avec une mauvaise estimation de la charge des leptons représentent environ 30% (20%) du bruit de fond total dans le canal  $2\ell SS0\tau ee (e\mu)$ .

### 5.8.4 Test de fermeture

Les taux de mauvaise identification de la charge des électrons sont obtenus à partir d'événement avec deux électrons dont la masse invariante est proche de la masse du boson Z sans aucun critère sur le nombre de jets étiquetés  $b$ . Pour valider l'estimation de nos taux, une comparaison est faite entre la prédiction du nombre d'événement dans une région avec deux électrons de masse invariante proche de la masse du boson Z et un jet étiqueté  $b$  et ce qui est observé dans les données pour cette région.

La comparaison entre les nombres d'événements prédits et observés est donnée dans la Table 5.2.



Bdf MC (VV, $t\bar{t}V, t\bar{t}$ )	QMisId	Total	Data	Data/Pred
$8.46 \pm 1.52$	$183.9 \pm 18.5$	$192.4 \pm 18.6$	200	$1.04 \pm 0.07(\text{donnees}) \pm 0.10(\text{sys})$

TABLE 5.2 – Nombre d’événements observés dans les données dans une région enrichie en boson Z et un jet étiqueté  $b$  comparé aux prédictions. Les incertitudes sur les prédictions MC sont statistiques tandis que l’ensemble des incertitudes décrites précédemment est utilisé pour la prédiction du nombre d’événement avec une mauvaise reconstruction de la charge.

## 5.9 Estimation des taux de mauvaise estimation de la charge des électrons à partir des données pour $36.1fb^{-1}$ de données à 13TeV

Une seconde estimation des taux de mauvaise identification de la charge des électrons a été effectuée dans le cadre de l’analyse des données produite à 13 TeV pour une luminosité intégrée de  $36.1fb^{-1}$  ce qui correspond aux données totales des années 2015 et 2016. Dans cette analyse les taux de mauvaise identification de la charge des électrons ont été obtenus en fonction des paramètres  $\eta$ ,  $p_T$  et la qualité des électrons.

### 5.9.1 Validation de la méthode par maximum de vraisemblance

La comparaison entre les taux obtenus par maximum de vraisemblance ou en utilisant les informations Monte-Carlo vraies est présentée sur la Fig. 5.10. Un bon accord dans les incertitudes est observé exception faite d’une région qui montre une différence plus large (bas  $p_T$  second bin en  $\eta$  pour des électrons *tight*). La différence entre les deux méthodes est prise comme incertitude systématique.

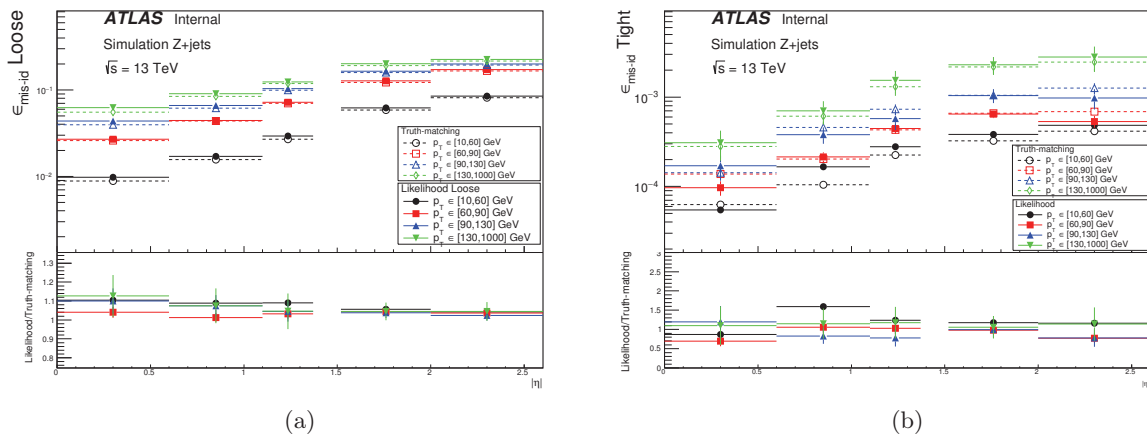


FIGURE 5.10 – Rapport entre les taux de mauvaise identification de la charge des électrons obtenus par méthode de maximum de vraisemblance et à partir des informations Monte-Carlo vraies pour des électrons *antiTight* (a) et des électrons *Tight* (b).

### 5.9.2 Incertitudes systématiques

Les composantes principales des incertitudes varient entre les régions considérées en  $\eta$  et  $p_T$  et selon la qualité des électrons. Ainsi la composante statistique est globalement la plus importante

pour les électrons *Tight*. La contribution de chacune des incertitudes est décrite sur les Fig. 5.11 et 5.12 pour les différentes régions en  $\eta$  et  $p_T$ , respectivement pour les électrons *antiTight* et *Tight*.

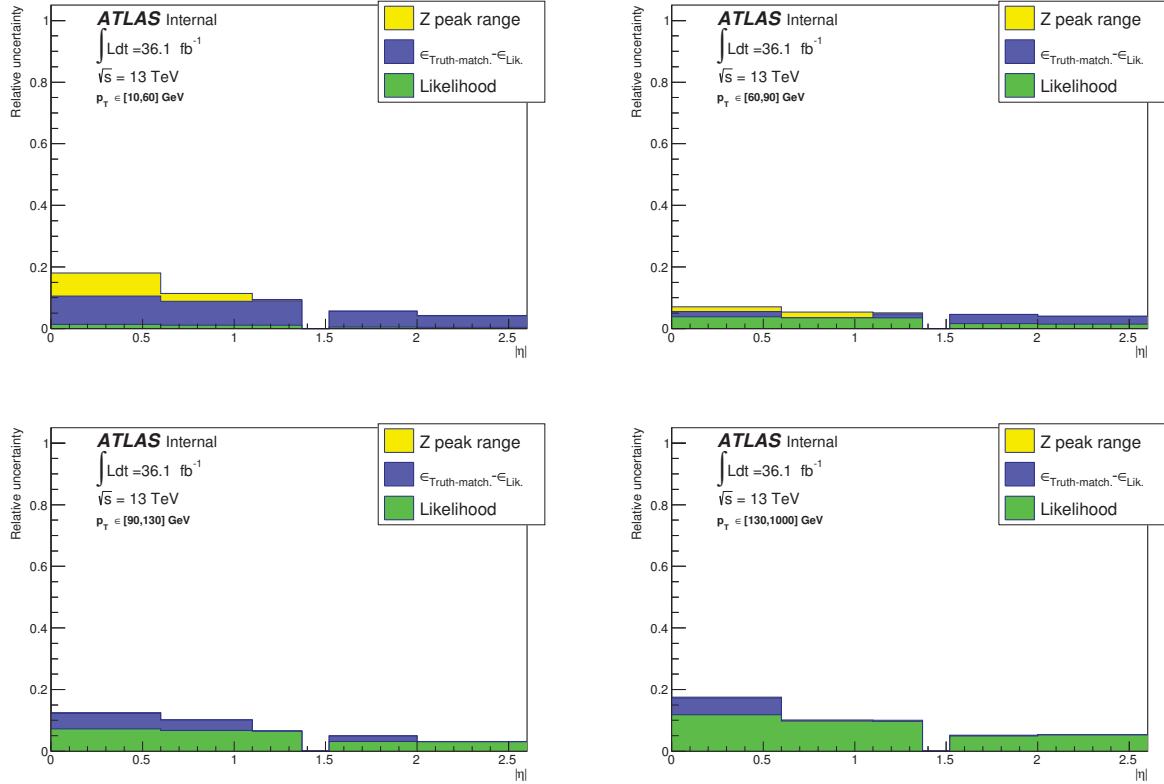


FIGURE 5.11 – Somme quadratique des contributions à l’incertitude systématique relative sur les taux de mauvaises identifications de la charge, pour les différentes régions en  $\eta$  et  $p_T$  pour des électrons *antiTight*

Les incertitudes varient selon les bins en  $\eta$  et  $p_T$  de 10% à 20% pour les électrons *anti-tight* et de 20% à 60% pour les électrons *tight*. Dans ce deuxième cas, la composante principale est le nombre d’événements présents dans les régions utilisées pour l’extraction des taux. L’augmentation de ces incertitudes par rapport à l’analyse à  $13.2 fb^{-1}$  vient de l’utilisation du BDT pour limiter la présence de ce processus, ce qui est compensé par une forte diminution des taux.

### 5.9.3 Taux observés

Les taux de mauvaise identification de la charge mesurés dans les données sont présentés sur les Fig. 5.13, où toutes les incertitudes sont présentes.

Le nombre d’événements avec un électron dont la charge est mal reconstruite est dans les régions de signal de l’ordre de :

- canal  $2lSS0\tau$  ( $ee$ ) avec deux électrons *tight* :  $19.2 \pm 0.2(stat) \pm 6.1(syst)$ .
- canal  $2lSS0\tau$  ( $e\mu$ ) avec deux électrons *tight* :  $14.7 \pm 0.2(stat) \pm 4.8(syst)$ .

Les régions de signal sont différentes de celle de l’analyse à  $13.2 fb^{-1}$  et sont décrites dans le Chapitre 9. Au total la mauvaise identification de la charge des électrons représente moins de 5% du nombre d’événements de bruit de fond dans le canal avec deux leptons de même charge.

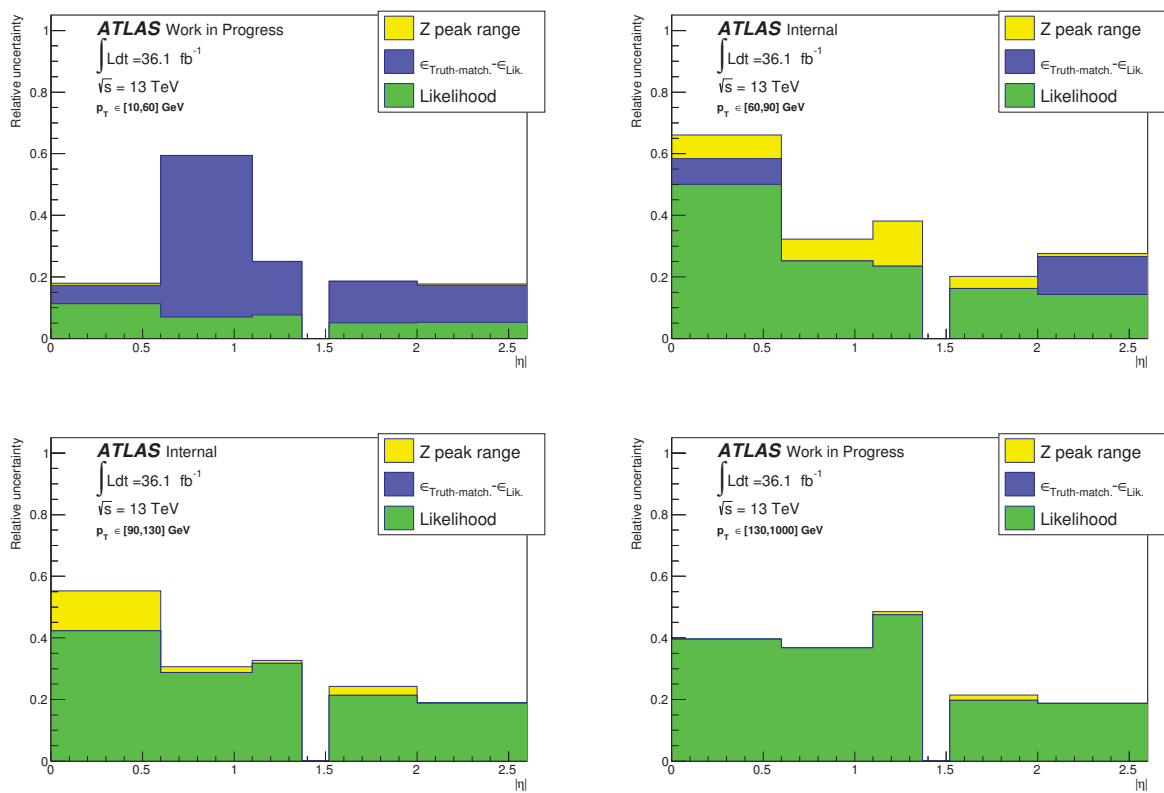


FIGURE 5.12 – Somme quadratique des contributions à l’incertitude systématique relative sur les taux de mauvaises identifications de la charge, pour les différentes régions en  $\eta$  et  $p_T$  pour des électrons *Tight*

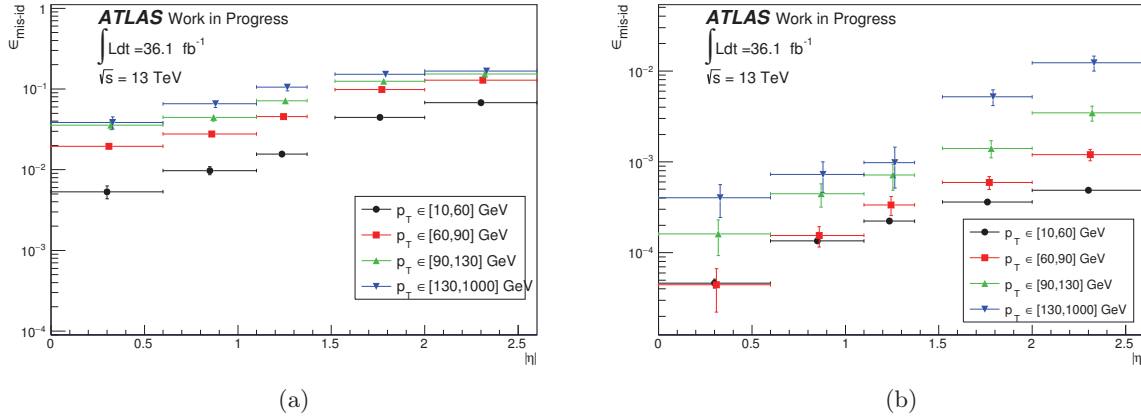


FIGURE 5.13 – Taux de mauvaise identification de la charge des électrons mesurés dans les données par la méthode du maximum de vraisemblance pour des électrons *antiTight* (a) et des électrons *Tight* (b).

### 5.9.4 Test de fermeture

Un test de fermeture est effectué en comparant le nombre d'événements prédits après l'application des taux sur des événements avec deux électrons de charges opposées dans une région riche en boson Z à ce qui est obtenu dans les données dans cette même région avec deux électrons de même charge.

Dans les données, le nombre d'événements observés pour des électrons *antiTight* est de 73 919 pour une prédiction de  $77\,263^{+4790}_{-4827}$  événements. Pour des électrons *Tight* le nombre d'événements observés est de 4 382 pour une prédiction de  $4\,244^{+1173}_{-1173}$  événements. Dans les deux cas les prédictions sont donc compatibles avec ce qui est observé. La masse invariante des deux électrons peut être observée sur la Fig. 5.14.

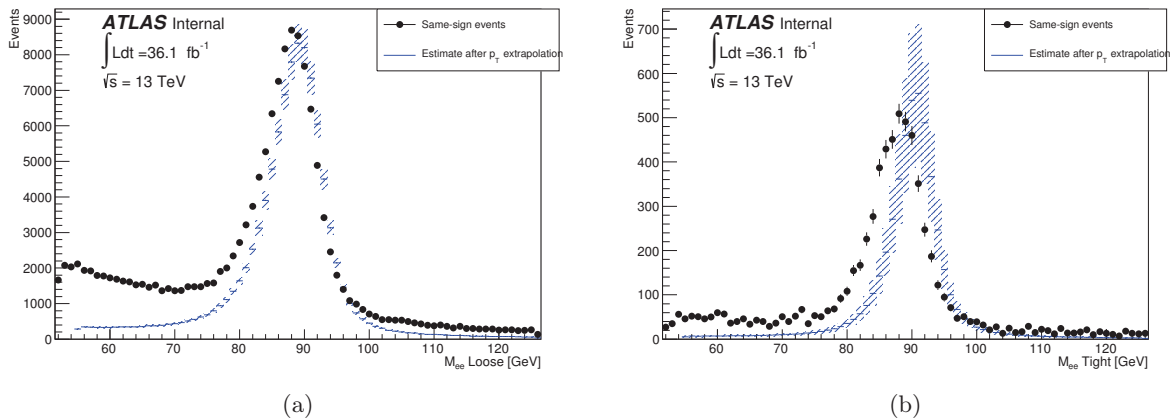


FIGURE 5.14 – Test de fermeture pour des événements avec deux électrons *antiTight* (a) et *Tight* (b). Les cercles noirs présentent la distribution des événements observés dans les données avec deux électrons de même charge tandis que les points bleus correspondent à la distribution des événements des données avec deux électrons de charges opposées repondérés par les taux obtenus par méthode du maximum de vraisemblance.

## 5.10 Conclusion

La méthode d'estimation du nombre d'événements avec une mauvaise estimation de la charge par maximum de vraisemblance a été présentée dans ce chapitre. Cette méthode a été utilisée lors de l'analyse des données partielles correspondant à 2015 et au début de l'année 2016 puis de nouveau pour l'analyse de l'ensemble des données 2015 et 2016. Une différence importante entre ces deux analyses est l'inclusion d'une troisième dimension dans l'estimation des taux de mauvaises reconstructions de la charge qui est la qualité de l'électron. Cette troisième dimension a permis d'utiliser plus d'événements de données. La méthode de vraisemblance permet de mesurer les taux de mauvaise identification de la charge des électrons avec une bonne précision et permet de s'affranchir de la simulation Monte-Carlo. In fine, elle est utilisée pour prédire un nombre d'événements avec mauvaise identification de la charge dans des régions de signal. Elle permet également la modélisation des distributions d'observables via la repondération des événements avec des leptons de charges opposées.

Un défaut de cette approche provient de la différence systématique entre l'énergie des électrons reconstruits après un effet trident et l'énergie des électrons n'ayant pas subi cet effet. En effet, le spectre en impulsion transverse des électrons provenant d'événements avec deux leptons de charges opposées est utilisé pour modéliser les spectres en impulsion transverse des électrons dans les événements avec deux leptons de même charge. Il en résulte un biais systématique de quelque GeV. L'impact de ce biais sur l'analyse, notamment 5% sur l'acceptance, a été estimé et est largement couvert par les incertitudes.

Des méthodes alternatives ont été développées dans ATLAS telle que la repondération de la simulation Monte-Carlo. Cette méthode consiste à utiliser les taux mesurés par vraisemblance non pour repondérer des événements de données mais pour corriger l'amplitude de l'effet simulé. Elle présente l'avantage d'utiliser la simulation de l'effet trident. Néanmoins elle nécessite l'utilisation d'une simulation complète des événements avec mauvaise identification de la charge dans les régions signal ce qui peut induire de fortes incertitudes statistiques et des incertitudes systématiques supplémentaires.

La première approche sera préférée pour la recherche de production associée  $t\bar{t}H$ .



# Interprétation statistique des résultats

Ce chapitre décrit une méthode statistique couramment utilisée au LHC pour l'interprétation statistique des données.

## 6.1 Modèle Statistique

Un outil communément utilisé pour définir le modèle statistique est la fonction de vraisemblance  $\mathcal{L}$  [68] (*likelihood* en Anglais) qui doit condenser tous les détails d'une mesure en une seule équation traduisant le degré de compatibilité entre les données et une hypothèse. Cette fonction de vraisemblance devra prendre en compte toutes les incertitudes statistiques et systématiques et peut être écrite sous la forme :

$$\mathcal{L}(\mu, \theta) = \prod_{i=1}^{n_{bins}} \prod_{j=1}^{n_{canaux}} \left[ \frac{(\mu s_{ij} + b_{ij})^{N_{ij}}}{N_{ij}!} e^{-\mu s_{ij} + b_{ij}} \right] \times \prod_{k=1}^{n_{nuis}} f_k(\theta_k). \quad (6.1)$$

Dans l'équation 6.1, le premier terme correspond à la statistique de Poisson avec  $N_{ij}$  le nombre d'événements observés dans les données,  $s_{ij}$  et  $b_{ij}$  respectivement les nombres d'événements de signal et de bruit de fond prédits. Ces quantités sont obtenus pour chaque bin  $i$ <sup>1</sup> et chaque région enrichie en signal (ou canal). Le paramètre  $\mu$  est un paramètre multiplicatif appliqué au signal appelé la *force du signal*. La force du signal est le paramètre d'intérêt (ou *POI* pour *Parameter of Interest*) de l'analyse.

Les termes  $\theta_k$  sont les paramètres de nuisance (ou *NP* pour *Nuisance Parameter*) de l'analyse et représentent les incertitudes systématiques dans la fonction de vraisemblance. Les fonctions  $f_k$  sont des fonctions correspondant à la mesure et à la contrainte des paramètres de nuisance. L'utilisation de régions de contrôle dédiées à la contrainte des paramètres de nuisances peut être réalisée. Dans ce cas ces fonctions pourront être décrites sous forme de la statistique de Poisson.

Pour estimer les valeurs vraies du paramètre d'intérêt et des paramètres de nuisance, un ajustement aux données sera fait avec pour but de maximiser la fonction de vraisemblance. Les valeurs obtenues seront des estimations des valeurs vraies et notées  $\hat{x}$  avec  $x$  soit le paramètre d'intérêt, soit un paramètre de nuisance.

## 6.2 Incertitudes et paramètres de nuisances

### 6.2.1 Incertitude statistique

Le nombre d'événements modélisé par simulation Monte-Carlo est limité. Cet incertitude statistique peut être ajouté dans le modèle statistique sous forme d'un paramètre de nuisance par bin par exemple dans l'outil *HistFactory* [69]. Ce choix a été fait pour limiter le nombre de paramètres de nuisance de ce type. En effet, un paramètre de nuisance pourrait être défini par bin et par processus mais cela pose des problèmes de temps pour la maximisation de la fonction de vraisemblance. En considérant un bin d'index  $b$ , la contribution au modèle statistique est le facteur :

$$Pois(n_b | \gamma_b \nu_b(\mu, \theta)) Pois(.,_b | \gamma_b \tau_b) \quad (6.2)$$

1. Un bin est un intervalle utilisé lors de la construction d'un histogramme

où  $n_b$  est le nombre d'événements observés dans le bin,  $\nu_b(\mu, \theta)$  est le nombre d'événements estimé à partir de la simulation Monte-Carlo,  $\gamma_b$  le paramètre de nuisance relatif au fait que le nombre d'événements réel puisse être différent de la simulation Monte-Carlo. Le second terme poissonien correspond au fait de traiter l'échantillon Monte-Carlo total dans ce bin comme une mesure supplémentaire pour contraindre ce paramètre de nuisance.

### 6.2.2 Incertitude systématiques

Dans le cas des incertitudes systématiques, la contribution au modèle statistique peut s'écrire de différentes manières. Deux cas sont décrits dans cette section, le cas d'une incertitude systématique n'agissant que sur la normalisation d'un processus et d'une incertitude systématique ne portant que sur la forme des distributions utilisées pour la maximisation. Certaines systématiques porteront sur les éléments qui sont factorisés et traités séparément.

### 6.2.3 Incertitude systématiques sur la normalisation

Les incertitudes portant sur la normalisation d'un processus sont décrits dans le modèle statistique pour un paramètre de nuisance  $\theta$  par :

$$Gaus(\tilde{\theta}|\theta, 1)\kappa^\theta \quad (6.3)$$

où le terme gaussien correspond à la contrainte supplémentaire sur le paramètre de nuisance et le terme  $\kappa^\theta$  correspond à la paramétrisation du changement des termes de la fonctions de vraisemblance lors des variations du paramètre de nuisance  $\theta$ . Le facteur  $\kappa$  est obtenu en faisant varier le paramètre de nuisance d'une déviation standard en positif et en négatif. Si l'incertitude n'est pas symétrique alors  $\kappa^+ \neq \kappa^-$ .

### 6.2.4 Incertitude systématiques sur la forme des distributions

Dans le cas des incertitudes systématiques portant sur la forme des distributions entrant dans le modèle statistique, il est aussi possible de séparer la composante du paramètre de nuisance dans le modèle en deux parties. La première correspondant à la contrainte supplémentaire sur le paramètre de nuisance est aussi décrite par une gaussienne. La seconde portant sur la réponse de la fonction de vraisemblance aux variations du paramètre est décrite par une fonction linéaire par morceau.

## 6.3 Test statistique et Profile Likelihood ratio

La valeur d'une fonction de vraisemblance est arbitraire à une constante multiplicative près et ce qui est utilisé le plus souvent comme test statistique est la comparaison des valeurs obtenues par cette fonction en deux points de l'espace des paramètres.

Par exemple on peut utiliser le test statistique appelé *log-likelihood ratio* (LLR) qui est défini tel que :

$$LLR = -2\ln \frac{\mathcal{L}(N|\alpha_1)}{\mathcal{L}(N|\alpha_2)}. \quad (6.4)$$

Ce test statistique permet la distinction entre deux hypothèses  $\alpha_1$  et  $\alpha_2$  pour un ensemble de données  $N$ . Le terme  $\alpha_1$  est appelé hypothèse nulle tandis que le terme  $\alpha_2$  est l'hypothèse alternative. Le lemme de Neyman-Pearson [70] permet de montrer que, si il n'y a pas de paramètre inconnu, ce rapport a la plus grande puissance statistique ce qui correspond à la plus grande capacité à rejeter l'hypothèse nulle si elle est fausse.



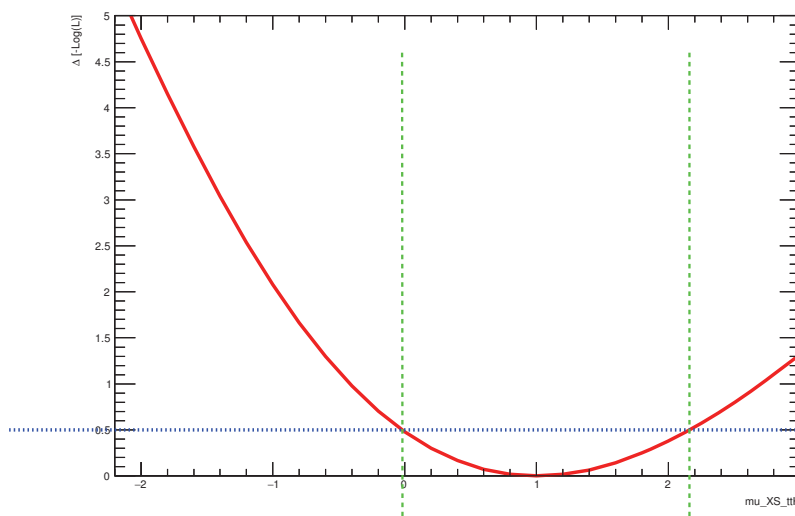


FIGURE 6.1 – Distribution de  $\ln(\lambda(\mu))$  en fonction de  $\mu$ . La valeur centrale de la force du signal obtenue par l’ajustement aux données est le minimum de la fonction et l’erreur sur cette valeur les abscisses correspondantes aux lignes pointillées vertes.

Un autre test statistique qui peut être employé est le *Profile Likelihood Ratio* (ou PLR)  $\lambda_\mu$  et qui est défini comme :

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta}_\mu)} \quad (6.5)$$

avec le numérateur correspondant à la fonction de vraisemblance obtenue en fixant la force du signal à une valeur précise et en ajustant les paramètres de nuisance et le dénominateur correspondant à un ajustement en laissant libre la force du signal et les paramètres de nuisance. Le numérateur correspond à la maximisation de la fonction de vraisemblance dans le cas où la valeur du paramètre d’intérêt est fixée et les valeurs des paramètres de nuisance sont variables.

Habituellement  $t_\mu = -2\ln(\lambda(\mu))$  est utilisé et la distribution de cette variable en fonction de la force du signal (valeur donnée dans la fonction de vraisemblance au numérateur) est représentée sur la Figure 6.1. Dans ce cas on cherchera à minimiser cette fonction  $t_\mu$ .

La fonction  $\lambda(\mu)$  a comme valeur maximale 1 qui est obtenue pour la valeur de  $\mu$  qui maximise la fonction de vraisemblance du dénominateur. Ce sera cette valeur de  $\mu$  qui sera citée comme force du signal mesurée. Cette force du signal mesurée peut être interprétée comme étant le rapport de la section efficace mesurée sur la section théorique du signal. Les erreurs sur la valeur de  $\mu$  mesurée vont être obtenues en prenant les abscisses correspondant à une ordonnée de 0.5. Par ailleurs une autre grandeur appelée *Signification statistique* peut être définie. Elle quantifiera l’écart entre l’hypothèse *bruit de fond seul* ( $\mu = 0$ ) et la valeur de  $\mu$  observée. Cette signification statistique est définie telle que le nombre de déviation standard entre les deux hypothèses et son unité est donc notée  $\sigma$ .

## 6.4 Données Asimov

Les analyses présentées dans ce document ont été optimisées en aveugle, c’est-à-dire sans regarder les données et en s’appuyant uniquement sur les prédictions du signal et du bruit de fond. Pour estimer la sensibilité de l’analyse, la maximisation de la fonction de vraisemblance est réalisée sur des pseudo-données, appelées données Asimov. Ces données sont définies tel que dans chaque bin le nombre d’événements de donnée est exactement le nombre d’événements prédits. Dans ce

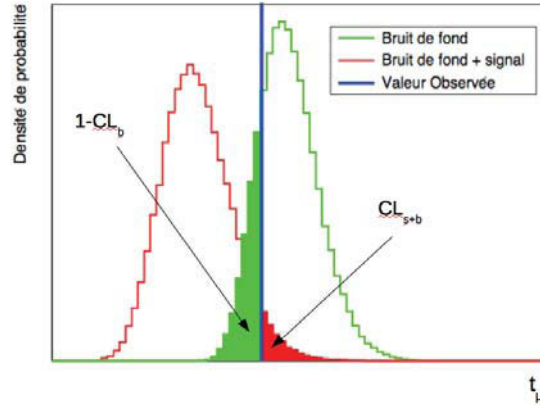


FIGURE 6.2 – Valeur observée du test statistique  $t_\mu$  et distributions de  $t_\mu$  pour les hypothèses H0 et H1.

cas le nombre d'événements peut ne pas être entier. La force du signal est par construction de 1 et l'erreur sur cette valeur donne une indication de la sensibilité de l'analyse.

## 6.5 Limite d'exclusion observées et attendues par méthode CLs

Dans le cas où il n'y a pas d'observation du processus recherché, une limite d'exclusion sur la force du signal va être réalisée. Pour cela des niveaux de confiance  $CL_b$  et  $CL_{s+b}$  sont définis et obtenus par calculs des  $p$ -value (Eq. 6.6 et 6.7) correspondant respectivement à l'hypothèse nulle, H0, pour une force du signal correspondant à la théorie ( $\mu = 1$ ) et l'hypothèse alternative, H1, pour l'hypothèse *bruit de fond seul* ( $\mu = 0$ ). La  $p$ -value représente la probabilité pour un test statistique d'avoir une valeur plus grande que celle observée.

$$p_0 = CL_{s+b} = P(t_\mu > t_{\mu Obs} | H0) = \int_{t_{\mu Obs}}^{\infty} P(t_\mu | H0) dt_\mu, \quad (6.6)$$

$$p_1 = CL_b = P(t_\mu > t_{\mu Obs} | H1) = \int_{t_{\mu Obs}}^{\infty} P(t_\mu | H1) dt_\mu \quad (6.7)$$

Les régions des distributions du test statistique  $t_\mu$  associées aux niveaux de confiance  $CL_b$  et  $CL_{s+b}$  sont représentées sur la Figure 6.2.

Ainsi plus la valeur de  $CL_{s+b}$  est petite et plus la compatibilité entre l'observation et l'hypothèse H0 sera faible. Dans le même temps une hypothèse H1 ne sera pas compatible avec les observations si la valeur de  $CL_b$  est proche de 1. Une compatibilité suffisamment faible suppose que le signal recherché est exclu.

Il est toutefois possible d'avoir une valeur de  $CL_{s+b}$  faible tout en ayant aussi une sensibilité de l'analyse faible si on observe une fluctuation basse du bruit de fond dans les données. Pour ne pas exclure le signal dans ce cas, le niveau de confiance  $CL_{s+b}$  est normalisé par celui de l'hypothèse *bruit de fond seul*. Le nouveau niveau de confiance est noté  $CL_s$  [71] :

$$CL_s = \frac{CL_{s+b}}{CL_b}. \quad (6.8)$$

L'hypothèse H1 sera alors exclue avec un niveau de confiance de 95% si  $CL_s = 0.05$ . Pour établir une limite supérieure d'exclusion sur la force du signal à 95% de niveau de confiance, on va faire varier celle-ci jusqu'à séparer suffisamment les distributions de  $t_\mu$  de chaque hypothèse et

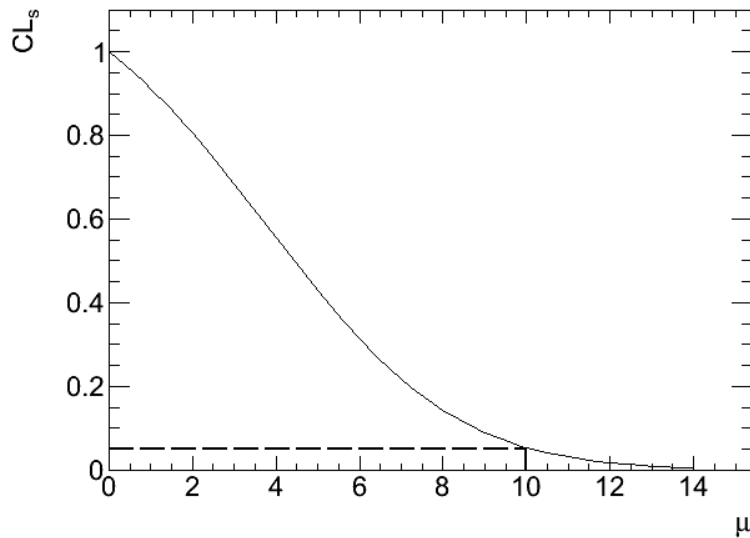


FIGURE 6.3 – Évolution du niveau de confiance  $CL_s$  en fonction de  $\mu$  et détermination de la valeur de la limite d'exclusion à 95%.

obtenir une valeur de  $CL_s = 0.05$ .

La distribution du niveau de confiance  $CL_s$  en fonction de  $\mu$  est construite, comme illustrée sur la Figure 6.3, à partir d'une valeur de  $t_\mu$  de référence telle que celle observée dans les données. Dans ce cas, la limite obtenue sera appelée limite d'exclusion observée.

Un limite d'exclusion attendue peut aussi être définie en prenant comme valeur de référence de  $t_\mu$ , la valeur médiane de  $t_\mu$  associée à l'hypothèse *bruit de fond seul*. Pour la limite attendue, cinq valeurs sont présentées correspondant à cinq quantiles de la distribution  $t_\mu$  de l'hypothèse H1. Ces cinq quantiles sont la médiane,  $\pm 1$ ,  $\pm 2$  déviations standards de cette distribution.

## 6.6 Représentation de l'impact des paramètres de nuisance sur la valeur de la force du signal

Pour évaluer l'impact des paramètres de nuisance sur le paramètre d'intérêt, une représentation couramment utilisée est de présenter les paramètres de nuisance rangées par ordre d'importance telle que sur la Figure 6.4. Cette figure correspond à l'analyse présentée dans le Chapitre ??.

Plusieurs éléments apparaissent sur cette figure. Tout d'abord les contraintes après ajustements aux données d'un paramètres de nuisances sont représentées par les barres noires. Si ces barres noires ont une longueur inférieure à 2, alors le paramètre de nuisance est contraint lors de l'ajustement aux données. Si la valeur centrale est différente de 0 alors cela signifie que la valeur nominale du bruit de fond associé est modifié lors de l'ajustement aux données.

Ensuite l'impact sur la force du signald'un paramètre de nuisance est représenté par les barres pleines (vides) avant (après) l'ajustement aux données. Cet effet est estimé en obtenant la valeur de la force du signal après avoir fixé la norme du bruit de fond à la valeur nominale  $\pm 1\sigma$  de cette incertitude et en comparant ces valeurs de la force du signalavec la valeur nominale. Dans le même temps l'impact des autres paramètres de nuisance est fixé.

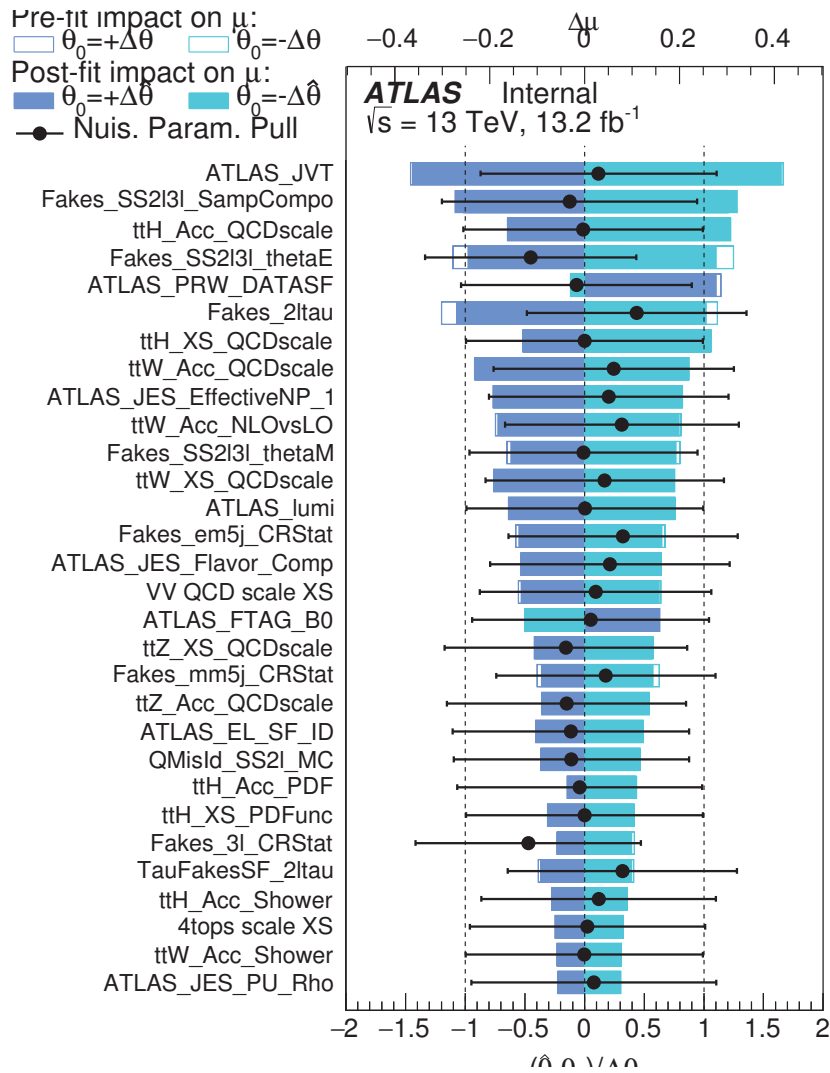


FIGURE 6.4 – Variation de la force du signal pour chaque paramètre de nuisance. Les paramètres sont classés par impact sur la force du signal. Les traits noirs représentent la contrainte sur les paramètres de nuisances opérées lors de l’ajustement aux données.

# Etude du couplage du boson de Higgs au quark *top* dans l'expérience ATLAS

---

## 7.1 Introduction

La recherche d'un état final avec deux quarks *top* et un boson de Higgs est présentée ici par le biais de la recherche d'un état final multilepton qui correspond à des désintégrations du boson de Higgs en une paire de bosons  $W^\pm W^\mp$ ,  $ZZ$  ou en une paire de leptons  $\tau\tau$ . Parmi les états finals leptoniques possibles<sup>1</sup>, le cas d'un état final avec deux leptons de même charge électrique sans  $\tau$  hadronique sera plus spécifiquement détaillé dans la suite du chapitre. Un diagramme correspondant à un tel état final est présenté sur la Figure 7.1. Outre cet état final, les canaux suivants sont aussi considérés : deux leptons de même charge électrique et un tau hadronique ( $2\ell SS1\tau$ ), trois leptons ( $3\ell$ ) et quatre leptons ( $4\ell$ ). L'analyse utilise les données enregistrées par le détecteur ATLAS auprès du LHC en 2015 et au début de 2016 pour une luminosité intégrée de  $13.2 \text{ fb}^{-1}$  et une énergie dans le centre de masse de 13 TeV.

Dans un premier temps, des précisions sont données sur les objets physiques utilisés dans la Section 7.2. Ensuite la Section 7.3 détaille les échantillons de données et de simulation utilisées pour cette analyse. La Section 7.5 décrit l'estimation de certaines catégories de bruit de fond à partir des données. La définition des régions enrichies en signal est décrite dans la Section 7.4. La Section 7.7 décrit la validation de la modélisation des différents bruits de fond dans une région de contrôle. Enfin les résultats seront décrits dans la Section 7.8. J'ai participé à l'estimation des bruits de fonds issus du processus  $t\bar{t}$  (leptons secondaires et événements avec des électrons dont la charge électrique est mal reconstruite) pour l'état final avec deux leptons de même charge et j'ai réalisé l'ajustement aux données pour l'ensemble des canaux multileptons.

## 7.2 Objets physiques et présélection

### 7.2.1 Objets physiques

Les objets physiques utilisés dans cette analyse correspondent à ceux décrits dans le Chapitre 4. Pour catégoriser l'état final étudié, le nombre de leptons est compté. Ce comptage est effectué sur des leptons reconstruits selon des critères lâches. Ainsi un électron est considéré si il passe le critère d'identification `LooseAndBLayer` et le critère d'isolation `Loose`. Les mêmes critères seront utilisés pour définir et compter les muons.

Pour réduire la proportion de leptons mal reconstruits, des critères supplémentaires sont ajoutés aux précédents selon le canal étudié. Dans le cas de deux leptons de même charge électrique, le critère d'isolation utilisé pour la définition d'un électron est `FixedCutTight`, tandis que celui utilisé pour les muons est `FixedCutTightTrackOnly`. En outre le critère `TightLH` est appliqué pour l'identification des électrons. Par ailleurs, seuls les événements dont les électrons sont reconstruits dans la région

$|\eta| < 1.37$  sont conservés pour réduire la proportion d'événements avec un électron dont la charge électrique est mal reconstruite. L'impulsion transverse des deux leptons devra être supérieure à

---

1. Par la suite le terme *canal* sera aussi utilisé

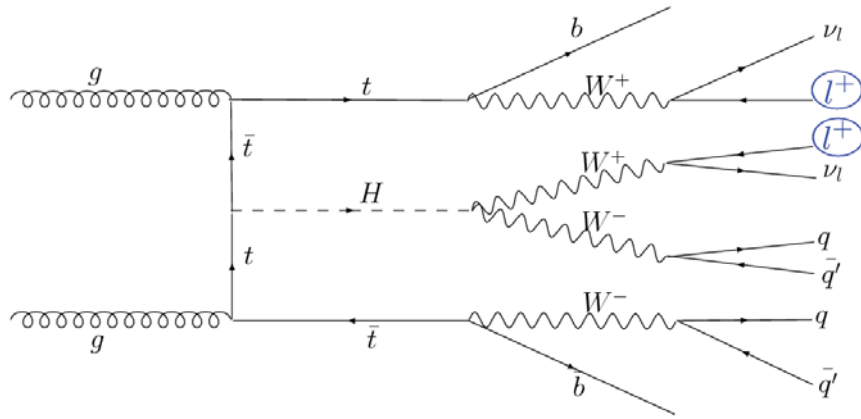


FIGURE 7.1 – Diagramme de production d'une paire de quarks *top* associés à un boson de Higgs puis de la désintégration de ces particules conduisant à un état final avec deux leptons de même charge électrique.

25 GeV, cette sélection ayant pour but de diminuer la proportion de leptons secondaires dans l'état final.

L'ensemble des critères d'isolation et d'identification des leptons utilisés pour les différentes signatures sont donnés dans la Table 7.1. Pour les canaux  $2\ell SS0\tau$ ,  $2\ell SS1\tau$  et  $4\ell$  les leptons sont ordonnés par impulsion transverse décroissante, le lepton 0 étant celui qui a la plus grande impulsion transverse. Pour le canal  $3\ell$ , le lepton 0 correspond au lepton dont la charge électrique est opposée à celle des deux autres. Ainsi dans le cas d'un événement produisant trois leptons dans l'état final, le lepton 0 est le lepton dont la probabilité d'être un lepton non prompt est la plus faible. Ensuite les deux leptons sont rangés par impulsion transverse croissante.

	$2\ell 0\tau$	$3\ell$	$2\ell 1\tau$	$4\ell$
electron		TightLH		Gradient
muon		FixedCutTight		Gradient
$\ell$ (e ou $\mu$ )	$p_T > 25\text{GeV}$	$p_{T\ 1(2)} > 20\text{GeV}, p_{T\ 0} > 10\text{GeV}$	$p_{T\ 0} > 25\text{GeV}, p_{T\ 1} > 25\text{GeV}$	$p_T > 10\text{GeV}$
tau			$Q_{lep} = -Q_\tau$	

TABLE 7.1 – Définition des leptons utilisés dans les différents canaux. La sélection utilisée pour le comptage des leptons est appliquée au préalable.

### 7.2.2 Exclusion du double comptage des objets

Lors de la reconstruction des objets, deux d'entre eux peuvent être reconstruits très près l'un de l'autre et il est probable que seul l'un des deux objets soit considéré. Ainsi un ensemble de critères est appliqué selon les cas pour savoir quels objets sont conservés. Ils sont présentés dans la Table 7.2. Ils sont définis à partir de la distance entre les deux objets reconstruits ou sur les traces partagées par les objets.

### 7.2.3 Présélection

Une présélection est effectuée avant la définition d'une région enrichie en signal. Les événements ne seront utilisés que dans le cas où ils contiennent au moins 2 jets dont un étiqueté *b*. En effet

Conservé	Enlevé	Distance ( $\Delta R$ ) ou trace
electron	tau	0.2
muon	tau	0.2
electron	CaloTagged muon	Partage une trace
muon	electron	Partage une trace
electron	jet	0.2
jet	electron	0.4
muon	jet	(0.2 ou ghost-matched au muon) et (nb de trace du jet $\leq 2$ )
jet	muon	0.4
tau	jet	0.2

TABLE 7.2 – Critères utilisés pour la sélection des objets, dans le cas de proximité topologique.

quelque soit le canal étudié, deux jets étiquetés  $b$  devraient être présents idéalement dans l'état final. En prenant en compte le fait que l'un des jets issu d'un quark  $b$  puisse ne pas être étiqueté  $b$ , les événements  $t\bar{t}H$  devraient avoir dans la majorité des cas au moins deux jets dont un étiqueté  $b$ . En outre cette préselection réduit grandement les bruits de fonds attendus tels que les événements  $V$ +jets et dibosons.

## 7.3 Échantillons de données, de simulation et description des bruits de fonds irréductibles

### 7.3.1 Données utilisées

L'analyse présentée utilise les données enregistrées par le détecteur ATLAS auprès du LHC en 2015 et pendant la première partie de 2016 soit une luminosité intégrée totale de  $13.2 \text{ fb}^{-1}$  pour une énergie dans le centre de masse de 13 TeV. Un ensemble de systèmes de déclenchement est utilisé pour ne conserver dans un premier temps que les événements possédant au moins un lepton. Ces critères sont différents selon l'année de la prise de données ou l'utilisation de simulations Monte-Carlo ou des données :

- $HLL\_e24\_lhmedium\_L1EM20VH$  ( $HLL\_e24\_lhmedium\_L1EM18VH$ ) OU  $HLL\_e60\_lhmedium$  OU  $HLL\_e120\_lhloose$  OU  $HLL\_mu20\_iloose\_L1MU15$  OU  $HLL\_mu50$  pour les données (resp. simulation MC) en 2015
- $HLL\_e24\_lhtight\_nod0\_ivarloose$  OU  $HLL\_e60\_lhmedium\_nod0$  OU  $HLL\_e140\_lhloose\_nod0$  OU  $HLL\_mu24\_ivarmedium$  OU  $HLL\_mu50$  en 2016

### 7.3.2 Simulation : Échantillons de signal

Les événements de signal décrivant la production d'une paire de quarks  $top$  et d'un boson de Higgs ( $t\bar{t}H$ ) et leurs désintégrations sont modélisés avec les générateurs Pythia 8 [72] et Mc@NLO [54]. Une simulation complète du détecteur est utilisée. La simulation d'événements  $t\bar{t}H$  est séparée en trois échantillonnages selon le nombre de leptons provenant des désintégrations des deux quarks  $top$ . Les sections efficaces correspondant à chacun de ces échantillons sont respectivement de  $53.4 \text{ fb}$ ,  $222.6 \text{ fb}$  et  $230.8 \text{ fb}$  pour des désintégrations de la paire  $t\bar{t}$  avec deux leptons, un lepton et aucun lepton dans l'état final.

### 7.3.3 Définition des bruits de fond et échantillons utilisés

Les processus qui peuvent produire des états finals avec deux leptons de même charge électrique se regroupent en deux catégories :

- Les processus physiques qui conduisent à deux leptons prompts de même charge électrique. On parlera par la suite de bruit de fond **irréductible**. Ils sont obtenus par simulation Monte-Carlo. Ces bruits de fond sont décrits dans la suite de cette section.
- Les processus physiques qui ne donnent pas deux leptons prompts de même charge électrique mais sont reconstruits comme tels. Ces bruits de fond seront qualifiés de **réductibles** ou **instrumentaux** et ils sont estimés à partir des données. Dans le cas de l'analyse présentée dans ce document et pour un état final avec deux leptons de même charge électrique, ces bruits de fond proviennent majoritairement d'événements  $t\bar{t}$  avec une désintégration semi-leptonique (dans le cas des leptons secondaires) ou dileptonique (dans le cas des événements avec une mauvaise reconstruction de la charge électrique).

Le plus important des bruits de fonds irréductibles provient de la **production associée d'une paire  $t\bar{t}$  et de bosons de jauge ( $t\bar{t}V$ )**. Dans le cas de deux leptons de même charge électrique, le boson de jauge en question sera le boson W. Dans le cas d'un état final avec 3 ou 4 leptons, la production d'événements  $t\bar{t}Z$  sera dominante. De la même façon que pour le signal, les échantillons sont générés à l'aide de Pythia 8 et MC@NLO. Les sections efficaces correspondantes à ces deux bruits de fond sont respectivement de 548.3 *fb* et 124.3 *fb* pour les processus  $t\bar{t}W$  et  $t\bar{t}Z$ . La séparation entre ces processus et le processus  $t\bar{t}H$  est difficile et la principale différence est un nombre de jet potentiellement plus faible pour ces processus. Pour ce bruit de fond une simulation MC suivant le modèle standard est utilisée malgré l'absence de mesure des sections efficaces de ces processus dans les données.

Une deuxième catégorie de bruit de fond irréductible provient de la **production de deux bosons de jauge**. Ces différents processus sont décrits par SHERPA [52]. La section efficace de production est très importante, de l'ordre de 100 *pb*. Ces processus seront toutefois minoritaires dans la région enrichie en signal car situés à plus basse multiplicité en nombre de jets et jets étiquetés *b*.

D'autres bruits de fond irréductibles existent dans le modèle standard tels que les processus  $WtZ$ , **4*tops***,  $t\bar{t}W^+W^-$  ou  $tZ$ . Ces bruits de fond représentent un petit nombre d'événements pour les régions considérées et cela même si leur multiplicité en nombre de jets et leptons peut être très proche de celle du signal<sup>2</sup>. En effet, leurs sections efficaces de production sont très faibles. Ces bruits de fond sont regroupés dans la suite sous la dénomination *Autres*.

Enfin le processus de **production d'une paire  $t\bar{t}$**  peut aussi être modélisé par simulation MC grâce au générateur POWHEG [55]. Cette simulation va être utilisée pour définir une région enrichie en signal à partir de la présélection (voir Section 7.2.3). Toutefois cette simulation n'est pas optimale et d'autres méthodes sont utilisées pour estimer les bruits de fond liés à ce processus et seront décrites dans la Section 7.5. Le bruit de fond  $t\bar{t}$  est caractérisé par un nombre de jets plus faible que le signal (entre 3 et 4 jets). En outre dans le cas d'une désintégration semi-leptonique de la paire  $t\bar{t}$  et de la présence d'un lepton non-prompt, un seul neutrino est présent dans l'état final contre deux pour le processus  $t\bar{t}H$  ce qui devrait conduire à une énergie transverse manquante plus faible.

---

2. Ainsi le processus  $t\bar{t}W^+W^-$  a exactement le même état final que le signal dans le canal  $2\ell SS0\tau$  mais une section efficace de l'ordre de 8 *fb*.



## 7.4 Définition de la région enrichie en signal

Pour définir la région enrichie en signal pour le canal  $2\ell SS0\tau$ , des critères supplémentaires peuvent être ajoutés sur le nombre de jets, le nombre de jets étiquetés  $b$  ou l'angle  $\eta$  des leptons. La stratégie d'optimisation est de définir la région qui maximise la sensibilité pour une analyse de comptage où le nombre d'événements de données dans cette région sera compté et comparé aux prédictions. Cette optimisation a été faite pour cette analyse au Run 1 du LHC pour une luminosité intégrée de  $20.2 \text{ fb}^{-1}$  et une énergie dans le centre de masse de 8 TeV. Les distributions des variables utilisées pour définir la région de signal sont présentées sur la Figure 7.2.

Lors de l'analyse faite au Run 1, la sélection finale consistait à demander 4 jets dont 1 étiqueté  $b$ . Pour cette analyse une optimisation supplémentaire est effectuée qui consiste à enlever les événements à 4 jets de la région enrichie en signal. En effet cette région à 4 jets contient peu d'événements de signal et peut être utilisée pour améliorer l'estimation du nombre d'événements avec des leptons secondaires. La Table 7.3 présente le nombres d'événements et la signification statistique pour des régions définies selon la saveur des leptons et le nombre de jets.

	$t\bar{t}H$	Diboson	$t\bar{t}W$	$t\bar{t}ll$	$t\bar{t}$	Bkg Total	$\sigma$ Cowan	S/B	$S\sqrt{B}$	$\sigma$ RooZ
$ee$ (5j, 1b)	$0.58 \pm 0.08$	$0.22 \pm 0.11$	$0.77 \pm 0.02$	$0.29 \pm 0.01$	$2.73 \pm 0.64$	$4.59 \pm 0.66$	0.25	0.13	0.27	0.01
$e\mu$ (5j, 1b)	$1.43 \pm 0.12$	$0.73 \pm 0.18$	$2.51 \pm 0.04$	$0.70 \pm 0.01$	$4.95 \pm 0.73$	$10.32 \pm 0.76$	0.41	0.14	0.44	0.16
$\mu\mu$ (5j, 1b)	$0.94 \pm 0.10$	$0.38 \pm 0.14$	$1.91 \pm 0.04$	$0.54 \pm 0.01$	$0.98 \pm 0.35$	$4.74 \pm 0.39$	0.42	0.20	0.43	0.20
$ee$ (4j, 1b)	$0.22 \pm 0.05$	$0.52 \pm 0.18$	$0.76 \pm 0.02$	$0.17 \pm 0.01$	$3.94 \pm 0.63$	$5.61 \pm 0.66$	0.08	0.04	0.09	-0.15
$e\mu$ (4j, 1b)	$0.62 \pm 0.07$	$0.46 \pm 0.13$	$2.23 \pm 0.04$	$0.46 \pm 0.01$	$7.24 \pm 0.90$	$11.02 \pm 0.91$	0.17	0.06	0.19	-0.06
$\mu\mu$ (4j, 1b)	$0.38 \pm 0.05$	$0.41 \pm 0.19$	$1.70 \pm 0.03$	$0.34 \pm 0.01$	$1.71 \pm 0.46$	$4.53 \pm 0.50$	0.16	0.08	0.18	-0.07

TABLE 7.3 – Nombre d'événements attendus dans le canal  $2\ell SS0\tau$ . La signification statistique est donnée pour des régions définies selon la saveur des leptons et le nombre de jets. Tous les bruits de fond sont modélisés par simulation MC et une incertitude de 50% est appliquée sur le bruit de fond diboson.

La sélection utilisée pour obtenir les résultats décrits dans la Section 7.8 est d'au moins 5 jets et 1 jet étiqueté  $b$ . Des optimisations similaires sont effectuées pour les autres canaux. Outre des sélections sur le nombre de jets, pour les canaux autres que  $2\ell SS0\tau$ , les événements avec deux leptons venant de désintégrations du boson Z sont aussi enlevés. L'ensemble des sélections dans les différents canaux est reporté dans la Table 7.4.

## 7.5 Estimation des bruits de fonds instrumentaux

Les bruits de fond réductibles présentent des états finals qui n'ont pas deux leptons de même charge électrique mais qui sont reconstruits comme tels. Deux catégories existent :

- deux leptons de charges électriques opposées sont présents dans l'état final mais l'un de ces leptons voit sa charge mal reconstruite. Une description plus complète de ce processus et des méthodes d'estimation est présentée au Chapitre 5. Ce bruit de fond est estimé par la même méthode pour le canal  $2\ell SS0\tau$  et le canal  $2\ell SS1\tau$ .
- un lepton reconstruit et sélectionné provient d'un processus secondaire (lepton secondaire) ou est un autre objet (jet ou photon). De tels leptons proviennent majoritairement de désintégrations de hadrons beaux. Ils peuvent aussi être produits lors de la production d'un trident d'électrons. La méthode utilisée pour estimer ce bruit de fond à partir des données est présentée dans la Section 7.5.2. Ce bruit de fond est estimé par la même méthode pour les canaux  $2\ell SS0\tau$  et  $3\ell$  et par d'autres méthodes pour les canaux  $2\ell SS1\tau$  et  $4\ell$ . Seule la méthode utilisée pour le canal  $2\ell SS0\tau$  sera décrite.

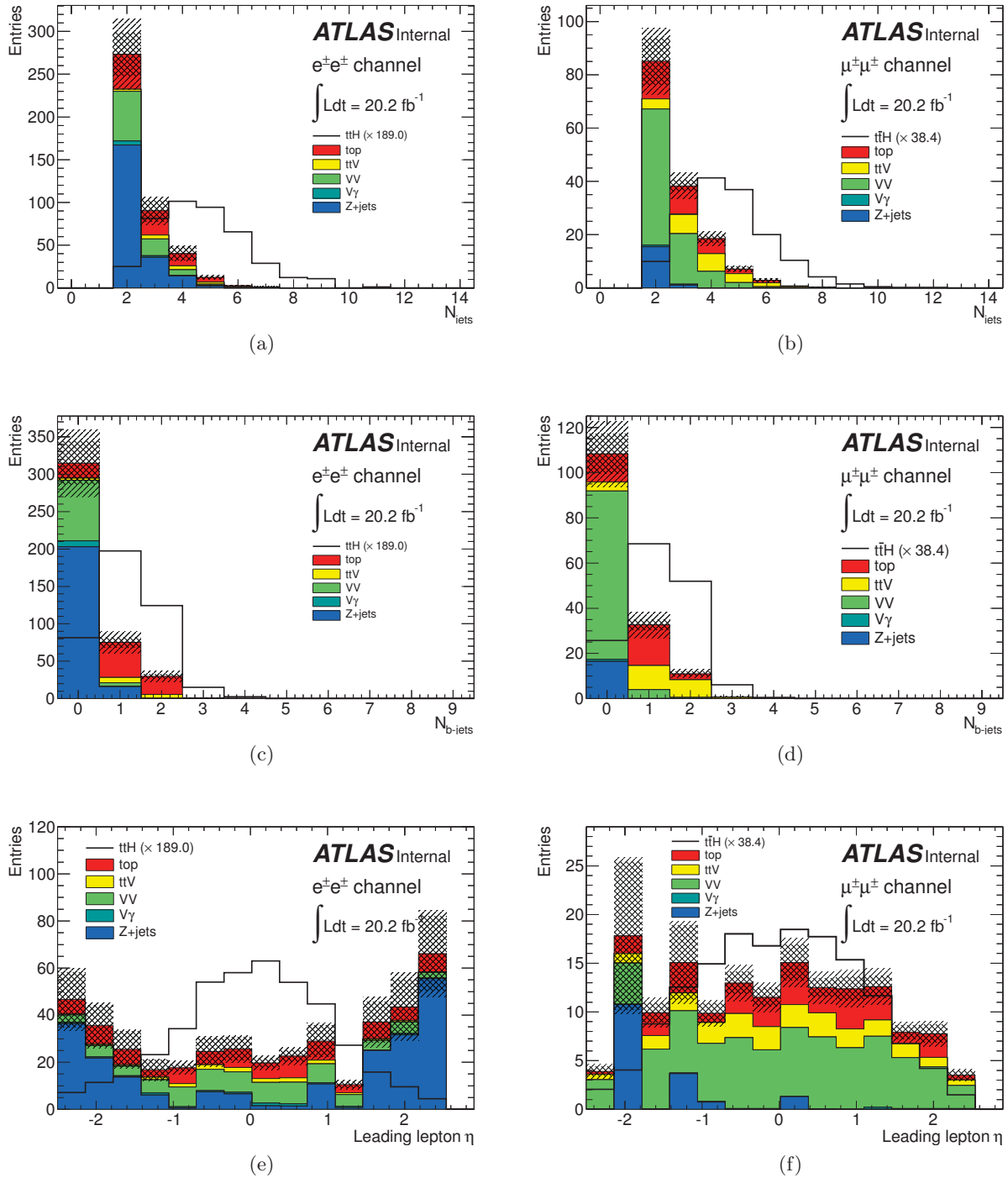


FIGURE 7.2 – Distribution pour le signal et les différents bruits de fond du nombre de jets pour le canal  $ee$  (a) et  $\mu\mu$  (b), du nombre de jets étiquetés  $b$  pour le canal  $ee$  (c) et  $\mu\mu$  (d) et des angles  $\eta$  pour le lepton de plus grande impulsion transverse pour le canal  $ee$  (e) et  $\mu\mu$  (f).

Région	Leptons	Taus Hadroniques	Jets	B-Tags	Saveur des leptons
$2\ell SS0\tau$	$\Sigma Q_{lep} = \pm 2$	$== 0$	$\geq 5$	$\geq 1$	ee
	$p_T^0 > 25 \text{ GeV}$				$e\mu$
	$p_T^1 > 25 \text{ GeV}$ $ \eta^{elec}  < 1.37$				$\mu\mu$
$3\ell$	$\Sigma Q_{lep} = \pm 1$	-	$N_{jet} \geq 4 \ \& \ N_b \geq 1$ or $N_{jet} == 3 \ \& \ N_b \geq 2$		-
	lep1 & lep2 : Tight, $p_T > 20 \text{ GeV}$				
	lep0 : Loose, $p_T > 10 \text{ GeV}$				
	$ M_{ll,OS-SF} - 91.2 \text{ GeV}  > 10 \text{ GeV}$				
	$M_{ll,OS-SF} > 12 \text{ GeV}$ $ M_{lll} - 91.2 \text{ GeV}  > 10 \text{ GeV}$				
$4\ell$	$\Sigma Q_{lep} = 0$	-	$\geq 2$	$\geq 1$	-
	$p_T > 10 \text{ GeV}$				
	Loose def + Gradient Iso				
	$ M_{ll,OS-SF} - 91.2 \text{ GeV}  > 10 \text{ GeV}$				
	$M_{ll,OS-SF} > 12 \text{ GeV}$ $M_{4l}$ entre [100, 350] GeV				
	$M_{4l}$ veto $\pm 5 \text{ GeV}$ autour $M_H = 125.0 \text{ GeV}$				
$2\ell SS1\tau$	$\Sigma Q_{lep} = \pm 2$	$== 1$	$\geq 4$	$\geq 1$	-
	$p_{T,lep}^{lead} > 25 \text{ GeV}$				
	$p_{T,lep}^{sub} > 15 \text{ GeV}$				
	$ M_{ee} - 91.2 \text{ GeV}  > 10 \text{ GeV}$				
	$Q_\tau = -Q_{lep}$				

TABLE 7.4 – Résumé des sélections effectuées dans les régions de signal.

### 7.5.1 Estimation du nombre d'événements avec des électrons dont la charge électrique est mal reconstruite

La méthode utilisée est présentée au Chapitre 5. Les taux de mauvaise identification de la charge des leptons est estimés dans des régions en  $\eta$  et  $p_T$ . L'estimation du nombre d'événements avec des électrons dont la charge est mal reconstruite est réalisée pour les régions de contrôles correspondant au processus  $t\bar{t}W$  et pour les régions de signal :

- $N_{Q_{MisId}}^{ee}(5j) = 6.9 \pm 1.3$ .
- $N_{Q_{MisId}}^{e\mu}(5j) = 7.1 \pm 1.7$ .
- $N_{Q_{MisId}}^{2\ell SS1\tau}(4j) = 0.24 \pm 0.03$ .

L'incertitude sur l'estimation dans les différentes régions est de l'ordre de 15%.

### 7.5.2 Estimation du nombre d'événements avec des leptons secondaires

La méthode utilisée pour extraire le nombre d'événements avec un lepton secondaire est appelée *Fake Factor Method*. Elle permet d'extraire le nombre d'événements dans une région de signal sans modélisation des variables. Cette méthode ne peut donc être utilisée que pour une analyse qui consistera à compter le nombre d'événements dans notre région enrichie en signal.

L'hypothèse de base de cette méthode est qu'une grandeur appelée *fake factor* et notée  $\theta$  est stable en fonction du nombre de jets. Pour extraire ce facteur deux régions sont définies :

- Une première région qui comprend deux leptons qui passent tous les critères d'isolation et d'identification décrits précédemment. Ces leptons sont qualifiés de *Tight* et notés  $\ell$ .
- Une deuxième région comprenant deux leptons dont l'un ne passe pas l'un des critères d'isolation ou de d'identification. Ce lepton est qualifié d'*AntiTight* et noté  $\ell$ .

Le facteur  $\theta$  est alors défini comme :

$$\theta_\ell = \frac{N_{\ell\ell}}{N_{\ell\bar{\ell}}} \quad (7.1)$$

où  $\ell$  est soit un électron soit un muon. Seuls des événements avec deux leptons de même saveur sont utilisés.

$N_{\ell\ell}$  et  $N_{\ell\ell}$  sont les nombres d'événements observés dans les données ( $N^D$ ) après soustraction des bruits de fond irréductibles ( $N^{Irred}$ ) et des événements dont l'un des leptons voit sa charge mal reconstruite ( $N^{QMisId}$  et  $N^{QMisId MC}$  pour les électrons). Le facteur  $\theta$  est extrait dans une région avec une faible multiplicité en nombre de jets, c'est-à-dire dans cette analyse pour des événements avec entre 2 et 4 jets. Les facteurs  $\theta$  pour électrons et muons peuvent donc être écrits sous la forme :

$$\theta_e = \frac{N_{ee}(\leq 4jets)}{N_{e\ell}} = \frac{N_{ee}^D - N_{ee}^{Irred} - N_{ee}^{QMisId}}{N_{e\ell}^D - N_{e\ell}^{Irred} - N_{e\ell}^{QMisId MC}} \quad (7.2)$$

$$\theta_\mu = \frac{N_{\mu\mu}(\leq 4jets)}{N_{\mu\ell}} = \frac{N_{\mu\mu}^D - N_{\mu\mu}^{Irred}}{N_{\mu\ell}^D - N_{\mu\ell}^{Irred}}. \quad (7.3)$$

La contribution des événements ayant deux leptons secondaires reconstruits comme prompts est considérée comme négligeable.

En faisant l'hypothèse que le facteur  $\theta$  est indépendant de la multiplicité en jets, on peut écrire pour la région enrichie en signal :

$$N_{\ell\ell}(> 5jets) = N_{\ell\ell}(> 5jets) \times \theta_\ell. \quad (7.4)$$

En suivant l'équation 7.4, le nombre d'événements contenant au moins un lepton secondaire reconstruit comme lepton prompt est extrait pour les régions avec deux leptons de même saveur. L'estimation du nombre d'événements dans le cas d'un électron et d'un muon est donnée par :

$$N_{e\mu}(> 5jets) = N_{e\mu}(> 5jets) \times \theta_\mu + N_{\mu\ell}(> 5jets) \times \theta_e \quad (7.5)$$

Par ailleurs les incertitudes systématiques suivantes sont associées à ces estimations :

- **Stat** : l'incertitude statistique due à la taille de la région à haute multiplicité en nombre de jet et avec un lepton *AntiTight*.
- $\theta_\ell$  **stat** : l'incertitude statistique due à la taille de la région à basse multiplicité en nombre de jets et avec un lepton *AntiTight*. Cette incertitude sera séparée pour les électrons et les muons.
- $\theta_\ell$  **syst** : l'incertitude liée à la validation de l'extrapolation des facteurs  $\theta$  des régions à basse multiplicité du nombre de jet à haute multiplicité avec des événements MC  $t\bar{t}$ .
- **QMisID** : l'incertitude liée à la soustraction des événements avec un électron dont la charge est mal reconstruite. Cette incertitude est anti-corrélée avec l'incertitude appliquée sur l'estimation du nombre d'événements avec un électron de charge mal reconstruite dans la région enrichie en signal.
- **Composition de la région à basse multiplicité (Comp.)** : estimée dans le cas où, dans la région à basse multiplicité, des processus autres que ceux présents à haute multiplicité (par exemple  $W + jets$  vis-à-vis de  $t\bar{t}$ ) contribuent au nombre d'événements avec leptons secondaires. La Figure 7.3 présentent le nombre d'événements avec des leptons non-prompts attendus en variant la définition de la région à basse multiplicité en nombre de jets pour changer la composition de cette région.

Un test de fermeture a été réalisé pour vérifier la validité de l'estimation. Les événements  $e\mu$  avec une faible multiplicité en jets sont choisis car ne participant pas à l'estimation des facteurs  $\theta$ . Les nombres d'événements obtenus sont rapportés dans la Table 7.5. Une compatibilité entre les prédictions et les données est observée.

Le nombre d'événements pour les différents saveurs de leptons est présenté dans la Table 7.6.

Dans le cas du canal  $3\ell$  les facteurs  $\theta$  dérivés dans le canal  $2\ell SS0\tau$  sont appliqués sur des événements contenant trois leptons dont l'un ne passe pas la sélection *Tight*. De même que pour le

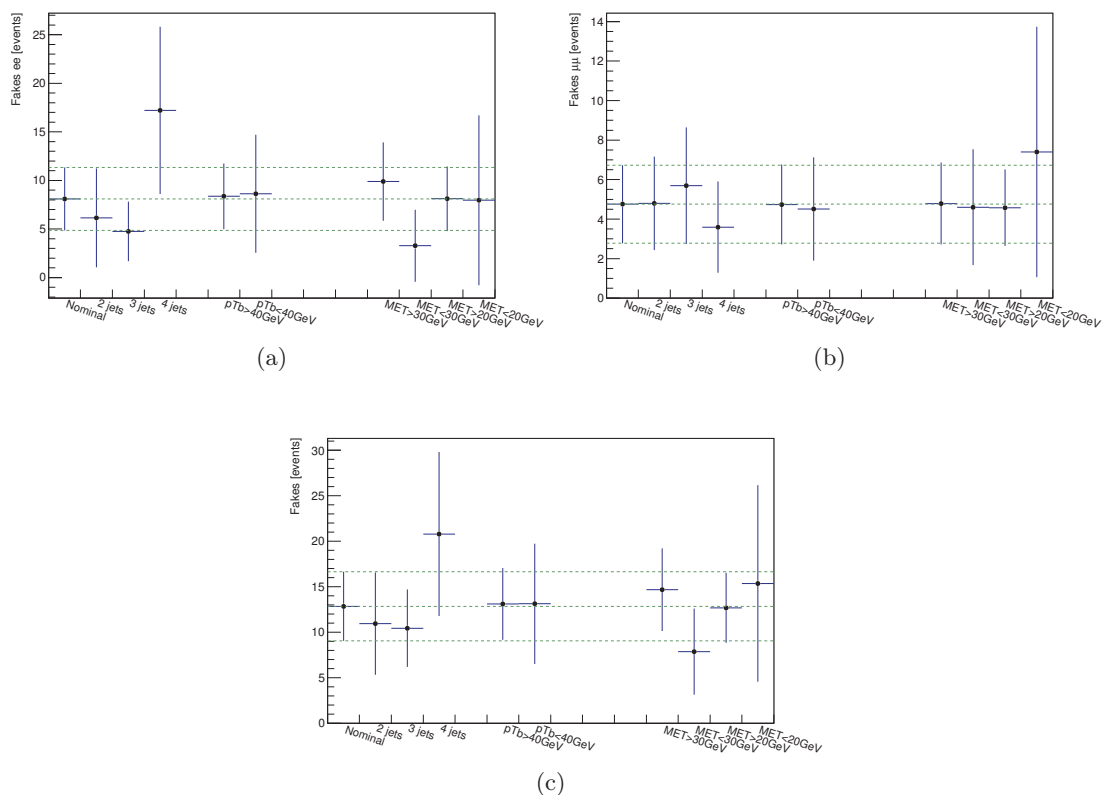


FIGURE 7.3 – Stabilité de la prédiction du nombre d'événements avec des leptons non-prompt dans le canal  $ee$  (en haut à gauche), le canal  $\mu\mu$  (en haut à droite) et le total (en bas) obtenue en variant la définition de la région à basse multiplicité en nombre de jets.

Non-prompt	Prédiction totale	Données	Données/Prediction
$91.84 \pm 22.22$ (stat)	$186.36 \pm 29.14$	206	$1.11 \pm 0.08$ (données) $\pm 0.17$ (non-prompt, QMisID)

TABLE 7.5 – Comparaison de la prédiction du nombre d'événements avec des leptons non-prompt dans le canal  $e\mu$  (2-4 jets) aux données.

Saveurs	Nombre d'evt	$\pm$ Stat	$\pm \theta_e$ stat	$\pm \theta_e$ syst	$\pm \theta_\mu$ stat	$\pm \theta_{mu}$ syst	$\pm$ QMisID	$\pm$ Comp
$ee =$	12.07	$\pm 2.40$	$\pm 2.92$	$\pm 2.99$	$\pm 0$	$\pm 0$	$\mp 2.29$	$\pm 2.29$
$\mu\mu =$	8.72	$\pm 2.40$	$\pm 0$	$\pm 0$	$\pm 1.62$	$\pm 0.85$	$\pm 0$	$\pm 1.66$
$e\mu =$	12.39	$\pm 2.60$	$\pm 2.26$	$\pm 2.32$	$\pm 0.57$	$\pm 0.30$	$\mp 1.77$	$\pm 1.87$

TABLE 7.6 – Estimation du nombre d'événements avec un lepton secondaire dans les régions à haute multiplicité en jet et deux leptons *Tight*.

canal  $2\ell SS0\tau$ , un test de fermeture est réalisé. La région utilisée est définie par trois leptons, un jet étiqueté *b* et 2 ou 3 jets. La Table 7.7 présente les nombres d'événements prédits et observés. Le test de fermeture montre un bon accord entre la prédiction et les données.

Non-prompt	Prédiction totale	Données	Données/Pred
$58.5527 \pm 10.1504$ (stat)	$75.7363 \pm 10.1504$	65	$0.85824 \pm 0.10645$ (stat) $\pm 0.14878$ (fakes)

TABLE 7.7 – Comparaison de la prédiction du nombre d'événements avec des leptons non-prompts dans le canal  $3\ell$  (2-3 jets) aux données.

Le nombre d'événements avec un lepton secondaire prédit pour le canal  $3\ell$  est donné dans la Table 7.8.

Lepton secondaire	Prédiction
	$N \pm CR \pm \theta$ stat $\pm \theta$ syst
Electron	$= 10.6697 \pm 2.31711 \pm 2.5777 \pm 2.64522$
Muon	$= 9.30877 \pm 2.4828 \pm 1.73349 \pm 0.908551$

TABLE 7.8 – Prédiction du nombre d'événements dans le canal  $3\ell$  avec un électron ou un muon secondaire.

## 7.6 Incertitudes systématiques

Plusieurs incertitudes systématiques doivent être prises en compte. Elles peuvent être regroupées en plusieurs catégories :

- Les incertitudes **théoriques** portent sur la normalisation des bruits de fond due aux échelles de renormalisation et de factorisation considérées dans le calcul des sections efficaces. Les réglages utilisés pour la modélisation des processus lors de leur génération sont aussi comptabilisés dans ces incertitudes.
- Les incertitudes **instrumentales** couvrent toutes les incertitudes liés aux objets utilisés dans l'analyse.
- Les incertitudes qui portent sur les bruits de fond estimés à partir des données : les événements avec des leptons secondaires ou avec un électron dont la charge est mal reconstruite. Pour ces bruits de fond, les incertitudes appliquées sont détaillées dans la Section 7.5.2 pour les leptons secondaires et dans le Chapitre 5 pour les événements avec un électron dont la charge électrique est mal reconstruite.

### 7.6.1 Les incertitudes théoriques

Les incertitudes théoriques dépendent des processus :

- Pour le bruit de fond **Diboson**, une incertitude de 50% est appliquée sur la normalisation. Cette incertitude conservative est ajoutée du fait que la contribution principale de ce bruit passe par le processus  $WZ + b\bar{b}$  qui n'a jamais été mesuré au LHC. La première mesure a été obtenue avec une incertitude proche de 100%, or ce processus représente environ 50% des dibosons observés dans les régions de signal. Une incertitude de 50% a été appliquée à ce bruit de fond.

- Pour les bruits de fond  $t\bar{t}V$ , une première incertitude est extraite à partir du calcul de la section efficace de production au niveau NLO. Cette incertitude est séparée en une contribution de l'échelle QCD et une seconde liée à la PDF et est décrite dans la Table ???. Un autre groupe de systématiques théoriques portant sur les échelles de renormalisation et de factorisation ainsi que sur l'ajustement des paramètres du générateur PYTHIA8 est présenté dans les Tables 7.10 et 7.11. Une dernière incertitude est ajoutée pour le processus  $t\bar{t}W$  à partir d'une comparaison d'un générateur NLO avec un générateur LO.
- Pour le signal  $t\bar{t}H$ , des incertitudes similaires à celles dérivées pour les processus  $t\bar{t}V$  sont aussi considérées et décrites dans les Tables 7.12 et 7.13.

Process	Section efficace [fb]		
	NLO 13 TeV		
$pp \rightarrow t\bar{t} W^\pm$	600.8	$+12.9\%$ $-11.5\%$	$\pm 2.7\%$
$pp \rightarrow t\bar{t} Z$	839.3	$+9.6\%$ $-11.3\%$	$\pm 2.8$

TABLE 7.9 – Section efficace NLO pour la production des processus  $t\bar{t}V$  à  $\sqrt{s} = 13$  TeV.

Channel	Shower	Scale var	PDF var
2l1tau	$\pm 5.5\%$	$+8.0\%$ $-10.4\%$	$+1.7\%$ $-1.8\%$
ee5j	$\pm 6.6\%$	$+6.7\%$ $-8.8\%$	$+1.6\%$ $-3.1\%$
mumu5j	$\pm 9.5\%$	$+7.6\%$ $-10.2\%$	$+1.6\%$ $-2.3\%$
emu5j	$\pm 2.5\%$	$+8.8\%$ $-11.7\%$	$+1.5\%$ $-2.6\%$
trilep	$\pm 1.5\%$	$+8.8\%$ $-10.7\%$	$+1.5\%$ $-2.3\%$
quadlep	$\pm 1.7\%$	$+10.1\%$ $-11.2\%$	$+1.3\%$ $-2.4\%$

TABLE 7.10 – Incertitudes théoriques pour le processus  $t\bar{t}l\bar{l}$ .

Channel	Shower	Scale var	PDF var	NLO vs LO
2l1tau	$\pm 3.1\%$	$+17.5\%$ $-14.1\%$	$+1.8\%$ $-2.4\%$	$\pm 30.5\%$
ee5j	$\pm 6.6\%$	$+11.3\%$ $-11.0\%$	$+2.0\%$ $-3.1\%$	$\pm 17.3\%$
mumu5j	$\pm 6.5\%$	$+14.6\%$ $-13.3\%$	$+1.4\%$ $-3.2\%$	$\pm 6.7\%$
emu5j	$\pm 2.5\%$	$+16.2\%$ $-14.1\%$	$+2.8\%$ $-3.6\%$	$\pm 13.3\%$
trilep	$\pm 8.0\%$	$+12.7\%$ $-11.3\%$	$+2.1\%$ $-2.7\%$	$\pm 1.6\%$
quadlep	$\pm 14.9\%$	$+14.3\%$ $-12.8\%$	$+3.1\%$ $-3.1\%$	–

TABLE 7.11 – Incertitudes théoriques pour le processus  $t\bar{t}W$ . Pour le canal  $4\ell$ , la comparaison des générateurs LO vs NLO n'est pas utilisée à cause d'un nombre d'événements trop faible.

Section efficace $t\bar{t}H$ à $\sqrt{s} = 13$ TeV [pb]	
Valeur après calcul théorique	NLO QCD+EW
CERN Yellow Report 4 [73]	$0.5071^{+5.8\%+3.6\%}_{-9.2\%-3.6\%}$ (PDF incl. $\alpha_S$ )

TABLE 7.12 – Section efficace total pour le processus  $t\bar{t}H$  à  $\sqrt{s} = 13$  TeV telle que calculée dans [73] à NLO QCD+EW avec l'approximation d'une largeur nulle.

Channel	Shower	Scale var	PDF var	Herwig vs Pythia
2l1tau	+5.4% -5.4%	+8.0% -10.4%	+1.7% -1.8%	±3.6%
ee5j	+5.8% -5.8%	+4.8% -12.4%	-1.7% -6.2%	±16.5%
mumu5j	+5.8% -5.8%	+6.9% -10.8%	+1.0% -2.8%	±7.3%
emu5j	+7.2% -7.2%	+5.8% -9.6%	+0.7% -2.9%	±7.1%
trilep	+4.1% -4.1%	+7.3% -10.3%	+1.6% -3.1%	±0.1%
quadlep	+1.8% -1.8%	+6.9% -10.3%	+2.5% -4.3%	±0.1%

TABLE 7.13 –  $t\bar{t}H$  Shower variations

## 7.6.2 Incertitudes systématiques instrumentales

Les incertitudes instrumentales sont appliquées sur tous les objets reconstruits de l'analyse. Toutes les incertitudes sont évaluées événement par événement. Elles portent sur les efficacités du système de déclenchement et de l'étiquetage des jets de *b*, la reconstruction et l'identification des leptons, la calibration de l'énergie des jets.

Ces incertitudes peuvent être appliquées de deux manières, soit par repondération du poids appliqué à l'événement, soit par un renormalisation de l'énergie ou de l'impulsion de l'objet ( $p_T$  Correction). Les incertitudes instrumentales présentes dans cette analyse sont regroupées dans les Tables 7.14, 7.15 et 7.16. Pour chacune de ces systématiques, la méthode d'application est indiquée.

Une autre incertitude instrumentale prise en compte est celle sur la mesure de la luminosité intégrée lors des collisions. Elle est de 2.9% et est appliquée à tous les processus simulés par Monte-Carlo y compris le signal.

Une dernière incertitude porte sur la repondération des événements simulés afin que la modélisation de l'effet d'empilement reproduise l'effet observé. Cette incertitude, PRW\_DATASF, est elle aussi appliquée sur tous les processus simulés.

## 7.7 Validation de la simulation des bruits de fond

Avant toute observation de données dans la région signal, des régions enrichies en différents bruits de fond sont utilisés pour tester leur modélisation. Ce travail est fait pour tous les bruits de fonds majoritaires dans chacun des canaux considérés. La Table 7.17 donne la définition de régions de validation définies pour le processus  $t\bar{t}W$  dans le canal  $2\ell SS0\tau$ , pour le processus  $t\bar{t}Z$  dans le canal  $3\ell$  et le processus  $WZ + bb/cc$  aussi pour le canal  $3\ell$ . Les distributions des prédictions MC et des données dans chacune de ces régions selon le nombre d'électrons et le nombre de jets sont présentées sur la Figure 7.4.

## 7.8 Résultats

La validation des bruits de fond et leurs estimations étant effectuées, une comparaison est faite entre la prédiction et les données dans la région enrichie en signal. Les nombres d'événements prédit et observé pour les différents canaux sont présentés dans la Table 7.18 tandis que les Figures 7.5, 7.6, 7.7, 7.8, 7.9 et 7.10 montrent les distributions des événements de données et de simulation selon différentes observables. Ces distributions ne sont pas utilisées lors de l'ajustement aux données.



<b>Incertitudes instrumentales sur les leptons</b>			
Type	Description	Nom	Application
<b>Système de déclenchement</b>			
Facteurs d'échelles	Efficacité du système	lepSFTrigTight_MU(EL)_SF_Trigger_STAT(SYST)	Repondération
<b>Muons</b>			
Efficacités	Reconstruction et Identification	lepSFObjTight_MU_SF_ID_STAT(SYST)	Repondération
	Isolation	lepSFObjTight_MU_SF_Isol_STAT(SYST)	Repondération
	Track To Vertex Association	lepSFObjTight_MU_SF_TTVA_STAT(SYST)	Repondération
Échelle sur $p_T$	Échelle sur $p_T$	MUONS_SCALE	$p_T$ Correction
Résolution	Détecteur Interne	MUONS_ID	$p_T$ Correction
	Résolution en énergie		
	Spectromètre à muon	MUONS_MS	$p_T$ Correction
	Résolution en énergie		
<b>Electrons</b>			
Efficacités	Reconstruction	lepSFObjTight_EL_SF_ID	Repondération
	Identification	lepSFObjTight_EL_SF_Reco	Repondération
	Isolation	lepSFObjTight_EL_SF_Isol	Repondération
Facteur d'échelle	Échelle d'énergie	EG_SCALE_ALL	Ener.Correction
Résolution	Résolution en énergie	EG_RESOLUTION_ALL	Ener.Correction
<b>Taus Hadroniques</b>			
Efficacités	Reconstruction	tauSFLoose_TAU_SF_RECO_TOTAL	Repondération
	BDT Identification	tauSFTight_TAU_SF_JETID_TOTAL	Repondération
	BDT Electron Veto	tauSFTight_TAU_SF_ELEOLR_TOTAL	Repondération
Facteur d'échelle	Échelle en $p_T$	TAUS_TRUEHADTAU_SME_TES_TOTAL	$p_T$ Correction

TABLE 7.14 – Incertitudes systématiques instrumentales appliquées aux leptons (tau hadronique compris).

Incertitudes instrumentales sur les Jets et la MET			
Type	Origine	Nom	Application
<b>Jets</b>			
Jet Vertex Tagger		JVT	Repondération
Échelle d'énergie	Méthode de calibration ( <i>e.g.</i> In-Situ)	JET_19NP_	$p_T$ Correction
		JET_EffectiveNP_1,2,3,4,5,6restTerm	$p_T$ Correction
		JET_GroupedNP_1	$p_T$ Correction
	$\eta$ inter-calibration	JET_EtaIntercalibration_Modelling	$p_T$ Correction
		JET_EtaIntercalibration_TotalStat	$p_T$ Correction
	Jets à haut $p_T$	JET_SingleParticle_HighPt	$p_T$ Correction
	Effet d'empilement	JET_Pileup_OffsetNPV	$p_T$ Correction
		JET_Pileup_OffsetMu	$p_T$ Correction
		JET_Pileup_PtTerm	$p_T$ Correction
		JET_Pileup_RhoTopology	$p_T$ Correction
Test de Fermeture	JET_PunchThrough_MC15	$p_T$ Correction	
Saveur	JET_Flavor_Response	$p_T$ Correction	
	JET_BJES_Response	$p_T$ Correction	
	JET_Flavor_Composition	$p_T$ Correction	
Résolution		JET_JER_SINGLE_NP	Repondération
<b>MET</b>			
Soft Tracks Terms	Résolution	MET_SoftTrk_ResoPerp	$p_T$ Correction
	Résolution	MET_SoftTrk_ResoPara	$p_T$ Correction
	Échelle	MET_SoftTrk_ScaleUp	$p_T$ Correction

TABLE 7.15 – Incertitudes systématiques appliquées sur la calibration de l'énergie des jets et de l'énergie transverse manquante.

Systématiques instrumentales sur l'étiquetage des b-jets		
Type	Origine	Nom
b-tags		
Facteurs d'échelle	Efficacité sur MV2c20 b-tagger sur les jets étiquetés <i>b</i> (bins en $\eta$ )	MV2c20_77_EventWeight_B0-5
	Efficacité sur MV2c20 b-tagger sur les jets étiquetés <i>c</i> (bins en $\eta$ )	MV2c20_77_EventWeight_C0-3
	Efficacité sur MV2c20 b-tagger sur les jets issus de quarks légers dans des bins en $\eta$ et $p_T$	MV2c20_77_EventWeight_Light0-11
	MV2c20 b-tagger Efficacité d'extrapolation	MV2c20_77_EventWeight_extrapolation MV2c20_77_EventWeight_extrapolation_from_charm

TABLE 7.16 – Incertitudes systématiques sur la reconstruction et l'étiquetage des jets provenant de hadrons beaux.

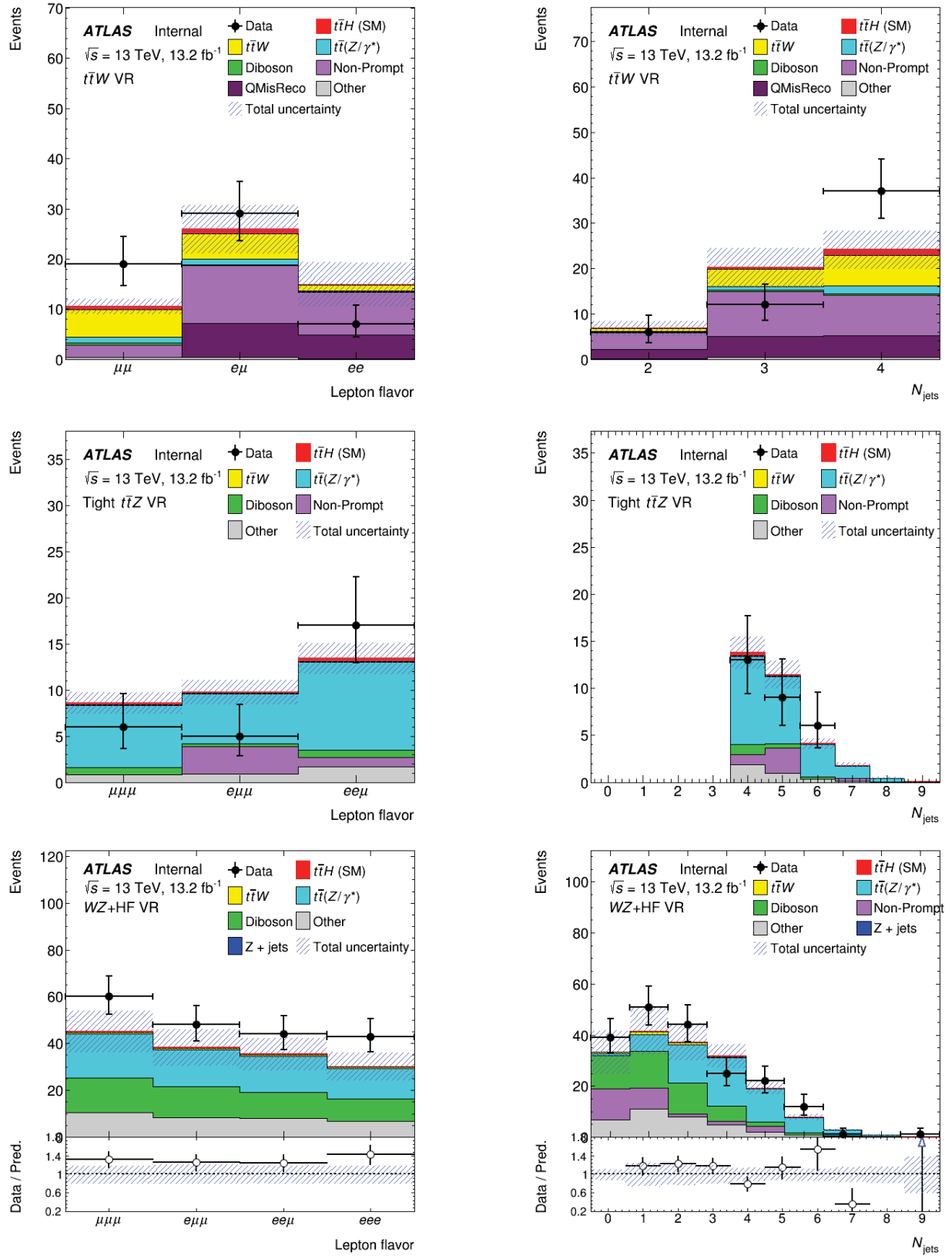


FIGURE 7.4 – Distribution du nombre de jets et du nombre d'électrons dans les régions de contrôle visant le processus  $t\bar{t}W$ (a) et (b), le processus  $t\bar{t}Z$ (c) et (d) et le processus  $WZ + bb$  (e) et (f).

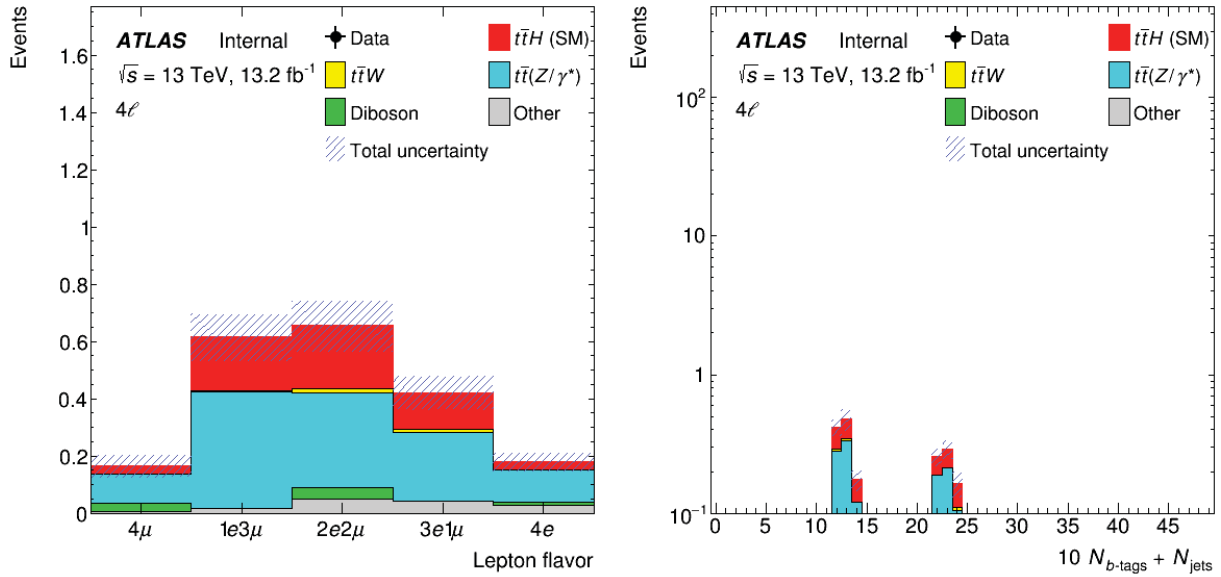


FIGURE 7.5 – Distribution des événements simulés et de données pour le canal  $4\ell$ . Nombre d'électrons (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite). Il n'y a pas de donnée observée dans cette région.

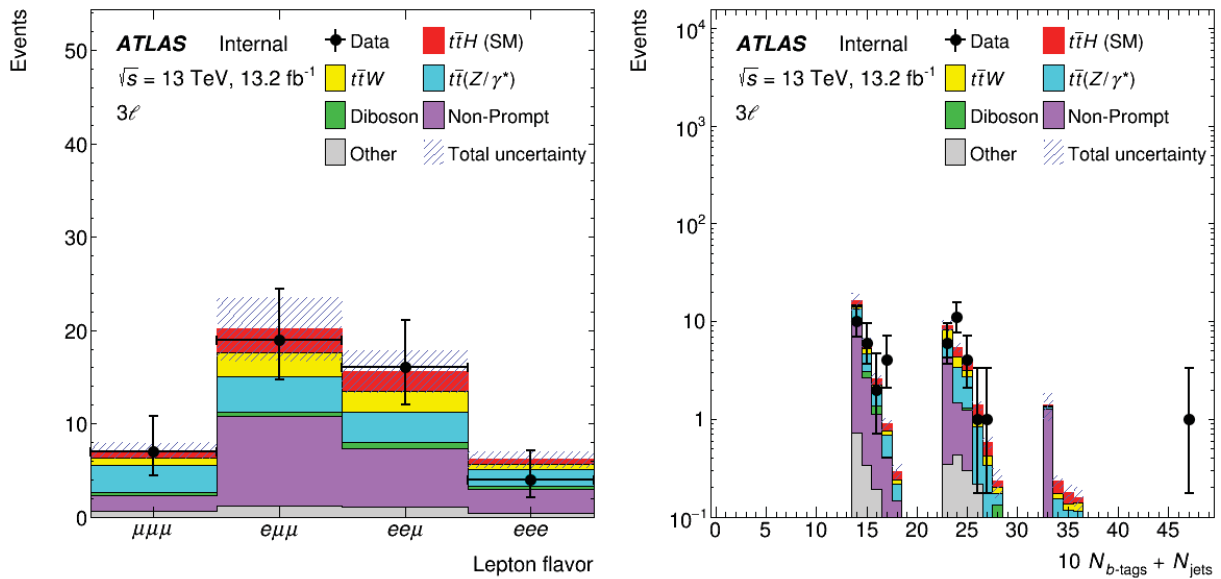


FIGURE 7.6 – Distribution des événements simulés et de données pour le canal  $3\ell$ . Nombre d'électrons (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite).

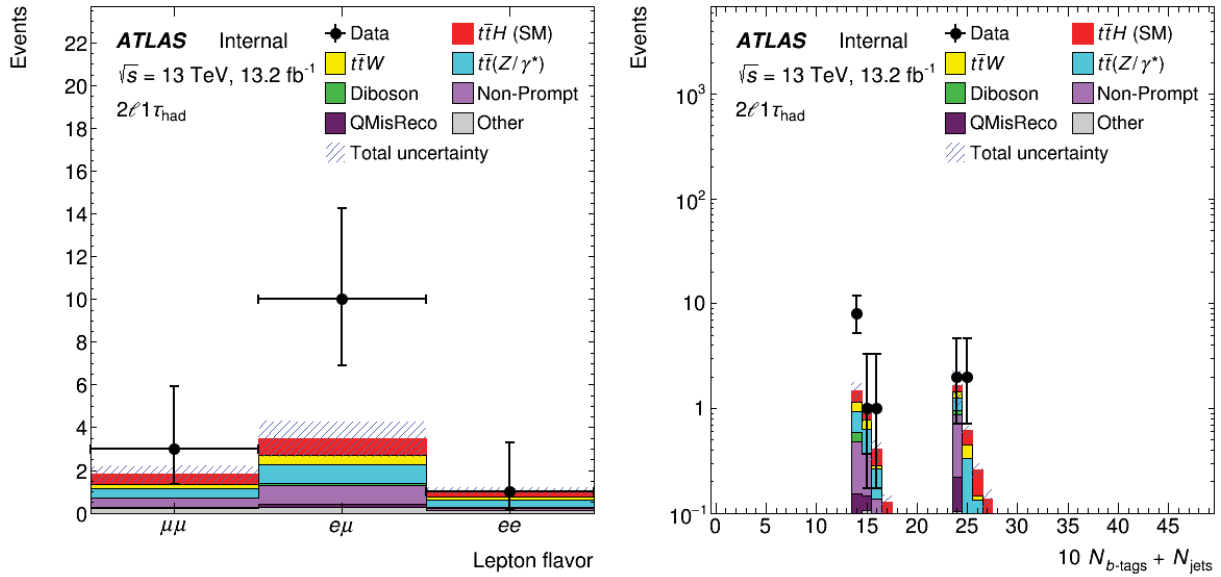


FIGURE 7.7 – Distribution des événements simulés et de données pour le canal  $2\ell SS1\tau$ . Nombre d’électrons (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite).

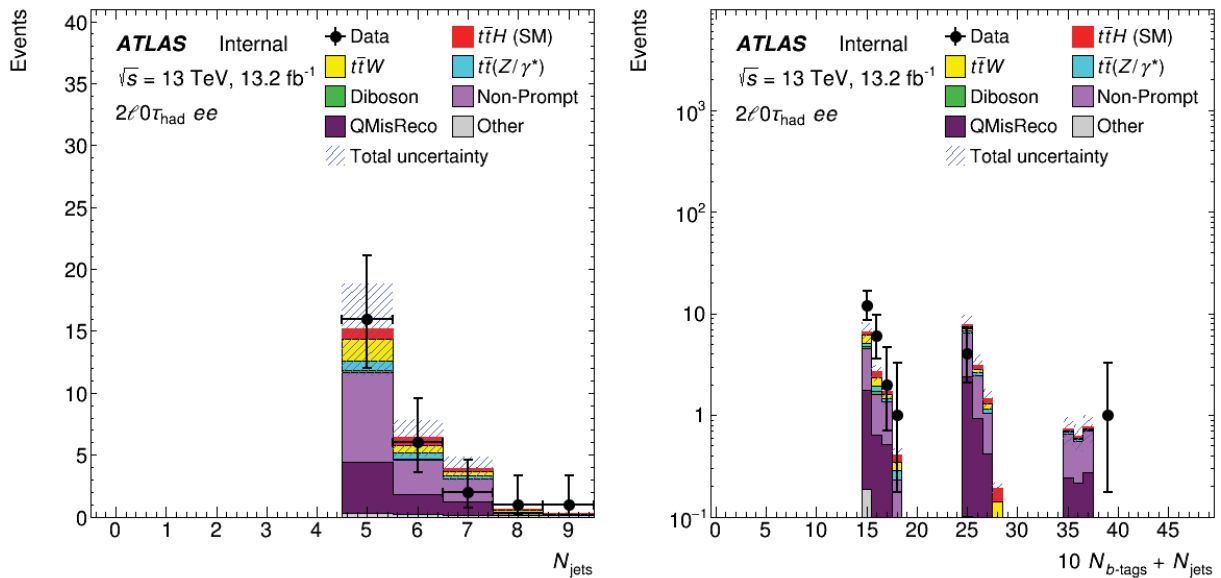


FIGURE 7.8 – Distribution des événements simulés et de données pour le canal  $ee$ . Nombre de jets (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite).

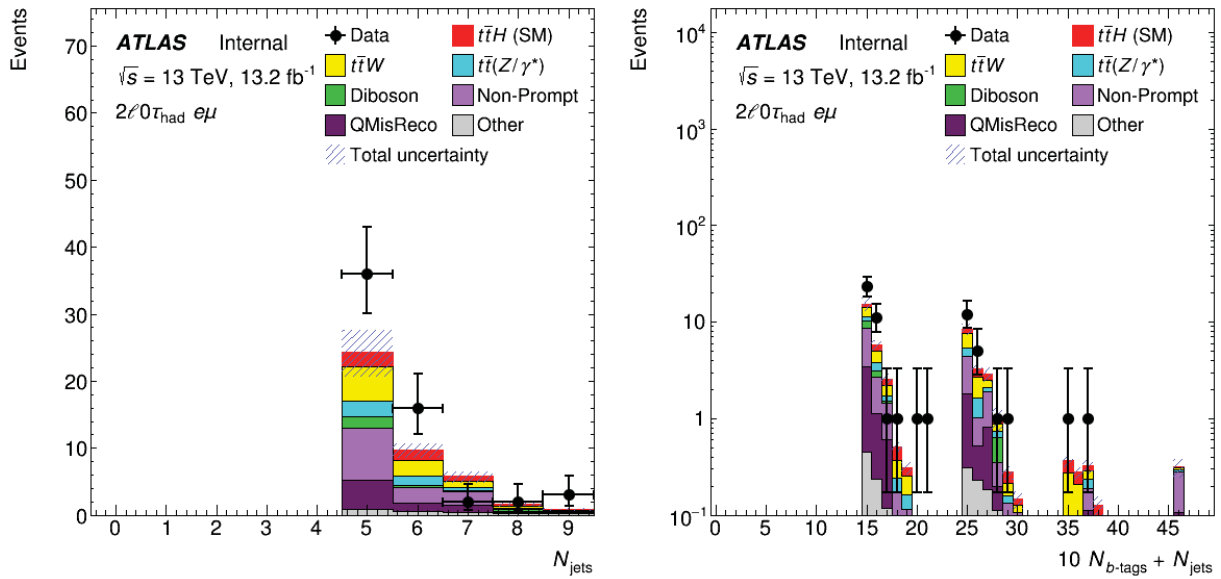


FIGURE 7.9 – Distribution des événements simulés et de données pour le canal  $e\mu$ . Nombre de jets (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite).

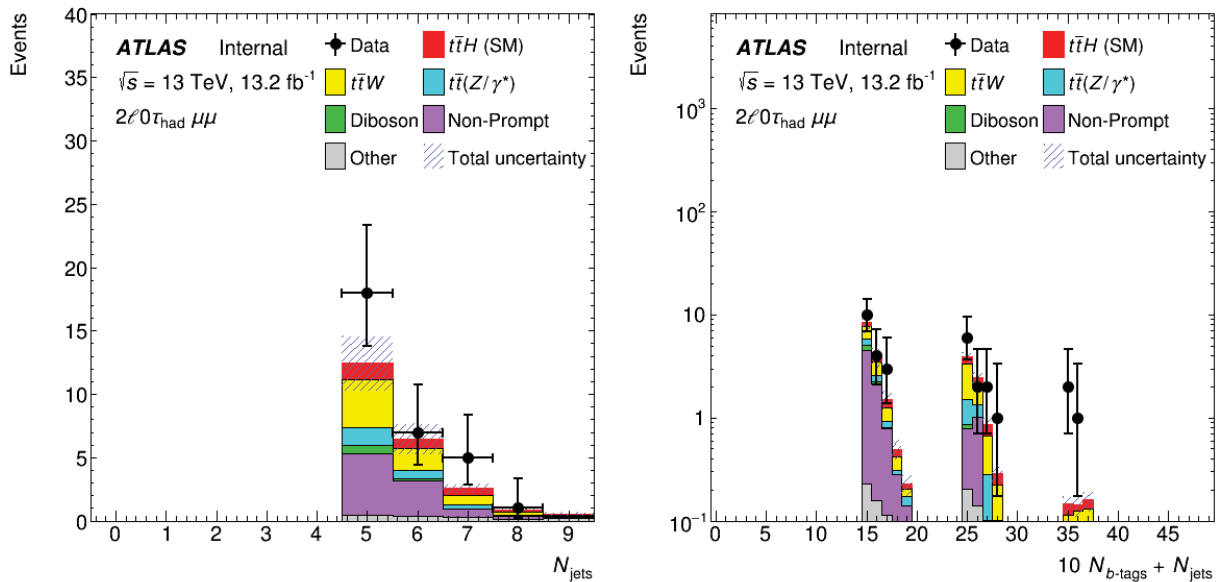


FIGURE 7.10 – Distribution des événements simulés et de données pour le canal  $\mu\mu$ . Nombre de jets (gauche) and nombre de jets +  $10\times$  nombre de jets étiquetés  $b$  (droite).

Région de validation	Sélection
$t\bar{t}W$	$2\ell SS0\tau$ sélection des leptons $\geq 2$ jets étiquetés $b$ et $\leq 4$ jets $H_T(\text{jets}) > 220$ GeV dans les cas $ee$ et $e\mu$ $M_{ee} \notin [75,105]$ GeV et $E_T^{\text{miss}} > 50$ GeV dans le cas de deux électrons
$WZ + 1b\text{-tag on-shell}$	$3\ell$ sélection des leptons  Masse invariante des deux leptons de même saveur dans une fenêtre de 10GeV autour de celle du boson Z exactement un jet étiqueté $b$
$t\bar{t}Z$	$3\ell$ sélection des leptons et jets Masse invariante des deux leptons de même saveur dans une fenêtre de 10GeV autour de celle du boson Z Exactement 4j dont 2 étiquetés $b$

TABLE 7.17 – Description des regions de validation utilisées pour valider la modélisation des différents bruits de fond.

TABLE 7.18 – Nombres d'événements attendus et observés dans les six régions enrichies en signal. Toutes les sources d'incertitudes sont considérées.

	$2\ell0\tau ee$	$2\ell0\tau e\mu$	$2\ell0\tau \mu\mu$	$2\ell1\tau$	$3\ell$	$4\ell$
$t\bar{t}W$	$2.9 \pm 0.7$	$9.1 \pm 2.5$	$6.6 \pm 1.6$	$0.8 \pm 0.4$	$6.1 \pm 1.3$	—
$t\bar{t}(Z/\gamma^*)$	$1.55 \pm 0.29$	$4.3 \pm 0.9$	$2.6 \pm 0.6$	$1.6 \pm 0.4$	$11.5 \pm 2.0$	$1.12 \pm 0.20$
Diboson	$0.38 \pm 0.25$	$2.5 \pm 1.4$	$0.8 \pm 0.5$	$0.20 \pm 0.15$	$1.8 \pm 1.0$	$0.04 \pm 0.04$
Leptons secondaires	$12 \pm 6$	$12 \pm 5$	$8.7 \pm 3.4$	$1.3 \pm 1.2$	$20 \pm 6$	$0.18 \pm 0.10$
Mauvaise Reco Charge	$6.9 \pm 1.3$	$7.1 \pm 1.7$	—	$0.24 \pm 0.03$	—	—
Autres	$0.81 \pm 0.22$	$2.2 \pm 0.6$	$1.4 \pm 0.4$	$0.63 \pm 0.15$	$3.3 \pm 0.8$	$0.12 \pm 0.05$
Bruit de fond total	$25 \pm 6$	$38 \pm 6$	$20 \pm 4$	$4.8 \pm 1.4$	$43 \pm 7$	$1.46 \pm 0.25$
$t\bar{t}H(\text{MS})$	$2.0 \pm 0.5$	$4.8 \pm 1.0$	$2.9 \pm 0.6$	$1.43 \pm 0.31$	$6.2 \pm 1.1$	$0.59 \pm 0.10$
Données	26	59	31	14	46	0

### 7.8.0.1 Pre-fit

Le paramètre d'intérêt considéré dans l'étude statistique est la force du signal  $\mu$  définie comme le rapport entre la section efficace observée et la section efficace théorique du processus  $t\bar{t}H$ . Pour extraire cette valeur et son incertitude, un ajustement aux données est effectué par maximisation de vraisemblance [74] tel que décrit dans le Chapitre 6.

Dans un premier temps, un ajustement est fait en supposant que les données correspondent exactement à la prédiction du signal et du bruit de fond. Les pseudo-données ainsi générées sont appelées données *Asimov*. La force du signal obtenue est par construction de 1 et l'incertitude donne la précision attendue de l'analyse. Le résultat pour l'analyse présentée est  $\mu=1_{-1.0}^{+1.2}$  et le résultat obtenu canal par canal est donné dans la Table 7.19. Les incertitudes sont séparées selon la composante statistique liée à la taille limitée des échantillons MC et selon la composante systématique due aux incertitudes décrites auparavant. Ainsi les deux canaux apportant le plus de sensibilité sont les canaux  $2\ell SS0\tau$  et  $3\ell$  avec des incertitudes systématiques qui participent autant que les incertitudes statistiques.

Par la suite un ajustement aux données est effectué, cette fois sur les données observées du

Canal	$\mu$	Stat	Syst	Tot
Canal $4\ell$				
$4\ell$	1.	+3.0 -1.9	+0.7 -0.4	+3.1 -1.9
Canal $3\ell$				
$3\ell$	1.	+1.2 -1.1	+0.9 -1.3	+1.8 -1.7
Canal $2\ell SS1\tau$				
$2\ell 1\tau$	1	+1.9 -1.5	+1.4 -0.9	+2.4 -1.7
Canal $2\ell SS0\tau$				
$2\ell 0\tau$	1	+1.0 -0.9	+1.3 -1.2	+1.6 -1.5
Canal $2\ell SS0\tau$ divisé par saveurs				
$ee$ (5j,1b)	1.	+2.8 -2.4	+3.0 -3.4	+4.1 -4.1
$e\mu$ (5j,1b)	1.	+1.4 -1.3	+1.3 -1.2	+1.9 -1.8
$\mu\mu$ (5j,1b)	1.	+1.8 -1.6	+1.6 -1.5	+2.4 -2.2
Total	1.0	+0.7 -0.6	+0.9 -0.8	+1.2 -1.0

TABLE 7.19 – Force du signal ( $\mu$ ) attendue pour un ajustement effectué sur des données Asimov.

LHC. Le résultat final est une force du signal  $\mu = 2.5^{+1.3}_{-1.1}$ . La Figure 7.11 présente le résultat ainsi que la force du signal obtenue canal par canal. Au final, les données observées correspondent à un processus  $t\bar{t}H$  dont la section efficace de production serait 2.5 fois plus grand que la prédiction du modèle standard. Les incertitudes et la quantité de données ne permettent pas de conclure quant à la mise en évidence de la production de bosons de Higgs en association avec une paire de quarks *top*.

Une limite est extraite et la méthode CLs est utilisée. Elle donne une limite haute à la valeur que peut prendre la force du signal selon les données observées. Les limites sur la force du signal sont regroupées dans la Table 7.20. Les différentes valeurs correspondent à la limite observée, à la limite médiane correspondant à des données Asimov et une hypothèse bruit de fond seul<sup>3</sup>. Enfin la dernière valeur correspond à la limite attendue pour des données Asimov avec l'hypothèse bruit de fond et signal.

L'impact des différentes incertitudes systématiques sur le résultat de l'ajustement aux données est présenté dans la Table 7.21. Les incertitudes les plus importantes sont celles appliquées aux estimations des bruits de fond instrumentaux ainsi que sur l'association d'un jet au vertex primaire (*JVT* pour *Jet Vertex Tagger*) et sur l'étalonnage de l'énergie des jets. L'impact des différentes systématiques est aussi présenté sur la Figure 7.12 avec plus de détails. L'incertitude la plus importante porte sur la *JVT* puis sur l'estimation des leptons non-prompts dans les canaux  $2\ell SS0\tau$  et  $3\ell$ .

## 7.9 Conclusion

L'analyse présentée dans ce Chapitre exploite  $13.2 \text{ fb}^{-1}$  de données collectées en 2015 et 2016. Sept régions de signal proches de celles définies pour l'analyse des données du Run 1 du LHC sont utilisées en se basant sur des expériences de comptage. L'estimation des événements de bruits de fond dus à la mauvaise identification de la charge des électrons a été utilisée dans les canaux avec deux leptons de même charge électrique. Une estimation des bruits de fond dus aux leptons non-prompts par la méthode *Fake Factor* a été développée et appliquée aux canaux  $2\ell SS0\tau$  et

3. Dans ce cas les données Asimov sont générées telles que correspondant à la prédiction du bruit de fond seul ( $\mu = 0$ ).



Canal	Observée	Médiane	Intervalle 68% CL	Intervalle 95% CL	Avec injection de signal
Canal $4\ell$					
$4\ell$	5.2	6.6	[5.2;9.4]	[5.0;11.7]	4.3
Canal $3\ell$					
$3\ell$	3.9	3.5	[2.5;5.0]	[1.9;6.4]	4.3
Canal $2\ell 1\tau$					
$2\ell 1\tau$	12.9	5.9	[4.3;8.8]	[3.2;13.8]	6.3
Canal $2\ell 0\tau$					
$2\ell 0\tau$	7.8	3.5	[2.5;5.2]	[1.9;7.9]	4.2
Canal $2\ell 0\tau$ divisé par saveurs					
$ee$ (5j,1b)	9.3	8.7	[6.2;13.0]	[8.7;21.0]	
$e\mu$ (5j,1b)	8.9	4.5	[3.3;6.6]	[4.5;10.]	
$\mu\mu$ (5j,1b)	9.2	5.2	[3.7;7.7]	[2.8;9.7]	
Total	4.9	2.3	[1.7;3.4]	[1.2;5.1]	3.1

TABLE 7.20 – Limites attendues et observées pour les différents canaux.

TABLE 7.21 – Résumé de l'impact des systématiques sur  $\mu$ .

Source	$\Delta\mu$	
Leptons non-prompt et mauvaise reconstruction de la charge	+0.56	-0.64
Jet-vertex association, pileup modelling	+0.48	-0.36
Incertitude $t\bar{t}W$ théorique	+0.29	-0.31
Incertitude $t\bar{t}H$ théorique	+0.31	-0.15
Échelle d'énergie des jets et résolution	+0.22	-0.18
Incertitude $t\bar{t}Z$ théorique	+0.19	-0.19
Luminosité	+0.19	-0.15
Incertitude Diboson théorique	+0.15	-0.14
Etiquetage des jets de $b$	+0.15	-0.12
ID, isolation, trigger des lepton légers ( $e, \mu$ ) et $\tau_{\text{had}}$	+0.12	-0.10
Incertitude théorique sur les autres bruits de fond	+0.11	-0.11
Total	+1.1	-0.9

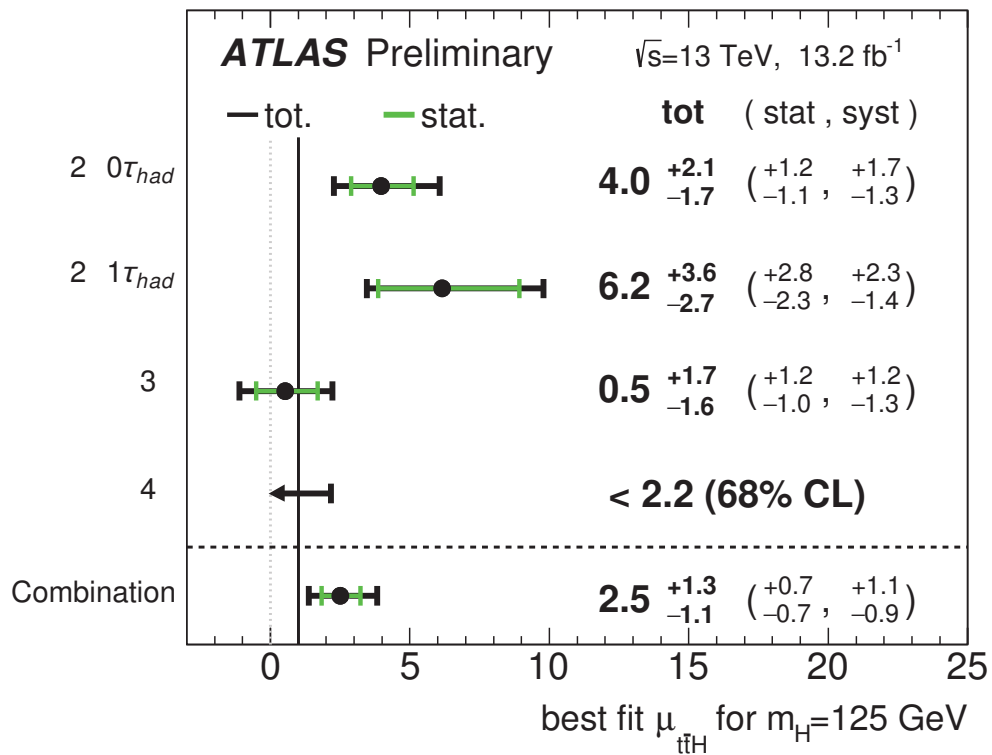


FIGURE 7.11 – Force du signal ( $\mu$ ) mesurée dans les données pour les différents canaux ainsi que pour leur combinaison.

3ℓ. Cette méthode a été choisie pour sa simplicité qui convient pour une analyse de comptage. Les incertitudes associées à ces prédictions ont été étudiées et de l'ordre de 40%. Elles ont pour source principale la taille des données.

L'interprétation statistique de l'ensemble des canaux leptoniques a également été effectuée. La force du signal a été déterminée comme étant  $\mu=1^{+1.2}_{-1.0}$ . La précision de l'analyse et le volume des données disponible ne permettent pas une mise en évidence de la production associée  $t\bar{t}H$ . Cette analyse peut être améliorée par l'utilisation d'information topologiques et d'outils statistiques décrits dans le Chapitre 9.

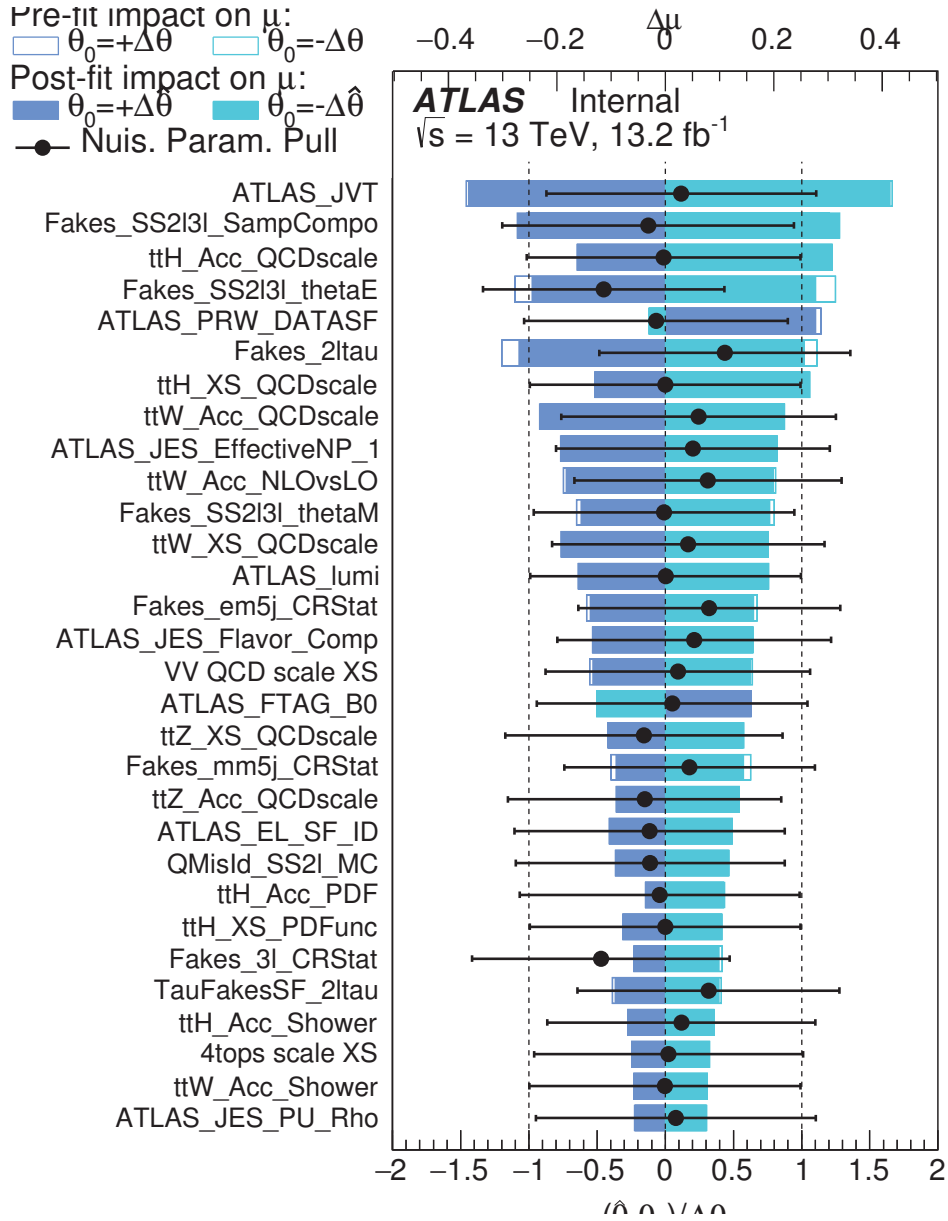


FIGURE 7.12 – Classement des paramètres de nuisance selon leur impact sur la force mesurée du signal lors de l’ajustement aux données de tous les canaux.



# Reconstruction cinématique des événements

---

## 8.1 Introduction

Pour améliorer la sensibilité de l'analyse, la reconstruction des différentes particules (quarks *top*, boson de Higgs, boson W) a été étudiée dans le cadre d'un état final avec deux leptons de même charge électrique. Pour mettre en place, une stratégie pour la reconstruction d'un état final, le choix d'un processus considéré doit être fait. Deux options ont été considérées :

- Reconstruction du processus  $t\bar{t}H$
- Reconstruction du processus  $t\bar{t}$  pour un état final semileptonique  $\rightarrow$  l'un des leptons est donc un lepton non-prompt

Toutefois la reconstruction du processus  $t\bar{t}H$  est rendue plus complexe par la présence de deux neutrinos dans l'état final ainsi que d'un boson W provenant du boson de Higgs n'étant pas sur sa couche de masse dans 50% des cas. Ce chapitre traite donc de la reconstruction du processus  $t\bar{t}$ . Pour cette reconstruction les événements utilisés ont au moins 4 jets dont 1 étiqueté *b*. Un ajustement cinématique est utilisé [75].

## 8.2 Combinatoire dans la reconstruction des quarks *top*

Pour un état final semileptonique provenant du processus  $t\bar{t}$ , le nombre de jets attendus est de 4 dont 2 étiquetés *b*. En outre un lepton et un neutrino sont présents dans cet état final. Or dans le cas considéré, deux leptons sont présents dans l'état final et le plus souvent plus de quatre jets sont reconstruits. On appelle un *top* hadronique un quark *top* dont le boson W associé se désintègre en paire  $q\bar{q}$  et un *top* leptonique dont le boson W associé se désintègre en un lepton et un neutrino. Pour reconstruire le quark *top* hadronique trois jets sont nécessaires et il faut donc retrouver la bonne combinaison de jets pour le reconstruire. De même dans le cas du quark *top* leptonique, la bonne combinaison de lepton et de quark doit être choisie pour la reconstruction. La combinatoire du  $t\bar{t}$  est l'ensemble des combinaisons possibles de quarks et leptons pour la reconstruction de la paire  $t\bar{t}$ . La combinatoire du boson W hadronique, du quark *top* hadronique et du boson W leptonique existent de la même façon.

La combinatoire du  $t\bar{t}$  est réduite en considérant les jets étiquetés *b*. En effet sachant que l'un des jets est étiqueté *b*, il n'entre pas dans la combinatoire du boson W hadronique. Si deux jets sont étiquetés *b*, ces deux jets sont considérés comme correspondant aux deux quarks *b* attendus. Dans le cas où plus de deux jets sont étiquetés *b*, toutes les combinaisons sont testées pour la reconstruction en supposant que les jets correspondant aux quarks *b* sont présents parmi les jets étiquetés *b*.

Pour la reconstruction du quark *top* leptonique, la reconstruction du vecteur impulsion du neutrino est réalisée à partir de contrainte sur la masse du boson W. En effet, pour un état final  $t\bar{t}$  semileptonique un seul neutrino est attendu provenant de la désintégration d'un boson W. À partir de l'énergie transverse manquante (*MET*), on peut reconstruire les composantes *x* et *y* de l'impulsion du neutrino. Pour reconstruire la composante selon l'axe *z*, la masse du boson est connue et à partir de là, une équation du second degré est dérivée. Parmi les solutions, celle qui donne la

meilleure masse pour le quark  $top$  associé est choisie. Pour le choix du lepton, deux possibilités ont été testées :

- Seul le lepton de plus haut  $p_T$  est considéré pour la reconstruction car le cas le plus fréquent est que ce soit le lepton de plus bas  $p_T$  qui soit non-prompt et qui proviennent de la désintégration d'un hadron beau.
- Toutes les combinaisons sont testées et les deux leptons sont considérés pour trouver la bonne combinaison pour la reconstruction du quark  $top$  hadronique.

Deux possibilités existent aussi pour le traitement du lepton qui n'est pas testé dans la reconstruction. Soit le lepton n'est pas utilisé du tout, soit le lepton est fusionné avec le jet étiqueté  $b$  le plus proche, son énergie et son impulsion étant ajoutées à celles du jet.

Les options conservées dans la suite du Chapitre sont que les combinaisons avec les deux leptons soient faites et que le lepton non utilisé soit associé au jet de  $b$  le plus proche.

### 8.3 Ajustement cinématique

L'ajustement cinématique modifie des quantités cinématiques liées aux objets de l'état final afin de vérifier des contraintes. Cet ajustement prend en compte la résolution sur les quantités ajustées. Un ajustement cinématique améliore la résolution associée à ces quantités en tirant profit des contraintes. L'ajustement cinématique va, de manière générale, tester la compatibilité entre l'état final considéré et les contraintes. Il permet alors de choisir une combinaison d'objets voire de tester la compatibilité d'un événement de donnée avec la contrainte cinématique.

#### 8.3.1 Méthode

Pour chaque événement et pour chaque combinaison possible d'objets, un  $\chi^2$  est défini :

$$\chi^2 = \sum_j \frac{(E_j + \Delta E_j - E_j^0)^2}{\sigma_{E_j}^2} + \frac{(\phi_j + \Delta\phi_j - \phi_j^0)^2}{\sigma_{\phi_j}^2} + \frac{(\eta_j + \Delta\eta_j - \eta_j^0)^2}{\sigma_{\eta_j}^2} + \frac{(P_{MET} + \Delta P_{MET} - P_{MET}^0)^2}{\sigma_{P_{MET}}^2} + \frac{(m_W^{h(l)} - m_W^{PDF})^2}{\Gamma_W^2} + \frac{(m_{top}^{h(l)} - m_{top}^{PDF})^2}{\Gamma_{top}^2} \quad (8.1)$$

Les paramètres libres de ce  $\chi^2$  sont des quantités corrigées pour chaque objet :  $E_j$ ,  $\phi_j$  et  $\eta_j$  pour l'énergie, l'angle  $\phi$  et la pseudorapidité des jets,  $P_{MET}$  pour les trois composantes de l'impulsion du neutrino,  $m_W^{h(l)}$  pour la masse du boson W hadronique (leptonique) et  $m_{top}^{h(l)}$  pour la masse du quark  $top$  hadronique (leptonique).

Les paramètres  $E_j^0$ ,  $\phi_j^0$ ,  $\eta_j^0$  et  $P_{MET}^0$  sont les grandeurs mesurées.

Les termes  $\Delta$  et  $\sigma$  sont respectivement les résolutions et les biais attendus. Ils sont obtenus sur simulation Monte-Carlo.

Les termes  $\Gamma$  sont arbitraires et déterminent les contraintes. Les termes de masse sont utilisés pour contraindre l'ajustement avec une repondération par les constantes  $\Gamma$ . Les valeurs de ces constantes ont été ajustées pour avoir le meilleur compromis entre le gain en résolution et le taux de convergence.

La combinaison d'objets qui est conservée pour la reconstruction du processus  $t\bar{t}$  est celle qui minimise le  $\chi^2$ . Plus celui-ci est faible et plus la reconstruction est bonne.

#### 8.3.2 Définition des résolutions et des biais sur les objets

Pour estimer la résolution en angle et énergie des jets, une simulation Monte-Carlo de paire  $t\bar{t}$  est utilisée. Chaque jet est associé au parton correspondant. La différence en énergie (angle) entre le parton et le jet reconstruit permet l'extraction du biais et de la résolution. Le biais représente la

différence moyenne sur la grandeur entre le parton et le jet. La résolution est obtenue par ajustement gaussien. Ils sont modélisés en fonction de l'énergie, la pseudorapidité et la saveur du jet (étiqueté  $b$  ou non). Les Figures 8.1, 8.2 présentent la modélisation des résolutions et biais sur l'énergie pour des jets légers et étiquetés  $b$  dans les tonneaux. La figure 8.3 montre la résolution en énergie pour les jets légers dans le crack ( $1.4 < |\eta| < 1.7$ ) et dans les bouchons ( $1.7 < |\eta| < 2.6$ ). Finalement la Figure 8.4 présente la résolution en  $\eta$  et  $\phi$  pour des jets légers dans les tonneaux.

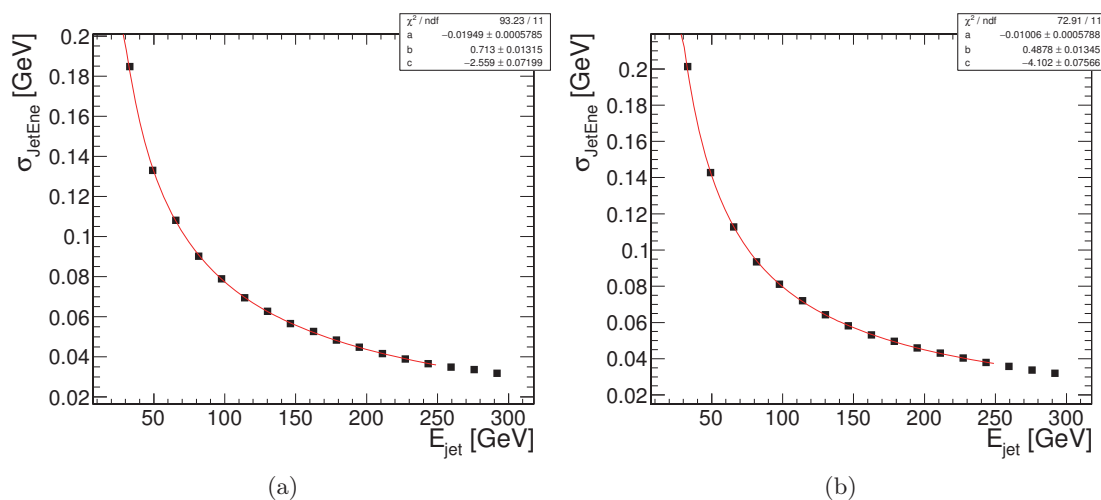


FIGURE 8.1 – Résolution en énergie des jets légers (a) et étiqueté  $b$  (b) dans la zone  $|\eta| < 1.4$ .

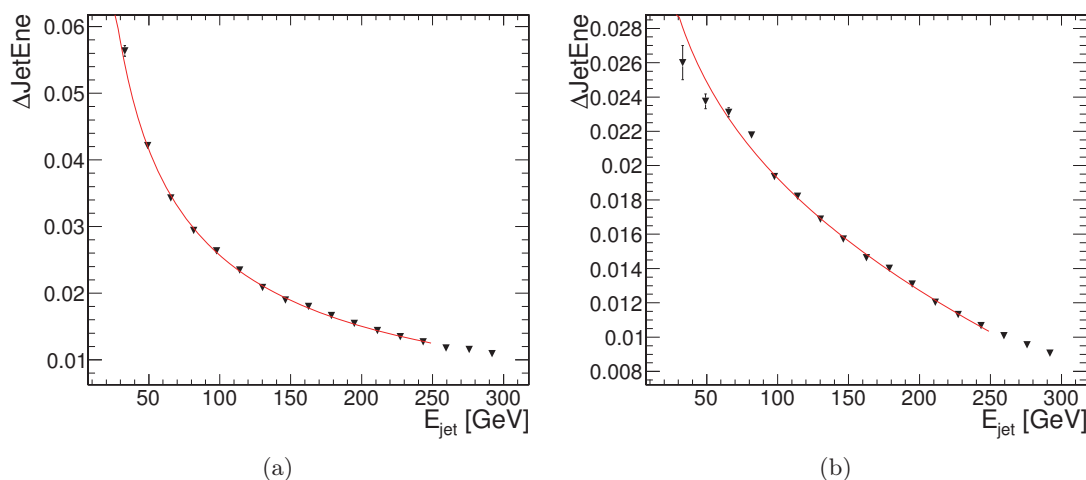


FIGURE 8.2 – Biais en énergie des jets légers (a) et étiqueté  $b$  (b) dans la zone  $|\eta| < 1.4$ .

## 8.4 Résultat

La reconstruction de l'événement est faite sous l'hypothèse du processus  $t\bar{t}$  et devrait donc être meilleure dans le cas du processus  $t\bar{t}$  que  $t\bar{t}H$ . La Figure ?? présente le  $\chi^2$  minimal par événement pour les processus  $t\bar{t}$  et  $t\bar{t}H$ . Comme attendu, la valeur moyenne est plus basse pour le processus  $t\bar{t}$  ce qui indique une meilleure reconstruction. Néanmoins, les événements  $t\bar{t}H$  ont une cinématique

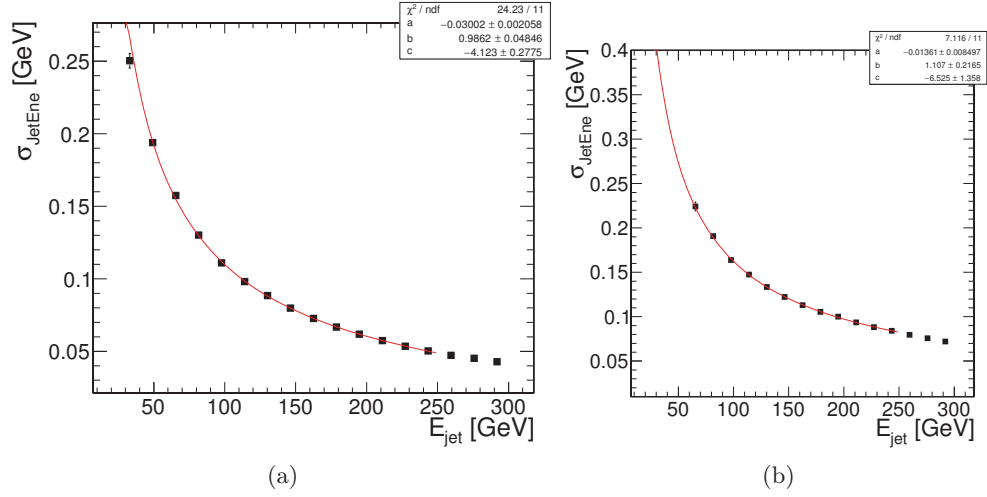


FIGURE 8.3 – Résolution en énergie des jets légers dans la zone  $1.4 < |\eta| < 1.7$  (a) et  $1.7 < |\eta| < 2.6$  (b).

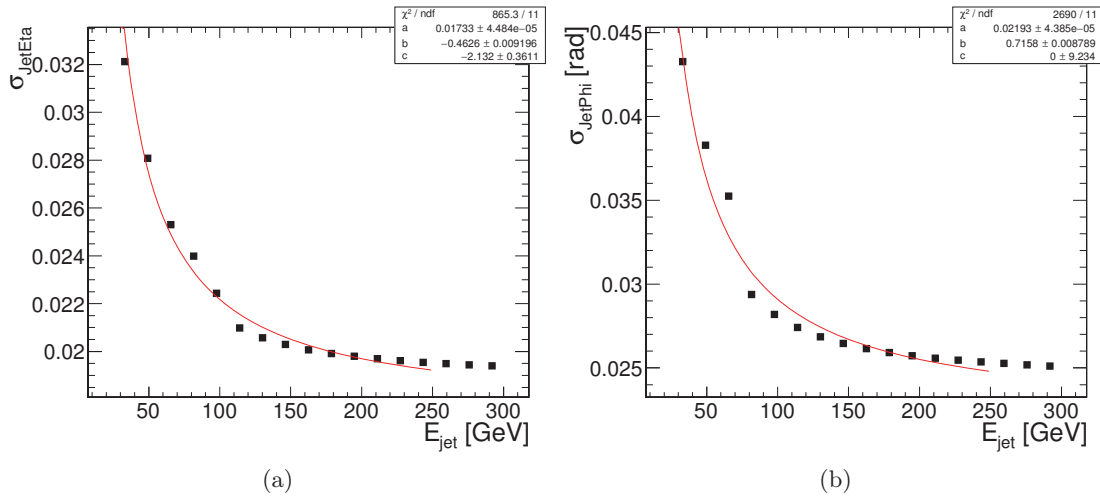


FIGURE 8.4 – Résolution en  $\eta$  (a) et  $\phi$  (b) des jets légers dans la zone  $|\eta| < 1.4$ .



proche car ils produisent également un top hadronique. Ces événements se distinguent par la présence d'un second neutrino. Ainsi la masse du quark *top* (et W) leptonique ne vérifie pas la contrainte dans le cas d'un événement  $t\bar{t}H$ .

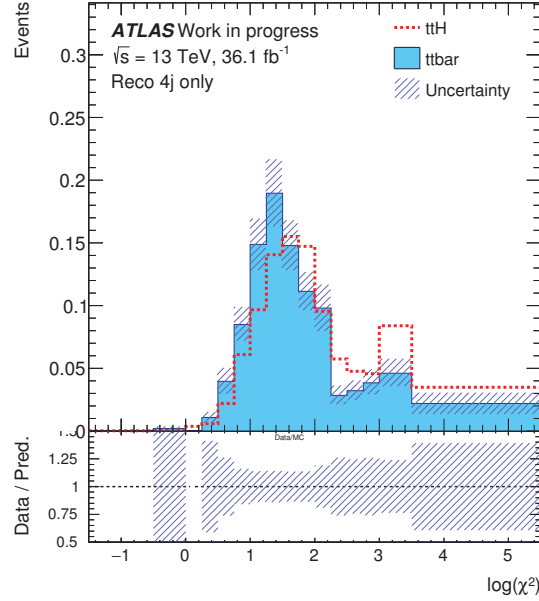


FIGURE 8.5 – Distribution du  $\text{Log}(\chi^2)$  obtenue à partir de la reconstruction de la paire  $t\bar{t}$  semileptonique pour des événements MC  $t\bar{t}$  et  $t\bar{t}H$ .

Un apport de l'ajustement cinématique est d'avoir une meilleure résolution sur les objets et donc sur d'autres grandeurs comme les masses invariantes des quark *top* et des bosons W ou la somme des impulsions transverses des objets de la combinaison ( $Ht$ ). Ainsi la variable  $Ht$  est représentée sur la Figure 8.6 pour les processus  $t\bar{t}$  et  $t\bar{t}H$ .

Cette variable, ainsi que d'autres variables telle que l'impulsion transverse totale du système ou d'un des quark *top*, montre un pouvoir discriminant pour améliorer la séparation entre le signal  $t\bar{t}H$  et le processus  $t\bar{t}$ .

La reconstruction du processus  $t\bar{t}$  permet aussi de vérifier le contenu d'une région enrichie en leptons non-prompts utilisée pour l'estimation de ce bruit de fond dans les régions de signal. Ainsi la Figure 8.7 présente la distribution du  $\chi^2$  de la reconstruction pour des prédictions et des données dans une région définie avec deux muons de même charge dont l'un est *antiTight* et une haute multiplicité en nombre de jets. Le bon accord entre données et prédictions indique que cette région contient majoritairement des leptons non-prompts provenant de processus  $t\bar{t}$  semi-leptonique tel que supposé.

## 8.5 Conclusion

La reconstruction d'un état final semileptonique pour le processus  $t\bar{t}$  par ajustement cinématique a été présentée dans ce chapitre. Le  $\chi^2$  et les variables obtenues par reconstruction ont un potentiel discriminant pour séparer le signal du processus  $t\bar{t}$  pour des événements avec deux leptons de même charges. Ces variables pourraient être considérées comme variables d'entrées pour des analyse multivariées. Ce n'est pas fait dans le cadre de l'analyse présentée au Chapitre 9 car d'autres variables corrélées avec la reconstruction comme le nombre de jet ou les angles des jets sont utilisées et l'impact des variables des reconstructions en est limité nécessitant des études complémentaires.

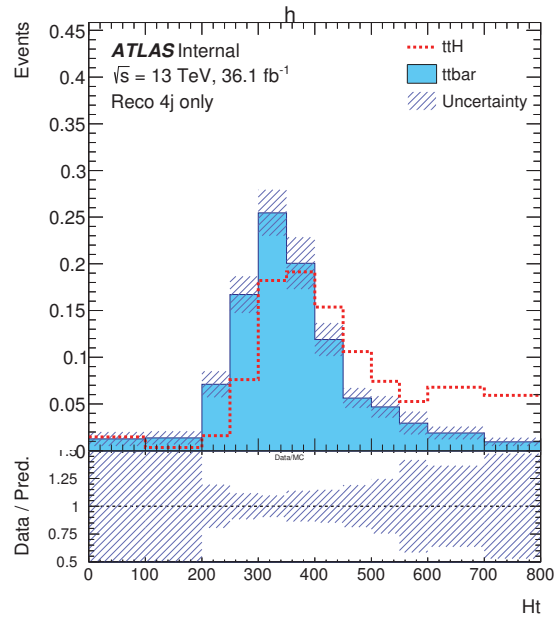


FIGURE 8.6 – Distribution de la variable  $H_t$  obtenue à partir de la reconstruction de la paire  $t\bar{t}$  semileptonique pour des événements MC  $t\bar{t}$  et  $t\bar{t}H$ .

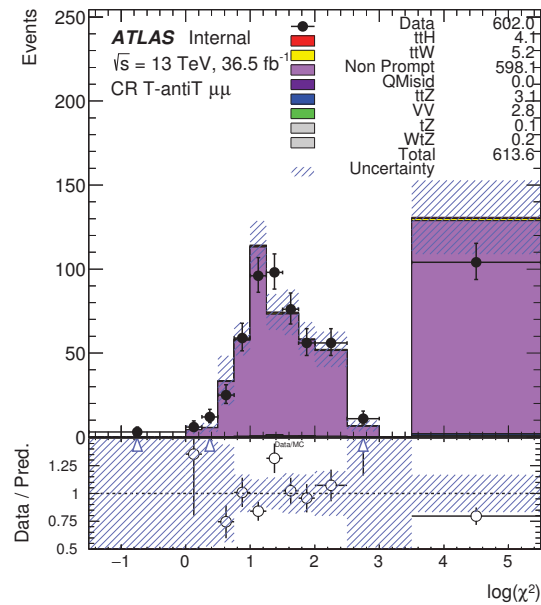


FIGURE 8.7 – Distribution de  $\log(\chi^2)$  obtenue à partir de la reconstruction de la paire  $t\bar{t}$  semileptonique dans une région avec deux muons de même charge, dont un est *antiTight* et au moins 5 jets.

La reconstruction par analyse multivariée représente une alternative permettant un gain en temps de calcul, l'utilisation de contraintes non gaussiennes et la prise en compte des corrélations mais en négligeant les résolutions associées aux objets.



# Extensions et améliorations de l'analyse

---

## 9.1 Introduction

Suite à l'analyse des données de 2015 et des premières données de 2016 pour une luminosité intégrée de  $13.2fb^{-1}$ , plusieurs changements ont été effectués pour l'analyse combinée des données des années 2015 et 2016 (luminosité intégrée de  $36.1fb^{-1}$ ). Ces améliorations sont listées dans les sections suivantes en commençant par l'ajout de nouveaux canaux, puis la redéfinition des objets physiques dans les sections 9.3 et 9.4. Par ailleurs les méthodes d'estimation des bruits de fond instrumentaux ont été adaptées à la nouvelle stratégie d'analyse et sont présentées dans la Section 9.5. Dans tous les canaux (sauf un décrit plus tard) des BDT sont utilisés pour l'extraction du signal et l'ajustement aux données. La définition des variables utilisées et les résultats de l'ajustement aux données sont présentés dans les Sections 9.6 et 9.8. Ainsi qu'il a été fait dans le Chapitre 7 une description plus détaillée de l'estimation des bruits de fond et des sélections est faite pour le canal avec deux leptons de même charge et aucun tau hadronique tandis qu'un résumé est présenté pour les autres canaux. Le travail que j'ai réalisé pour cette analyse a porté sur l'estimation du bruit de fond avec un lepton de charge mal reconstruite, la catégorisation dans le canal  $2\ell SS0\tau$  et la participation à la combinaison des différents canaux pour l'ajustement aux données final.

## 9.2 Motivation

La première analyse réalisée avec les données du Run 2 du LHC avait une philosophie très proche de celle utilisée pour l'analyse des données du Run 1 du LHC. Dans ce cadre, le simple comptage du nombre d'événements dans des catégories données était utilisé pour la robustesse de cette approche. Il était également difficile de contrôler la modélisation complète de variables discriminantes en particulier pour les bruits de fond instrumentaux. Avec la prise de données supplémentaires en 2016, le nombre d'événement est maintenant suffisant pour utiliser les distributions des variables tels le nombre de jets ou l'impulsion transverse des leptons dans des analyses multivariées. Ce changement de stratégie implique un certain nombre de changements dans le traitement des systématiques ou l'estimation des bruits de fond qui sont décrits dans les sections suivantes.

## 9.3 Ajout de nouveaux canaux

Suite à l'analyse réalisée avec  $13.2fb^{-1}$  de nouveaux canaux ont été ajoutés. Ces canaux supplémentaires contiennent tous des taus hadroniques :

- $3\ell 1\tau_{had}$  : canal avec un nombre d'événements attendus faible mais une bonne pureté en signal. Précédemment le canal  $3\ell$  ne contenait pas de comptage du nombre de taus hadroniques. Dorénavant deux canaux sont donc définis :  $3\ell 0\tau$  et  $3\ell 1\tau$ . Ce choix est réalisé pour améliorer la sensibilité de l'analyse au canal  $3\ell 1\tau$  par une sélection dédiée. Ce canal est le seul pour lequel un BDT n'est pas utilisé.
- $2\ell OS 1\tau_{had}$  : canal assez peu pur en signal mais présentant un nombre d'événements attendus important dû à la sélection de deux leptons légers de signes opposés.
- $1\ell 2\tau_{had}$  : canal assez peu pur en signal mais présentant un nombre d'événements attendus important.

Ces canaux scrutent une partie non couverte de l'espace des phases, excepté pour le canal  $3\ell 1\tau$ , ou permettent d'améliorer la sensibilité de l'analyse à la désintégration du boson de Higgs en leptons  $\tau$ .

## 9.4 Nouvelle définition des objets physiques et présélection

### 9.4.1 Système de déclenchement

Dans un premier temps la sélection basée sur le système de déclenchement a été modifiée pour conserver soit des événements avec au moins un lepton (*Single lepton trigger*) soit avec au moins deux leptons (*Dilepton trigger*). Dans ce deuxième cas, des leptons de plus basse impulsion transverse vont passer le système de déclenchement. La liste de l'ensemble des voies de système de déclenchement considérées selon l'année de la prise de donnée est décrite dans la Table 9.1.

	Single lepton triggers (2015)	Single lepton triggers (2016)
$\mu$	HLT_mu20_iloose_L1MU15, HLT_mu50	HLT_mu26_ivarmedium, HLT_mu50
$e$	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium, HLT_e120_1hloose	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0, HLT_e140_1hloose_nod0
	Dilepton triggers (2015)	Dilepton triggers (2016)
$\mu\mu$ (asymm.)	HLT_mu18_mu8noL1	HLT_mu22_mu8noL1
$ee$ (symm.)	HLT_2e12_1hloose_L12EM10VH	HLT_2e17_1hvloose_nod0
$e\mu, \mu e$ ( $\sim$ symm.)	HLT_e17_1hloose_mu14	HLT_e17_1hloose_nod0_mu14

TABLE 9.1 – Liste des systèmes de déclenchements utilisés selon l'année de la prise de donnée. Un événement doit passer l'un des critères *Single lepton* ou *dilepton*.

### 9.4.2 Objets physiques

La principale évolution dans la sélection des objets porte sur l'utilisation de BDT dans la définition des leptons définis avec des critères durs. Ces BDT sont décrits dans la Section 4, le premier portant sur la réjection des leptons non-prompts et le second sur celle des électrons ayant une mauvaise reconstruction de leur charge électrique. Les définitions des leptons définis avec des critères lâches (leptons *loose*) ou sévères (*tight*) sont données dans la Table 9.2.

TABLE 9.2 – Définitions des leptons *loose* et *tight*.

	Loose		Tight	
	$e$	$\mu$	$e$	$\mu$
BDT lepton non-prompt et isolation loose	-	-	Oui	Oui
Identification	Loose	Loose	Tight	Loose
BDT sur la charge	-	-	Oui	-
	( $2\ell SS0\tau, 2\ell SS1\tau, 3\ell 0\tau$ )			

Un autre changement des objets concerne la procédure de choix et d'exclusion des objets reconstruits proches. Les nouveaux critères utilisés sont décrits dans la Table 9.3. Dorénavant seules les distances entre les objets sont considérées. Le principal changement porte sur la réjection des muons proches d'un jet, la nouvelle sélection permettant de réduire la présence de muons provenant de désintégrations de hadrons beaux.

Conservé	Enlevé	Distance ( $\Delta R$ )
electron	electron (bas $p_T$ )	$< 0.1$
muon	electron	$< 0.1$
electron	jet	$< 0.3$
jet	muon	$\min(< 0.4, < 0.04 + 10[\text{GeV}]/p_T(\text{muon}))$
electron	tau	$< 0.2$
muon	tau	$0.2$
tau	jet	$< 0.3$

TABLE 9.3 – Critères utilisés pour la sélection des objets dans le cas de proximité topologique. La troisième colonne contient ces critères.

### 9.4.3 Présélection

Dans chacun des canaux une présélection est effectuée avant l'utilisation d'analyses multivariées pour extraire le signal. Ces présélections sont regroupées dans la Table 9.4. Les leptons sont ordonnés par impulsion transverse décroissante, le lepton 0 étant le lepton à plus haut  $p_T$ , excepté pour les canaux à 3 leptons où le lepton 0 est le lepton de charge électrique opposée à celle des deux autres. Les deux leptons de même charge sont alors rangés par impulsion transverse décroissante. Pour le canal  $3\ell 1\tau$ , la sélection décrite est la sélection finale car le nombre d'événements attendus est trop faible pour la mise en place d'une analyse multivariée. En comparaison avec l'analyse faite sur les données partielles de 2016, cette présélection est un peu plus lâche en terme de nombre de jets et d'impulsion transverse des leptons.

## 9.5 Estimation et validation des bruits de fond

Comme pour l'analyse précédente, les bruits de fond physiques sont décrits par simulation Monte-Carlo, tandis que les bruits de fonds instrumentaux sont estimés à partir des données. La Section 9.5.2 présente la méthode utilisée pour l'estimation des bruits de fond dans les canaux  $2\ell SS0\tau$  et  $3\ell 0\tau$ . L'estimation du nombre d'événements présentant un électron dont la charge électrique a été mal reconstruite repose sur la méthode par maximum de vraisemblance en 3 dimensions présentée dans le Chapitre 5. Les figures correspondant à la validation des bruits de fond sont présentées dans l'annexe ??

### 9.5.1 Estimation des événements avec un électron dont la charge électrique est mal reconstruite

Les taux de mauvaise identification de la charge des électrons sont obtenus en 3 dimensions, selon  $\eta$ ,  $p_T$  et la qualité du lepton. L'ajout de cette dernière permet d'estimer aussi la mauvaise reconstruction de la charge pour des événements *anti-tight* et *Loose* utilisée par la méthode de la matrice décrite après pour l'estimation des événements avec des leptons non-prompts. L'estimation du nombre d'événements dans les différentes régions est :

- canal  $2\ell SS0\tau$   $ee$  avec deux électrons *tight* :  $19.2 \pm 0.2$  (stat)  $\pm 6.1$  (syst).
- canal  $2\ell SS0\tau$   $e\mu$  avec deux électrons *tight* :  $14.7 \pm 0.2$  (stat)  $\pm 4.8$  (syst).
- canal  $2\ell SS0\tau$   $ee$  avec au moins un électrons *anti-tight* (utilisés par la méthode de la matrice) :  $1277.3 \pm 4.3$  (stat)  $\pm 105.2$  (syst).
- canal  $2\ell SS0\tau$   $e\mu$  avec au moins un électron *anti-tight* (utilisés par la méthode de la matrice) :  $747.2 \pm 4.3$  (stat)  $\pm 64.8$  (syst).
- canal  $2\ell SS1\tau$  :  $0.05 \pm 0.02$  (tot).

Région	Leptons	Taus Hadroniques	Jets	B-Tags
$2\ell SS0\tau$	==2 $\Sigma Q_{lep}=\pm 2$ $p_T^0 > 20$ GeV $p_T^1 > 20$ GeV	== 0	$\geq 4$	$\geq 1$ & $< 3$
	==3 $\Sigma Q_{lep}=\pm 1$ lep1 & lep2 : Tight, $p_T > 15$ GeV lep0 : Loose, $p_T > 10$ GeV $M_{ll,OS-SF} > 12$ GeV $ M_{ll} - 91.2$ GeV $  > 10$ GeV	== 0	$\geq 2$	$N_b \geq 1$
$4\ell$	==4 $p_T > 10$ GeV Lep2 et Lep3 tight $\Sigma Q_{lep}=0$ $ M_{ll,OS-SF} - 91.2$ GeV $  > 10$ GeV $M_{ll,OS-SF} > 12$ GeV $M_{4l}$ veto $\pm 5$ GeV autour $M_H = 125.0$ GeV	-	$\geq 2$	$\geq 1$
	==2 $\Sigma Q_{lep}=\pm 2$ $p_{T,lep}^{lead} > 15$ GeV $p_{T,lep}^{sub} > 15$ GeV $ M_{ee} - 91.2$ GeV $  > 10$ GeV	== 1 $Q_\tau = -Q_{lep}$	$\geq 4$	$\geq 1$
$1\ell 2\tau$	==1 $p_{T,lep} > 27$ GeV	== 1 $\Sigma Q_\tau = 0$	$\geq 3$	
$2\ell OS1\tau$	==2 $\Sigma Q_{lep}=\pm 0$ $p_{T,lep}^{lead} > 25$ GeV $p_{T,lep}^{sub} > 15$ GeV $ M_{ee} - 91.2$ GeV $  > 10$ GeV	== 1	$\geq 3$	$\geq 1$
	==3 $\Sigma Q_{lep}=\pm 1$ lep1 & lep2 : Tight, $p_T > 15$ GeV lep0 : Loose, $p_T > 10$ GeV $ M_{ll,OS-SF} - 91.2$ GeV $  > 10$ GeV $M_{ll,OS-SF} > 12$ GeV	== 1 $\Sigma Q_{lep+\tau}=\pm 0$	$\geq 2$	$N_b \geq 1$

TABLE 9.4 – Résumé des présélections effectuées dans les régions de signal. Il s'agit pour la région  $3\ell 1\tau$  de la sélection finale.



L'incertitude est de l'ordre de 30% dans les régions de signal et de 10% dans les régions utilisées par la méthode de la matrice. L'augmentation de l'incertitude est due à un nombre d'événements plus faible pour l'estimation des taux des électrons *tight* à cause du BDT utilisé pour rejeter la mauvaise identification de la charge. Toutefois cette augmentation est compensée par le fait que le nombre d'événements attendus avec une mauvaise reconstruction de la charge soit beaucoup plus faible qu'auparavant.

### 9.5.2 Estimation des événements avec des leptons non-prompts par méthode de la matrice

La méthode de la matrice est utilisée pour extraire le nombre d'événements contenant des leptons non-prompts. Elle permet également de prédire la distribution d'observables ce qui n'est pas possible avec la méthode présentée dans la section 7.5.2. Pour extraire ce nombre d'événements les leptons sont séparés en leptons *tight* et leptons *loose*, ces derniers contenant les leptons *tight* et les leptons *anti-tight* ne passant pas l'un des critères de sélection *tight*.

Dans un premier temps, le cas simple d'un événement à un seul lepton est étudié. Le nombre d'événements avec un lepton *tight*,  $N^T$ , ou un lepton *anti-tight*,  $N^{\mathcal{X}}$ , peut être exprimé en fonction de l'efficacité pour un lepton *loose* réel (prompt) ou faux (non-prompt) de passer la sélection *tight* :

$$*N^T = \varepsilon_r N^r + \varepsilon_f N^f \quad (9.1)$$

$$N^{\mathcal{X}} = \not\varepsilon_r N^r + \not\varepsilon_f N^f \quad (9.2)$$

où  $\varepsilon_r$  ( $\varepsilon_f$ ) est l'efficacité pour un lepton prompt (non-prompt) de passer la sélection *tight* et  $\not\varepsilon_r=(1-\varepsilon_r)$  ( $\not\varepsilon_f=(1-\varepsilon_f)$ ) la probabilité pour un lepton prompt (non-prompt) de passer la sélection *loose* mais pas la sélection *tight*. Les taux d'efficacités,  $\varepsilon$ , sont extraits de régions de contrôle à partir des données. Ces régions de contrôle sont définies à basse multiplicité en nombre de jets (2 ou 3 jets). L'hypothèse sous-jacente à l'utilisation de ces régions de contrôle est que les efficacités sont indépendantes du nombre de jets. Le choix des régions pour l'extraction des efficacités est tel que l'origine des leptons non-prompts est similaire dans la région de contrôle et la région de signal. En outre pour extraire l'efficacité  $r$ , des événements avec deux leptons de saveurs différentes de charges opposées sont utilisés, tandis que pour l'efficacité  $f$  des événements avec deux leptons de saveurs différentes et de même charge électrique sont sélectionnés. Dans les deux cas les leptons doivent passer la sélection *loose*. Les efficacités sont extraites selon différents bins en  $p_T$  et nombre de jets (pour les électrons) ou bins en  $p_T$  et distance entre le muon et le jet le plus proche (pour les muons).

La méthode décrite ci-après s'applique à des événements dileptoniques et est utilisée à la fois dans le cas du canal  $2\ell SS0\tau$  mais aussi du canal  $3\ell 0\tau$ . En effet pour ce dernier, la probabilité la plus importante est que ce soit l'un des deux leptons de même charge électrique qui soit un lepton non-prompt. Ces événements proviennent pour la plupart du processus  $t\bar{t}$  dont les produits de désintégrations comportent deux leptons de charges opposées avec un troisième lepton qui sera non-prompt.

Ainsi dans le cas de deux leptons, quatre régions peuvent être définies pour l'événement  $i$  :

- $TT_i$  : les deux leptons sont *tight* et le nombre d'événements de ce type est noté  $N^{TT}$ .
- $T\mathcal{X}_i$  : le lepton à plus grand  $p_T$  est *tight* et le second *anti-tight* et le nombre d'événements de ce type est noté  $N^{T\mathcal{X}}$ .
- $\mathcal{X}T_i$  : le lepton à plus grand  $p_T$  est *anti-tight* et le second *tight* et le nombre d'événements de ce type est noté  $N^{\mathcal{X}T}$ .
- $\mathcal{X}\mathcal{X}_i$  : les deux leptons sont *anti-tight* et le nombre d'événements de ce type est noté  $N^{\mathcal{X}\mathcal{X}}$ .

On peut définir des régions en fonction du caractère prompt ou non-prompt des leptons. Les nombres d'événements  $N^{rr}$ ,  $N^{rf}$ ,  $N^{fr}$  et  $N^{ff}$  correspondent respectivement à des événements avec deux leptons prompts, avec le lepton à plus grand  $p_T$  prompt et le second non-prompt, avec le lepton à

plus grand  $p_T$  non-prompt et le second prompt et finalement avec deux leptons non-prompts. Les nombres d'événements dans les différentes régions vont être mis en relation suivant l'équation :

$$\begin{pmatrix} N^{TT} \\ N^{T\cancel{X}} \\ N^{\cancel{X}T} \\ N^{\cancel{X}\cancel{X}} \end{pmatrix} = \begin{pmatrix} \varepsilon_{r,1}\varepsilon_{r,2} & \varepsilon_{r,1}\varepsilon_{f,2} & \varepsilon_{f,1}\varepsilon_{r,2} & \varepsilon_{f,1}\varepsilon_{f,2} \\ \varepsilon_{r,1}\cancel{\varepsilon}_{r,2} & \varepsilon_{r,1}\cancel{\varepsilon}_{f,2} & \varepsilon_{f,1}\cancel{\varepsilon}_{r,2} & \varepsilon_{f,1}\cancel{\varepsilon}_{f,2} \\ \cancel{\varepsilon}_{r,1}\varepsilon_{r,2} & \cancel{\varepsilon}_{r,1}\varepsilon_{f,2} & \cancel{\varepsilon}_{f,1}\varepsilon_{r,2} & \cancel{\varepsilon}_{f,1}\varepsilon_{f,2} \\ \cancel{\varepsilon}_{r,1}\cancel{\varepsilon}_{r,2} & \cancel{\varepsilon}_{r,1}\cancel{\varepsilon}_{f,2} & \cancel{\varepsilon}_{f,1}\cancel{\varepsilon}_{r,2} & \cancel{\varepsilon}_{f,1}\cancel{\varepsilon}_{f,2} \end{pmatrix} \begin{pmatrix} N^{rr} \\ N^{rf} \\ N^{fr} \\ N^{ff} \end{pmatrix} \quad (9.3)$$

L'inversion de la matrice  $4 \times 4$  donne l'expression du nombre d'événements  $N_{TT}^{rf}$ ,  $N_{TT}^{fr}$  et  $N_{TT}^{ff}$ . Le nombre d'événements avec un lepton non-prompt dans la région avec deux leptons *tight*  $N_{TT}^f$  est :

$$\begin{aligned} N_{TT}^f &= N_{TT}^{rf} + N_{TT}^{fr} + N_{TT}^{ff} \\ &= \varepsilon_{r,1}\varepsilon_{f,2}N^{rf} + \varepsilon_{f,1}\varepsilon_{r,2}N^{fr} + \varepsilon_{f,1}\varepsilon_{f,2}N^{ff} \\ &= (w_{TT}^{MM}N^{TT}) + (w_{T\cancel{X}}^{MM}N^{T\cancel{X}}) + (w_{\cancel{X}T}^{MM}N^{\cancel{X}T}) + (w_{\cancel{X}\cancel{X}}^{MM}N^{\cancel{X}\cancel{X}}) \end{aligned} \quad (9.4)$$

avec :

$$\begin{aligned} w_{TT}^{MM} &= (1 - \alpha\varepsilon_{r,1}\varepsilon_{r,2}\cancel{\varepsilon}_{f,1}\cancel{\varepsilon}_{f,2}) \\ w_{T\cancel{X}}^{MM} &= (\alpha\varepsilon_{r,1}\varepsilon_{r,2}\varepsilon_{f,1}\cancel{\varepsilon}_{f,2}) \\ w_{\cancel{X}T}^{MM} &= (\alpha\varepsilon_{r,1}\varepsilon_{r,2}\cancel{\varepsilon}_{f,1}\varepsilon_{f,2}) \\ w_{\cancel{X}\cancel{X}}^{MM} &= -(\alpha\varepsilon_{r,1}\varepsilon_{r,2}\varepsilon_{f,1}\varepsilon_{f,2}) \\ \alpha &= \frac{1}{(\varepsilon_{r,1} - \varepsilon_{f,1})(\varepsilon_{r,2} - \varepsilon_{f,2})} \end{aligned} \quad (9.5)$$

Un élément qui n'est pas présent dans les équations précédentes est que dans les régions utilisées pour l'extraction de la valeur final, tous les événements doivent présenter au moins un lepton non-prompt. Or ces nombres d'événements sont obtenus par comptage dans des régions de données après soustraction de la contamination en leptons prompts dans les données. Cette soustraction est faite à partir de la simulation Monte-Carlo pour les processus physiques ( $t\bar{t}V$ , Dibosons ...) et à partir de l'estimation par maximum de vraisemblance en 3 dimensions pour les événements avec un électron de charge électrique mal reconstruite.

Les efficacités sont paramétrées en deux dimensions ( $p_T$  et nombre de jets étiquetés  $b$  pour les électrons,  $p_T$  et la distance entre le lepton et le jet le plus proche pour les muons). Par ailleurs dans le cas des électrons un élément supplémentaire est la présence d'électrons non-prompt provenant de conversion de photon. La proportion d'électrons non-prompt dus à cette effet n'est pas la même dans la région à basse multiplicité en jets et celle à haute multiplicité. Un facteur multiplicatif obtenu à partir de la simulation MC est appliqué sur les efficacités des électrons pour corriger cet effet.

Plusieurs incertitudes sont estimées lors de l'extraction du nombre d'événements avec des leptons non-prompts par la méthode de la matrice. Ces incertitudes sont :

- les incertitudes statistiques liées au nombre d'événements dans les régions de contrôle définies pour l'extraction des efficacités.
- les incertitudes liées à la soustraction des bruits de fond prompts. Ces incertitudes sont anticorrélées avec les incertitudes mises sur ces bruits de fond dans la région de signal.
- une incertitude liée au test de fermeture comparant le nombre d'événements avec des leptons non-prompts prédits par la méthode de la matrice sur une simulation Monte-Carlo  $t\bar{t}$  et le nombre d'événements vrais de la simulation ayant des leptons non-prompts. Les distributions de la variable de sortie du BDT utilisée dans le canal  $2\ell SS0\tau$  est représentée dans le cadre de ce test de fermeture sur la Figure 9.1.

- les incertitudes statistiques liées au nombre d'événements dans les régions sur lesquelles les efficacités sont appliquées pour extraire le nombre d'événements de la région de signal.
- les incertitudes liées au facteur supplémentaire utilisé pour prendre en compte la proportion de conversion de photons.

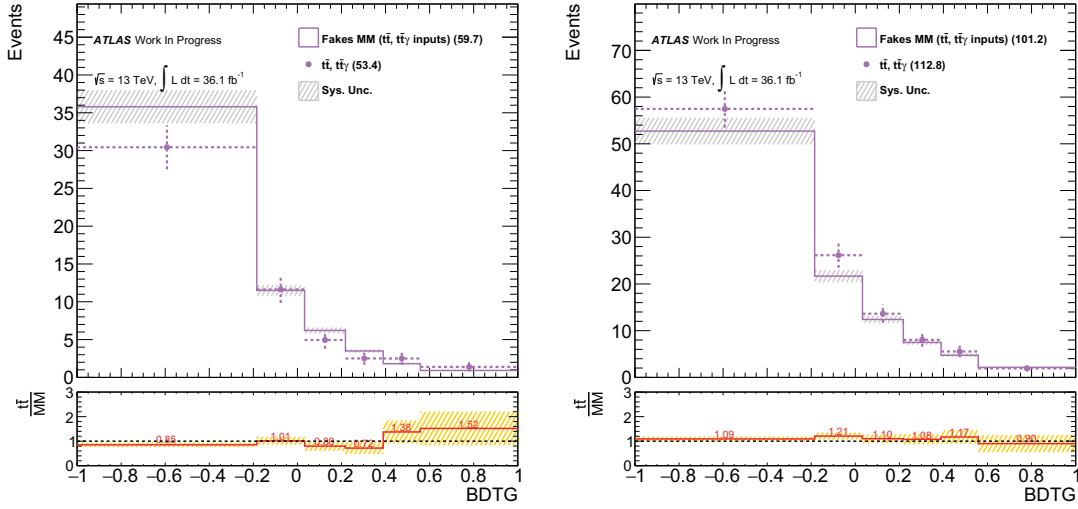
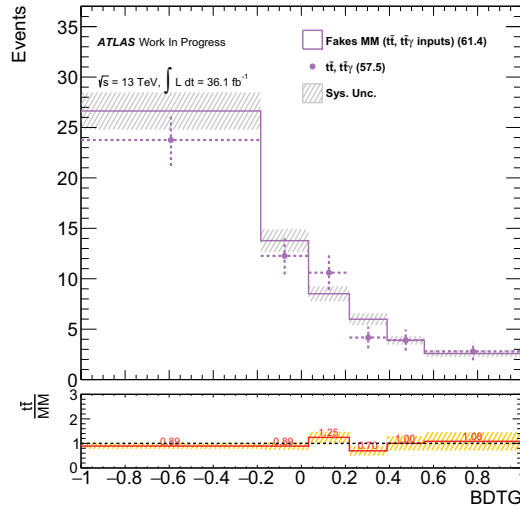
(a) *BDTG, ee*(b) *BDTG, OF*(c) *BDTG, mu mu*

FIGURE 9.1 – Distribution de la variable de sortie du BDT utilisée dans le canal  $2\ell SS0\tau$  pour la simulation MC  $t\bar{t}$  (point) et la prédiction faite avec la méthode de la matrice (trait plein) pour des événements avec au moins 4 jets et 1 jet étiqueté  $b$ , dans le cas de deux électrons (a), un électron et un muon (b) et deux muons (c).

Une fois l'estimation du nombre d'événements réalisée, une comparaison est faite entre les prédictions et les données dans une région riche en leptons non-prompts qui correspond à une région avec deux leptons *tight* et 2 ou 3 jets dont au moins un étiqueté  $b$ . La Figure 9.2 présente la distribution de l'impulsion transverse du lepton 0 (lepton à plus haute impulsion transverse) dans cette région.

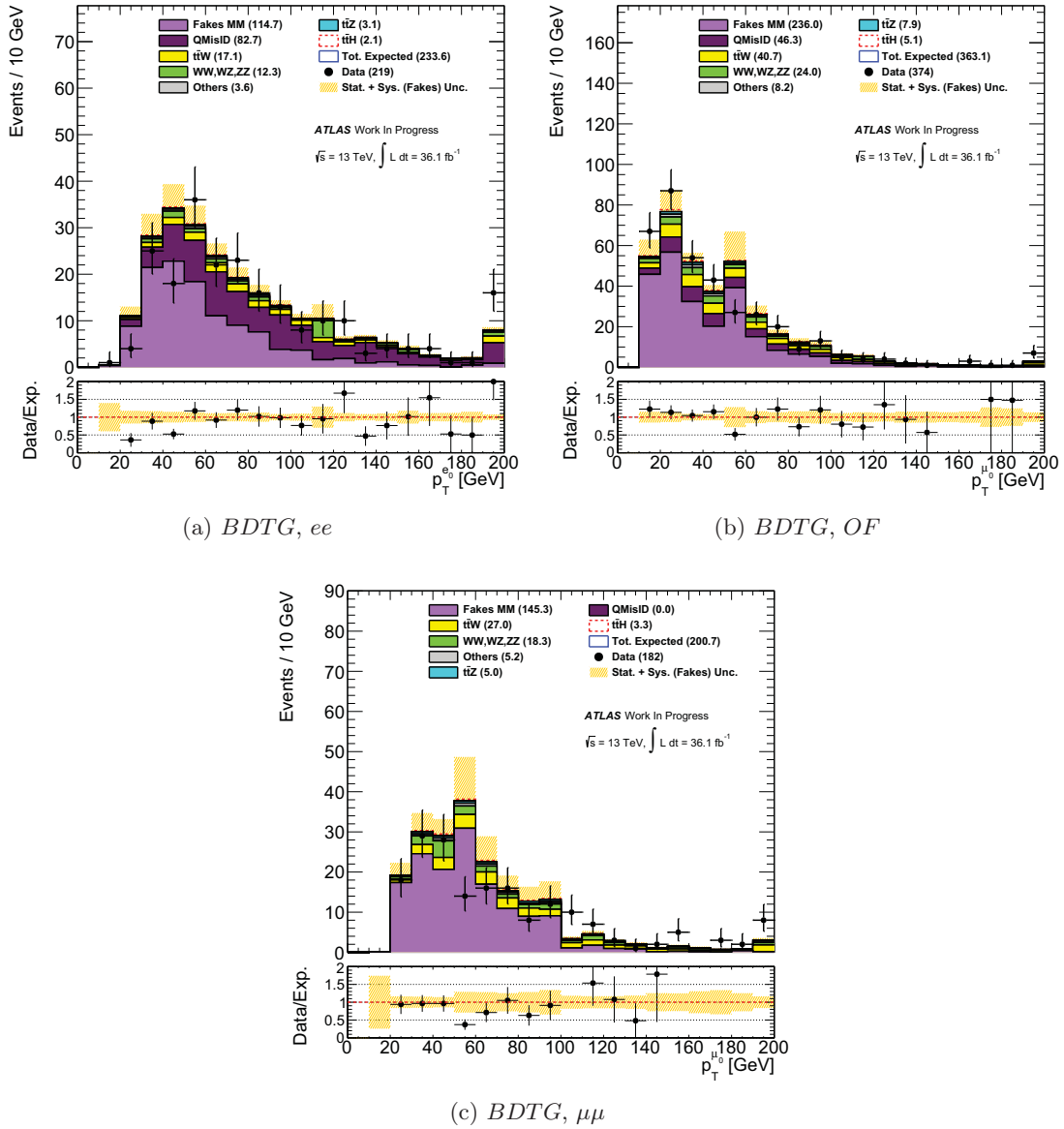


FIGURE 9.2 – Distribution du nombre d'événements prédits et observés selon le  $p_T$  du lepton à plus haut  $p_T$  pour des événements avec 2 ou 3 jets et 1 jet étiqueté  $b$ , dans le cas de deux électrons (a), un électron et un muon (b) et deux muons (c).

## 9.6 Sélection des événements

Des BDT sont utilisés dans tous les canaux, excepté pour le canal  $3\ell 1\tau$ , et donc aucune sélection n'est faite outre la présélection décrite dans la section 9.4.3. Selon les canaux, différentes variables sont utilisées pour construire les BDT. Ces variables sont toutes listées dans la Table 9.5.

### 9.6.1 Analyse multivariée dans le canal $2\ell SS0\tau$

Dans le cas du canal  $2\ell SS0\tau$ , 9 variables sont utilisées pour construire deux BDT, le premier ( $BDTG_{tt\bar{b}ar}$ ) pour séparer les événements  $t\bar{t}$  de  $t\bar{t}H$  et le second ( $BDTG_{ttW}$ ) pour séparer le processus  $t\bar{t}W$  de  $t\bar{t}H$ . Les variables choisies sont donc discriminantes vis-à-vis de l'un des de ces deux bruits de fond. Ainsi les distances entre les jets et les leptons et le nombre de jets sont parmi les variables les plus discriminantes pour les deux BDT. Dans le cas de la réjection du bruit de fond  $t\bar{t}$ , l'impulsion transverse du second lepton joue aussi un rôle important tandis que la masse invariante dilepton est discriminante pour le processus  $t\bar{t}W$ . Une sélection supplémentaire est ajoutée en rejetant les événements avec 3 jets étiquetés  $b$ . Ce choix est fait car le processus  $t\bar{t}H$  dans l'état final  $2\ell SS0\tau$  ne devrait pas avoir plus de 2 jets étiquetés  $b$ , et que cette région est assez mal modélisée. Le fait d'enlever ces événements ne modifie pas la sensibilité attendue de l'analyse. Le pouvoir discriminant de  $BDTG_{tt\bar{b}ar}$  est meilleur que celui de  $BDTG_{ttW}$  à cause des similarités du signal et du processus  $t\bar{t}W$ <sup>1</sup>. Les résultats de ces deux BDTs individuels sont combinés pour construire une variable globale  $BDTG = (BDTG_{tt\bar{b}ar} + BDTG_{ttW})/2$  dans six bins définis tel que chacun contient la même quantité de signal. Les distributions des 9 variables d'entrée et celles obtenues pour les BDT individuels et combiné sont présentées sur les Figures 9.3 et 9.4. L'ajustement aux données pour extraire le résultat final sera réalisée sur cette variable combinant les deux BDTs visant le  $t\bar{t}$  et  $t\bar{t}W$ .

### 9.6.2 Catégorisation des événements dans la région $2\ell SS0\tau$

Une méthode simple basée sur la catégorisation des événements est utilisée comme vérification mais aussi pour tester la sensibilité atteignable par l'analyse sans passer par des analyses multivariées. La méthode utilisée est similaire à ce qui a été fait pour l'analyse avec  $13.2fb^{-1}$  de données. Le point de départ est la présélection effectuée dans le cadre de l'analyse multivariée. La catégorisation se fait selon les variables discriminantes que sont :

- le nombre de jets : soit exactement 4 jets, soit plus de 5 jets.
- le nombre de b-jets : soit 1 jet étiqueté  $b$ , soit deux jets étiquetés  $b$  (les événements avec 3 jets étiquetés  $b$  ou plus sont enlevés).
- la saveur des leptons :  $ee$ ,  $e\mu$ ,  $\mu\mu$ .

Ainsi 12 catégories sont définies, les plus sensibles au signal étant les catégories avec au moins 5 jets car le bruit de fond non-prompt est particulièrement important dans la région 4 jets. Pour la même raison, les régions avec 2 jets étiquetés  $b$  sont plus sensibles. Enfin les catégories avec deux muons sont les plus sensibles car ne sont pas contaminées par les événements avec un électron dont la charge est mal reconstruite ou les leptons non-prompts provenant de conversions de photons. Le nombre d'événements dans chacune des catégories est donné dans la Table 9.6 tandis que la fraction des différents bruits de fond dans chacune des catégories est présentée sur la Figure 9.5.

### 9.6.3 Résumé des régions de signal

Des analyses multivariées sont utilisées dans 6 des 7 états finals de l'analyse tandis que des catégorisations sont faites en tant que vérification dans les canaux  $2\ell SS0\tau$  et  $3\ell 0\tau$ . Ainsi parmi les

1. Pour l'état final  $2\ell SS0\tau$ , le signal contient au final 2 quarks  $top$  et deux bosons W provenant de la désintégration du boson de Higgs quand le processus  $t\bar{t}W$  a aussi deux quarks  $top$  mais seulement 1 boson W.

TABLE 9.5 – Variables utilisées dans les analyses multivariées des différents canaux.

Variable	$2\ell SS0\tau$	$3\ell 0\tau$	$4\ell$	$1\ell 2\tau$	$2\ell SS1\tau$	$2\ell OS1\tau$
$p_T$ du lepton avec le plus haut $p_T$		×				
$p_T$ du lepton avec le second plus haut $p_T$	×	×			×	
$p_T$ du lepton avec le troisième plus haut $p_T$		×				
Masse invariante dilepton	×	×				×
Masse invariante de trois leptons		×				
Masse invariante de quatre leptons			×			
Masse invariante dilepton la plus proche de celle du Z			×			
Autre masse invariante dilepton de candidat Z			×			
Somme scalaire des $p_T$ des leptons			×			×
Isolation de la trace du lepton avec le second plus haut $p_T$					×	
$ \eta_\ell $ maximum entre deux leptons	×				×	
Saveur des leptons	×	×				
Charge des leptons		×				
Nombre de jets	×	×		×	×	×
Nombre de jets étiquetés $b$	×	×		×	×	×
$p_T$ du jet avec le plus haut $p_T$						×
$p_T$ du jet avec le second plus haut $p_T$		×			×	
$p_T$ du b-jet avec le plus haut $p_T$		×				
Somme scalaire du $p_T$ de tous les jets		×		×	×	×
Somme scalaire du $p_T$ de tous les b-jets						×
Poids lié à l'étiquetage $b$ du jet avec le second $p_T$		×			×	
Booléen : jet à plus haut $p_T$ avec le plus haut poids de b-tagging		×				
Booléen : jet à plus haut $p_T$ étiqueté $b$		×				
Booléen : jet à plus haut $p_T$ pas étiqueté $b$		×				
Poids lié à l'étiquetage $b$ du jet avec le troisième $p_T$					×	
$ \eta $ du jet avec le quatrième $p_T$					×	
$p_T$ du $\tau_{had}$ avec le premier $p_T$				×		×
$p_T$ du $\tau_{had}$ avec le second $p_T$				×		
Masse invariante di-tau				×		
Masse invariante entre le $\tau_{had}$ et le lepton le plus éloigné					×	
Distance lepton0-lepton1		×				
Distance lepton0-lepton2		×				
Distance lepton0-jet le plus proche	×	×				
Distance lepton0-jet à plus haut $p_T$		×			×	
Distance lepton0-b-jet le plus proche		×				
Distance lepton1-jet le plus proche	×	×				
Distance lepton2-jet le plus proche		×				
Distance lepton-jet la plus faible		×				×
Distance lepton-jet étiqueté $b$ la plus faible						×
Distance jet non étiqueté b-jet étiqueté $b$ la plus faible						×
Distance lepton0- $\tau_{had}$						×
Distance lepton1- $\tau_{had}$						×
Distance la plus faible entre les jets				×		
Distance entre les deux jets à plus haut $p_T$					×	
Énergie transverse manquante (MET)	×		×			
Séparation azimuthale entre jet à plus haut $p_T$ -MET		×				
Masse invariante des leptons-MET			×			
$\mathcal{P.M.E.}$			×			

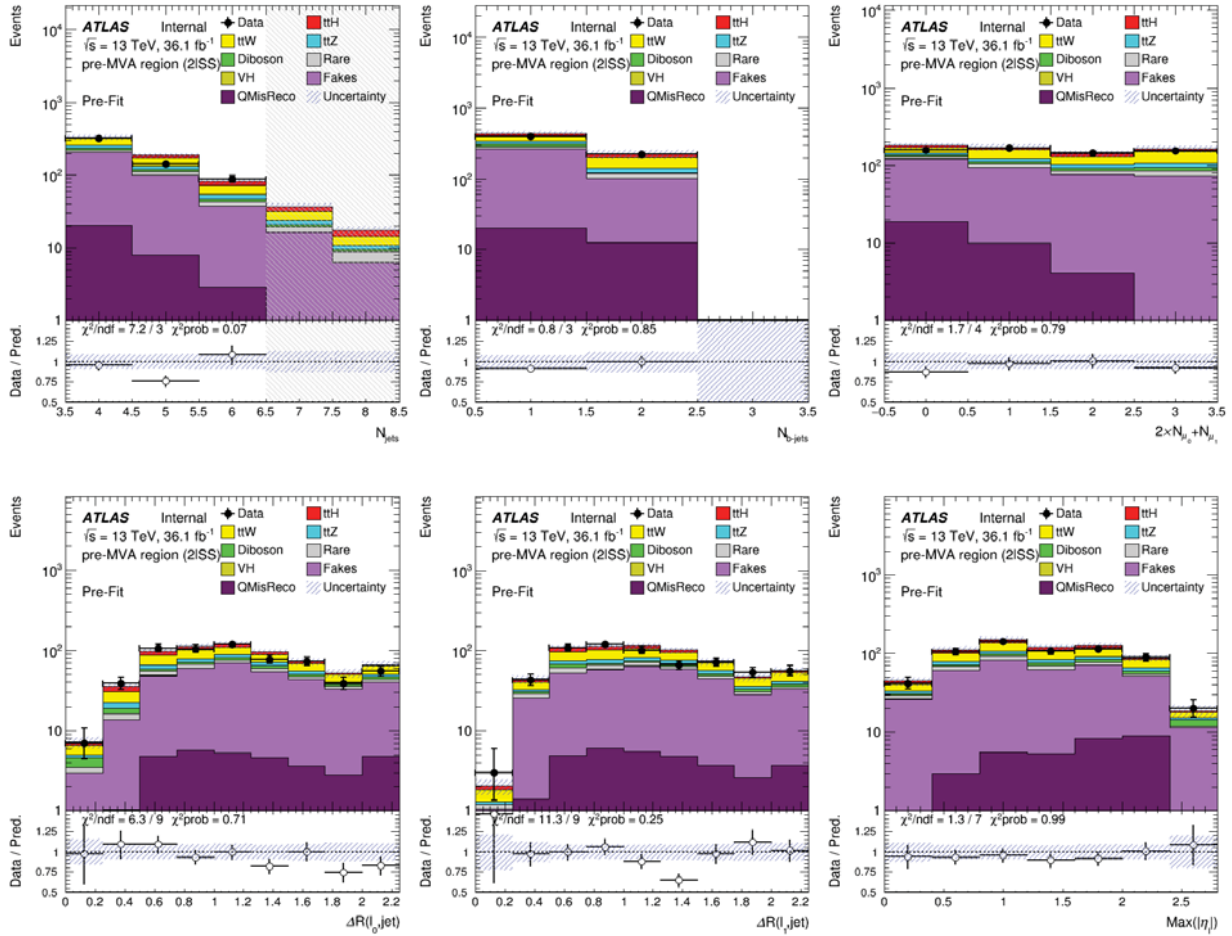


FIGURE 9.3 – Distributions des 6 premières variables utilisées dans les BDT du canal  $2\ell SS0\tau$ . Les données sont masquées dans les bins pour lesquels le signal est supérieur à 15% du bruit de fond. Les bruits de fond non-prompt et QMisReco (avec une mauvaise reconstruction de la charge d'un électron) sont obtenus par les méthodes d'estimation décrites précédemment.

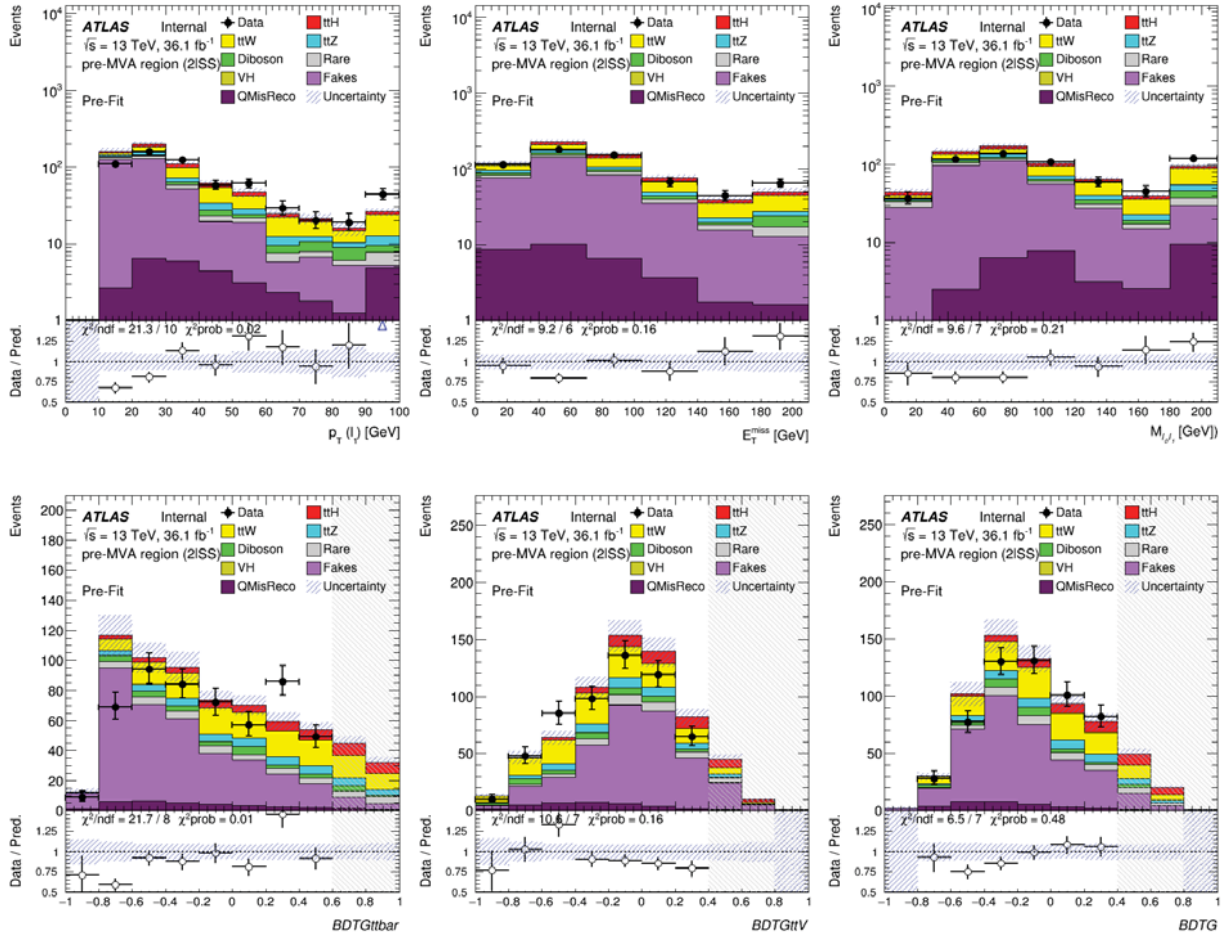


FIGURE 9.4 – Distributions des 3 dernières variables (Première ligne) utilisées dans les BDT du canal  $2\ell SS0\tau$  et des variables obtenus à partir de ces BDT (dernière ligne) pour les prédictions et les données. Les données sont masquées dans les bins pour lesquels le signal est supérieur à 15% du bruit de fond. Les bruits de fond non-prompt et QMisReco (avec une mauvaise reconstruction de la charge d'un électron) sont obtenus par les méthodes d'estimation décrites précédemment.



	$\mu\mu + 1b5j$	$e\mu + 1b5j$	$ee + 1b5j$	$\mu\mu + 2b5j$	$e\mu + 2b5j$	$ee + 2b5j$
MC BG	$20.68 \pm 4.40$	$32.28 \pm 6.05$	$13.51 \pm 3.27$	$16.01 \pm 2.36$	$25.02 \pm 3.51$	$9.28 \pm 1.31$
DD BG	$21.15 \pm 7.50$	$28.48 \pm 6.70$	$21.78 \pm 5.11$	$3.38 \pm 2.05$	$22.86 \pm 4.99$	$18.35 \pm 4.71$
Total BG	$41.83 \pm 7.56$	$60.76 \pm 6.94$	$35.29 \pm 5.14$	$19.39 \pm 2.98$	$47.88 \pm 5.20$	$27.62 \pm 4.58$
ttH	$5.30 \pm 0.68$	$7.80 \pm 1.02$	$3.23 \pm 0.48$	$4.09 \pm 0.54$	$6.75 \pm 1.66$	$2.65 \pm 0.42$
Données	38	79	39	28	44	25
	$\mu\mu + 1b4j$	$e\mu + 1b4j$	$ee + 1b4j$	$\mu\mu + 2b4j$	$e\mu + 2b4j$	$ee + 2b4j$
MC BG	$19.39 \pm 3.57$	$31.47 \pm 6.28$	$12.27 \pm 2.35$	$11.90 \pm 1.91$	$19.07 \pm 3.19$	$7.50 \pm 1.32$
DD BG	$20.53 \pm 6.39$	$54.77 \pm 14.08$	$32.12 \pm 7.21$	$1.60 \pm 0.99$	$18.08 \pm 3.66$	$22.41 \pm 5.44$
Total BG	$39.92 \pm 5.86$	$86.23 \pm 12.71$	$44.39 \pm 6.85$	$13.50 \pm 2.04$	$37.16 \pm 4.16$	$29.91 \pm 5.29$
ttH	$2.57 \pm 0.58$	$3.76 \pm 0.41$	$1.50 \pm 0.42$	$1.43 \pm 0.31$	$2.39 \pm 0.39$	$1.09 \pm 0.20$
Données	37	77	49	20	58	20

TABLE 9.6 – Nombre d'événements dans les différentes catégories définies pour l'état final  $2\ell SS0\tau$ . Le signal est décomposé par désintégrations du boson de Higgs tandis que les bruits de fond sont regroupés selon qu'ils sont obtenus par simulation MC (bruits de fond irréductibles) ou à partir des données (bruits de fond instrumentaux). Les incertitudes présentées contiennent toutes les incertitudes théoriques et systématiques.

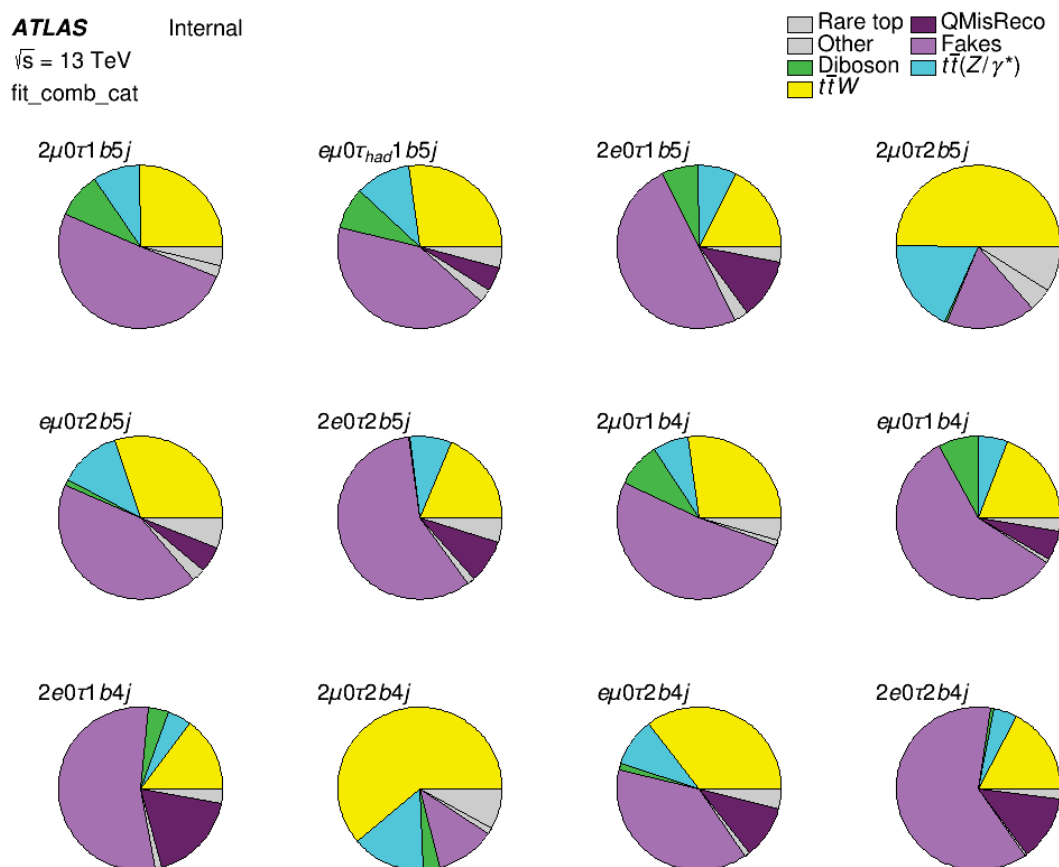


FIGURE 9.5 – Part des différents bruits de fond dans chacune des 12 catégories définies pour l'état final  $2\ell SS0\tau$ .

différents canaux, les régions et variables finales considérées pour l'ajustement aux données sont :

- $2\ell SS0\tau$  : la variable obtenue par la combinaison des deux BDTs visant à discriminer le  $t\bar{t}$  et

le  $t\bar{t}W$  du  $t\bar{t}H$ .

- $3\ell 0\tau$  : 5 catégories enrichies en  $t\bar{t}H$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $t\bar{t}\text{et}$  diboson. Pour chaque catégorie une variable correspondant au résultat d'un BDT est utilisé pour l'ajustement.
- $4\ell$  : 2 catégories dont l'une enrichie en  $t\bar{t}Z$  et l'autre où les événements avec deux leptons de même saveur et une masse proche du boson  $Z$  sont rejetés. Dans la première catégorie un BDT est utilisée pour séparer le signal du  $t\bar{t}Z$ , après une sélection sur la variable de sortie du BDT, le nombre d'événements est compté. Dans la seconde, pour laquelle très peu d'événements sont attendus, un comptage des événements est fait.
- $1\ell 2\tau$  : la variable de sortie d'un BDT ayant pour but de séparer le signal du  $t\bar{t}$  est séparée en deux régions avec comme critère que la variable extraite du BDT  $< 0.5$ . La région à haut BDT est utilisée comme région de signal, tandis que la région à bas BDT est utilisée pour extraire la proportion de jets reconstruit comme des  $\tau$ .
- $2\ell SS1\tau$  : la variable de sortie d'un BDT ayant pour but de séparer le signal du  $t\bar{t}$  est évaluée dans une région plus pure que la région de présélection avec au moins 4jets dont 1 est étiqueté  $b$ .
- $2\ell OS1\tau$  : la variable de sortie d'un BDT ayant pour but de séparer le signal du  $t\bar{t}$ .
- $3\ell 1\tau$  : le comptage des événements dans la région de signal définie dans la section 9.4.3.

Les résumés des sélections et de la stratégie utilisées pour chaque canal sont données dans la Table 9.7.

	Canaux sans $\tau_{had}$				Canaux avec $\tau_{had}$		
	$2\ell SS0\tau$	$3\ell 0\tau$	$4\ell$	$1\ell 2\tau$	$2\ell SS1\tau$	$2\ell OS1\tau$	$3\ell 1\tau$
Lepton léger	2T*	1L*, 2T*	2L, 2T	1T	2T*	2L <sup>†</sup>	1L <sup>†</sup> , 2T
$\tau_{had}$	0M	0M	-	1T, 1M	1M	1M	1M
$N_{jets}, N_{b-jets}$	$\geq 4, = 1, 2$	$\geq 2, \geq 1$	$\geq 2, \geq 1$	$\geq 3, \geq 1$	$\geq 4, \geq 1$	$\geq 3, \geq 1$	$\geq 2, \geq 1$
Bruit de fond	DD	DD	pseudo-DD	MC	DD	MC	MC
lepton non-prompt	(MM)	(MM)	(SF)	-	(FF)	-	-
Bruit de fond	-	-	-	DD	pseudo-DD	DD	pseudo-DD
tau fake	-	-	-	(SS data)	(SF)	(FF)	(SF)
BDT entraîné vs	Fakes et $t\bar{t}V$	$t\bar{t}, t\bar{t}W, t\bar{t}Z, VV$	$t\bar{t}Z / -$	$t\bar{t}$	tous	$t\bar{t}$	-
Discriminant	2×1D BDT	5D BDT	Comptage	BDT	BDT	BDT	Comptage
	shape	shape	des evts	shape	shape	shape	des evts
Régions de signal	1	1	1 / 1	1	1	1	1
Nombre de bins	6	5	1 / 1	2	2	10	1
Régions de contrôle	-	4	-	-	-	-	-
(1 bin pour chaque)							

TABLE 9.7 – Résumé des critères de présélection et des stratégies appliquées dans les différents canaux. pour la sélection des leptons,  $T$  signifie sélection *Tight* définie dans la Table 9.2 sans la sélection sur le BDT QMisId,  $T^*$  est la même définition avec la sélection sur le BDT QMisId,  $L$  correspond à la définition *Loose*,  $L^\dagger$  correspond à la définition *Loose* avec un critère d'isolation, et  $L^*$  correspond à la définition *Loose* avec le critère d'isolation *Loose*. Pour les leptons non-prompts et les  $\tau_{had}$  fakes  $DD$  signifie dérivé à partir des données, avec  $MM$  pour la méthode de la matrice et  $FF$  pour une méthode *Fake factor*.

## 9.7 Incertitudes systématiques

Avec l'utilisation des analyses multivariées dans cette analyse, la modélisation des variables utilisées est à prendre en compte. Les incertitudes instrumentales décrites pour l'analyse précédente couvrent cette composante qui n'était pas utilisée jusque là. L'estimation des incertitudes

théoriques portant sur la distribution des variables utilisées est faite en comparant ces variables dans la simulation utilisée avec la distribution des mêmes variables obtenues avec des variations de générateurs. Ainsi la Table 9.8 regroupent les différentes incertitudes en précisant le nombre de composants par type d'incertitudes et si elles couvrent la forme de la distribution (S pour *Shape*) ou la normalisation du processus (N).

Les incertitudes instrumentales sont les mêmes que celles présentes dans l'analyse du Chapitre 7. La différence porte sur les incertitudes théoriques puisque de nouvelles incertitudes ont été ajoutées pour prendre en compte la modélisation des processus  $t\bar{t}H$  et  $t\bar{t}V$  selon les variables discriminantes.

Les incertitudes théoriques et instrumentales sont corrélées entre tous les canaux excepté pour une incertitude sur l'énergie des jets<sup>2</sup> qui est séparée selon les canaux. Pour les bruits de fond réductibles, les incertitudes sont corrélées entre  $2\ell SS0\tau$  et  $3\ell0\tau$  pour l'estimation des leptons non-prompts de même qu'entre  $2\ell SS0\tau$  et  $2\ell SS1\tau$  pour l'estimation des événements avec une mauvaise reconstruction de la charge. Par ailleurs certaines incertitudes sur l'estimation des  $\tau_{had}$  *fakes*<sup>3</sup> sont corrélées pour les canaux  $2\ell SS1\tau$ ,  $2\ell OS1\tau$  et  $3\ell1\tau$ . Finalement les incertitudes sur les bruits de fond réductibles pour les canaux  $4\ell$  et  $1\ell2\tau$  sont décorréliées de celles des autres canaux.

Les tables présentes dans l'annexe B présentent les noms des systématiques instrumentales ainsi que leur méthodes d'applications.

## 9.8 Résultats

### 9.8.1 Nombres d'événements attendus et observés dans les régions de signal

Le nombre d'événements dans chaque canal est présenté dans la table 9.9. Au final ces nombres ne seront pas utilisés directement mais séparés dans des bins pour plusieurs régions. La sensibilité de l'analyse ne peut donc être déduite de cette table. La Fig. 9.6 présente la fraction de bruit de fond dans chaque canal.

### 9.8.2 Ajustement aux données

Après la définition des régions de signal et des variables d'intérêt, un premier ajustement aux données est réalisé sur des données Asimov pour extraire l'incertitude sur la force du signal et la signification statistique. Ces deux grandeurs sont données pour chaque canal dans la Table 9.10 dans le cas où seul la normalisation du processus  $t\bar{t}H$  est laissée flottante.

Un ajustement aux données est aussi réalisé pour la catégorisation dans le canal  $2\ell SS0\tau$  et la force du signal obtenue est de  $\mu = 1.00 \pm 0.82$  pour une signification statistique de  $1.2\sigma$ .

Un ajustement est ensuite réalisé sur les données observées et le résultat final sur la force du signal pour chaque canal et la combinaison est présentée dans la table 9.11. La signification statistique finale est de  $4.1\sigma$  pour une force du signal observée de  $1.6_{-0.3}^{+0.3}$ (stat.)  $_{-0.3}^{+0.4}$ (syst.). Ce résultat est le premier exemple d'observation<sup>4</sup> du processus  $t\bar{t}H$  au LHC.

Les Figures 9.7 et 9.8 présentent les comparaisons entre les données et les prédictions pour les variables utilisées lors de l'ajustement aux données. Les bruits de fond et le signal sont corrigés par la force du signal observée et les variations observées sur les systématiques lors de l'ajustement aux

2. Impact des jets légers sur l'échelle d'énergie des jets

3. Ce terme désigne des objets reconstruits comme des  $\tau_{had}$  mais qui n'en sont pas et peuvent être des jets légers ou des leptons.

4. Une signification statistique de  $3\sigma$  est considérée nécessaire pour parler d'observation tandis que la découverte de nouveaux processus ou particules doit correspondre à une signification statistique de  $5\sigma$ .

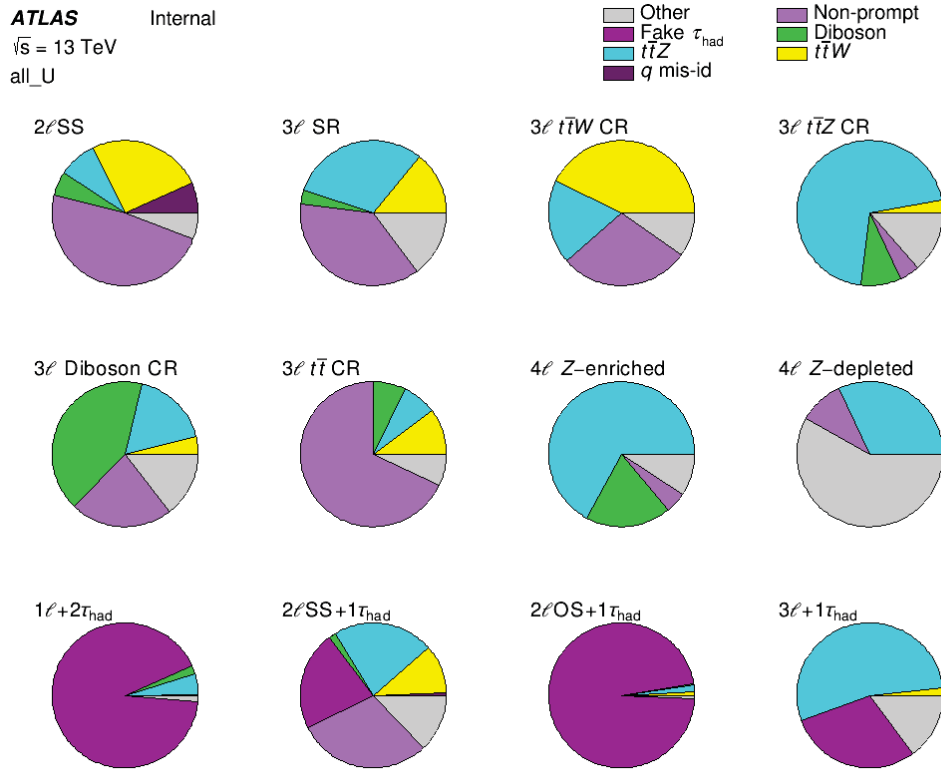


FIGURE 9.6 – Proportion de chaque bruit de fond dans les douze régions de l'analyse.

données. L'accord final entre les données et les prédictions corrigées par l'application de la force du signal est bon.

Les Figures 9.9 et 9.10 présentent le classement des incertitudes dans les différents canaux selon leur impact sur la force du signal pour chaque canal individuellement. La figure 9.11 présente ce même classement pour la combinaison des canaux sans  $\tau_{had}$ , avec  $\tau_{had}$  et finalement pour la combinaison complète. Finalement la Table 9.12 présente l'impact des systématiques regroupés par groupe sur la précision de la force du signal. Les incertitudes les plus importantes proviennent de l'estimation des  $\tau_{had}$  fakes et leur identification, des incertitudes théoriques sur le signal et des incertitudes sur l'énergie des jets.

## 9.9 Conclusion

L'analyse présentée dans ce Chapitre est une amélioration de l'analyse présentée au Chapitre précédent. De nouveaux canaux sont introduits ainsi que l'utilisation d'analyses multivariées pour séparer le signal  $t\bar{t}H$  des différents bruits de fond. L'estimation des bruits de fond dus à la mauvaise reconstruction de la charge des électrons a été utilisée dans les canaux  $2\ell SS0\tau$  et  $2\ell SS1\tau$ . Une catégorisation pour le canal  $2\ell SS0\tau$  a été mise en place avec pour but de comparer avec le résultat obtenu par l'analyse multivariée. Finalement le nombre d'événements de données considérées est aussi quasiment multiplié par trois puisque toutes les données enregistrées dans ATLAS en 2016 sont utilisées.

L'interprétation statistique de l'ensemble des canaux leptoniques a été effectuée et la force du signal a été déterminée pour la combinaison des canaux comme étant  $\mu = 1.6^{+0.5}_{-0.4}$  pour une signification statistique observée de  $4.1\sigma$  ( $2.8\sigma$  attendue). Ce résultat final montre une observation du processus  $t\bar{t}H$  compatible avec la prédiction du modèle standard. Une mesure de la section efficace du

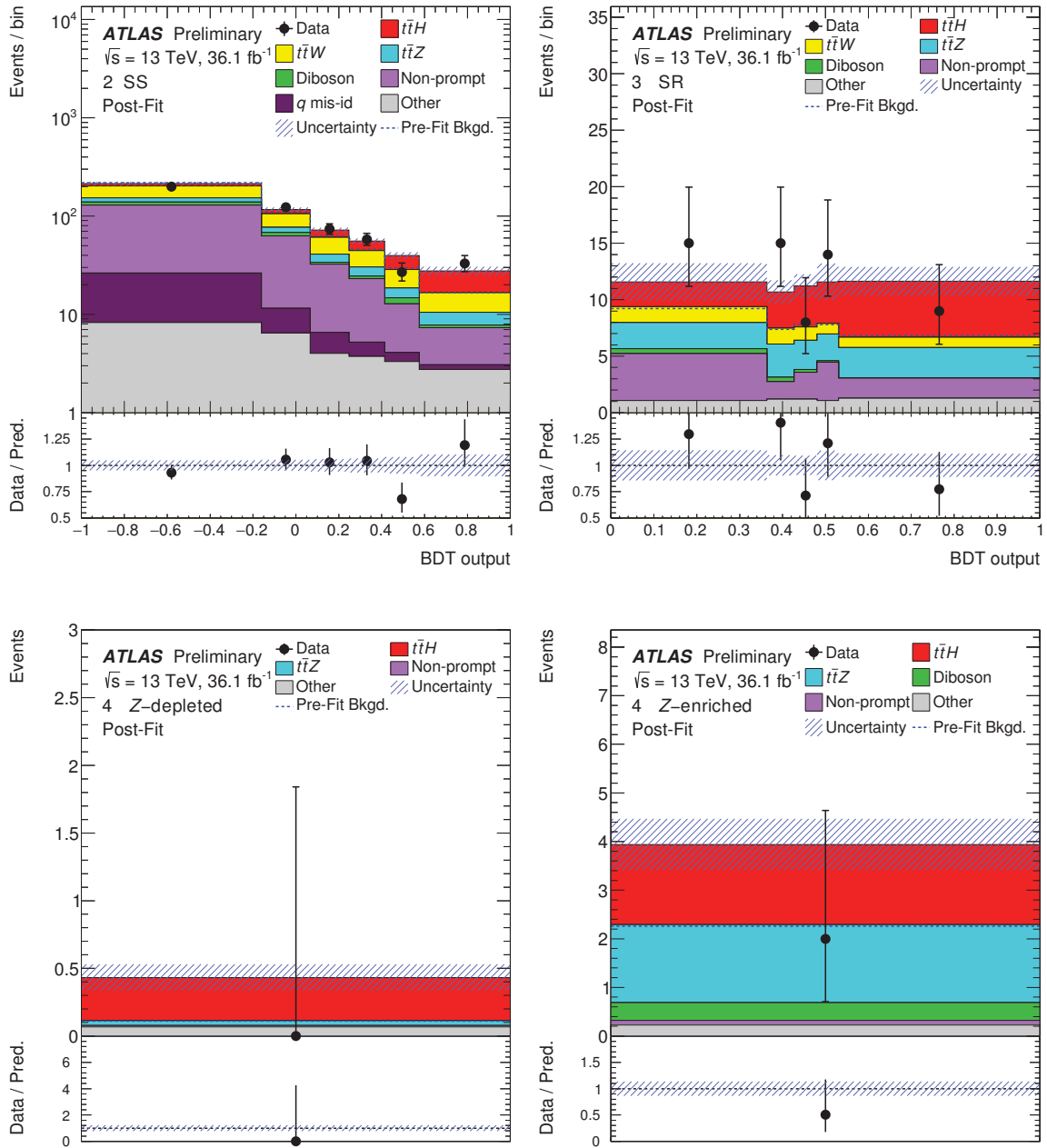


FIGURE 9.7 – Distribution des variables discriminantes utilisées pour les données et les prédictions pour les régions de signal (a)  $2\ell SS0\tau$ , (b)  $3\ell 0\tau$ , (c)  $4\ell$ (enrichi en  $Z$ ) et (d)  $4\ell$ (avec réjection des  $Z$ ) après ajustement aux données.

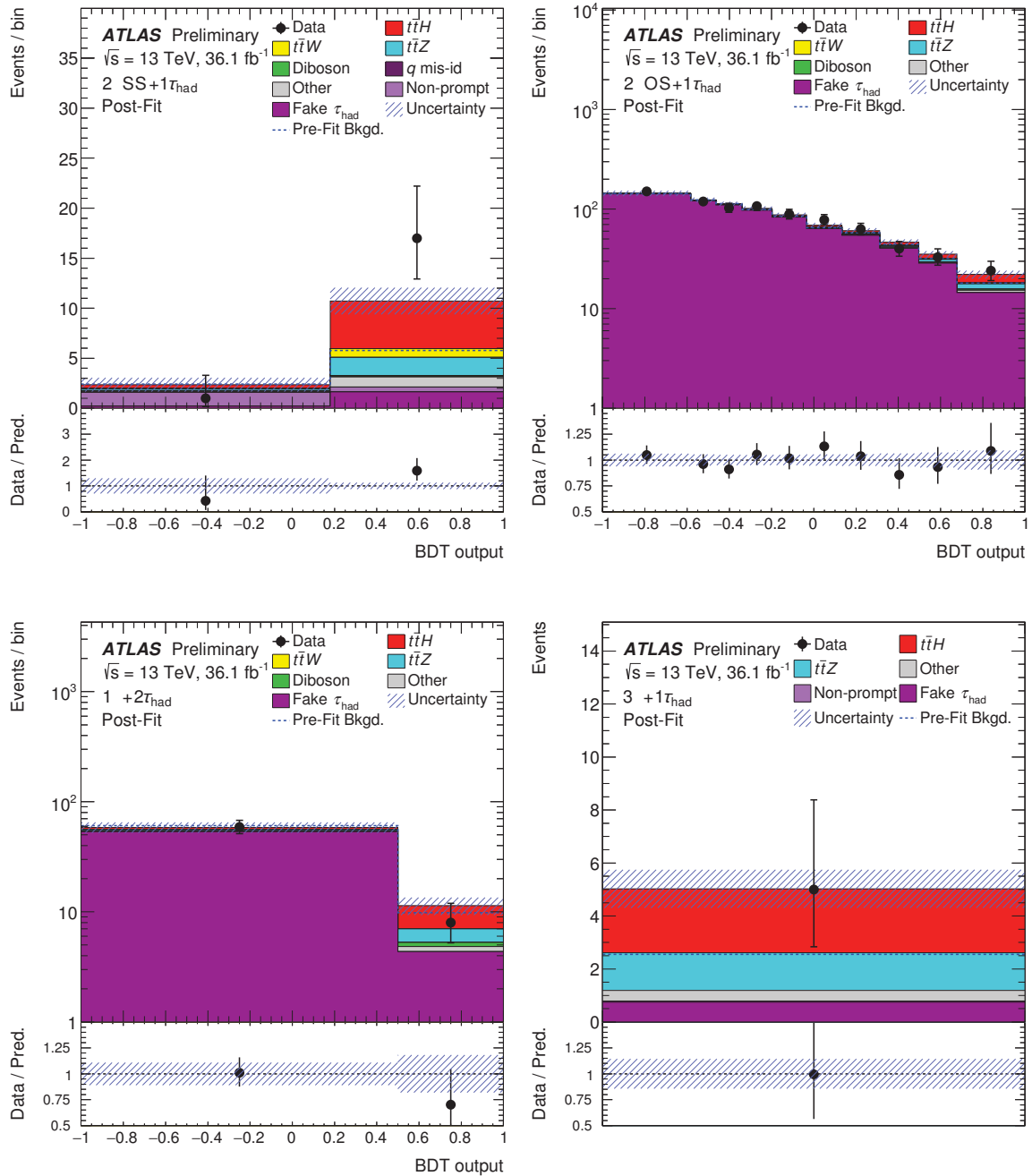


FIGURE 9.8 – Distribution des variables discriminantes utilisées pour les données et les prédictions pour les régions de signal (a)  $2\ell\text{SS}1\tau$ , (b)  $2\ell\text{OS}1\tau$ , (c)  $1\ell 2\tau$  et (d)  $3\ell 1\tau$  après ajustement aux données.

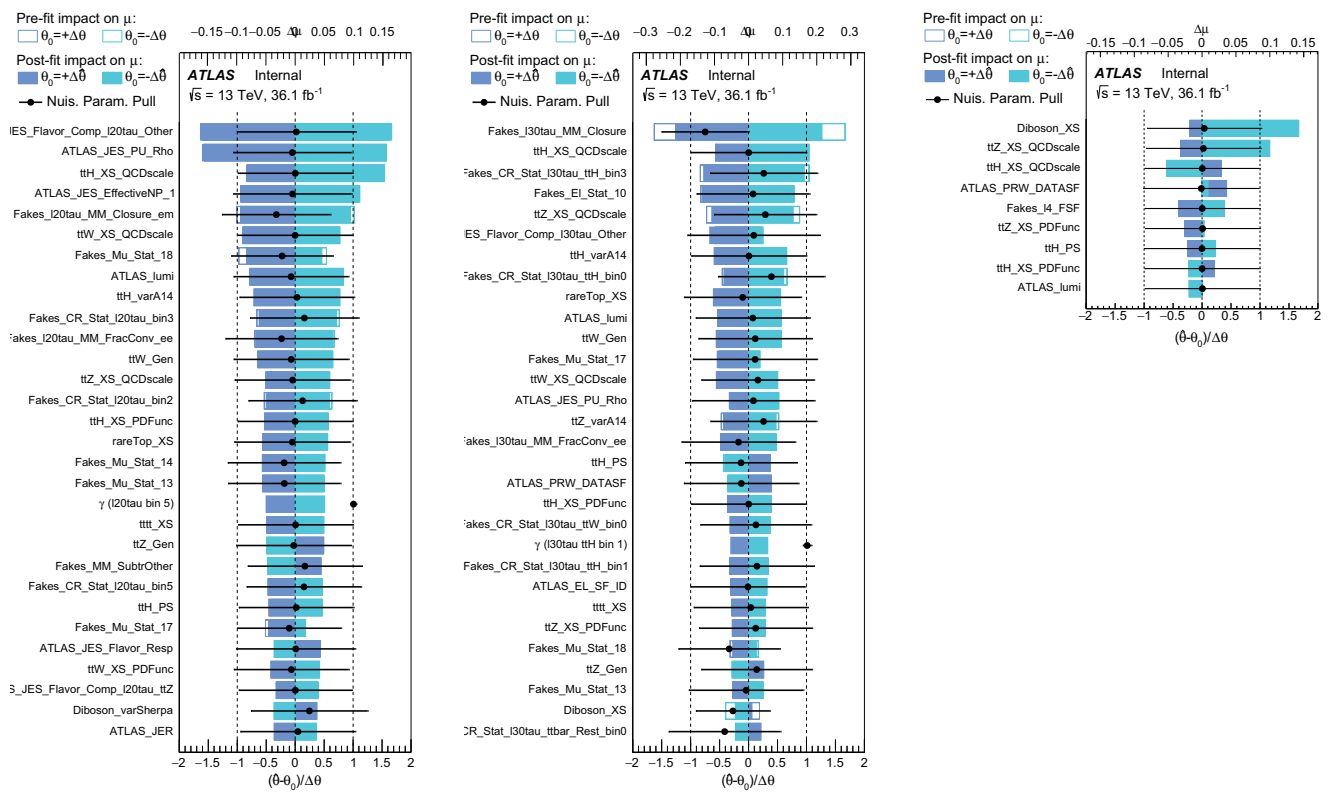


FIGURE 9.9 – Classement des paramètres de nuisance selon leur impact sur la force mesurée du signal lors de l’ajustement aux données observées pour le canal  $2\ell SS0\tau$  (gauche),  $3\ell 0\tau$  (centre) et  $4\ell$  (droite).

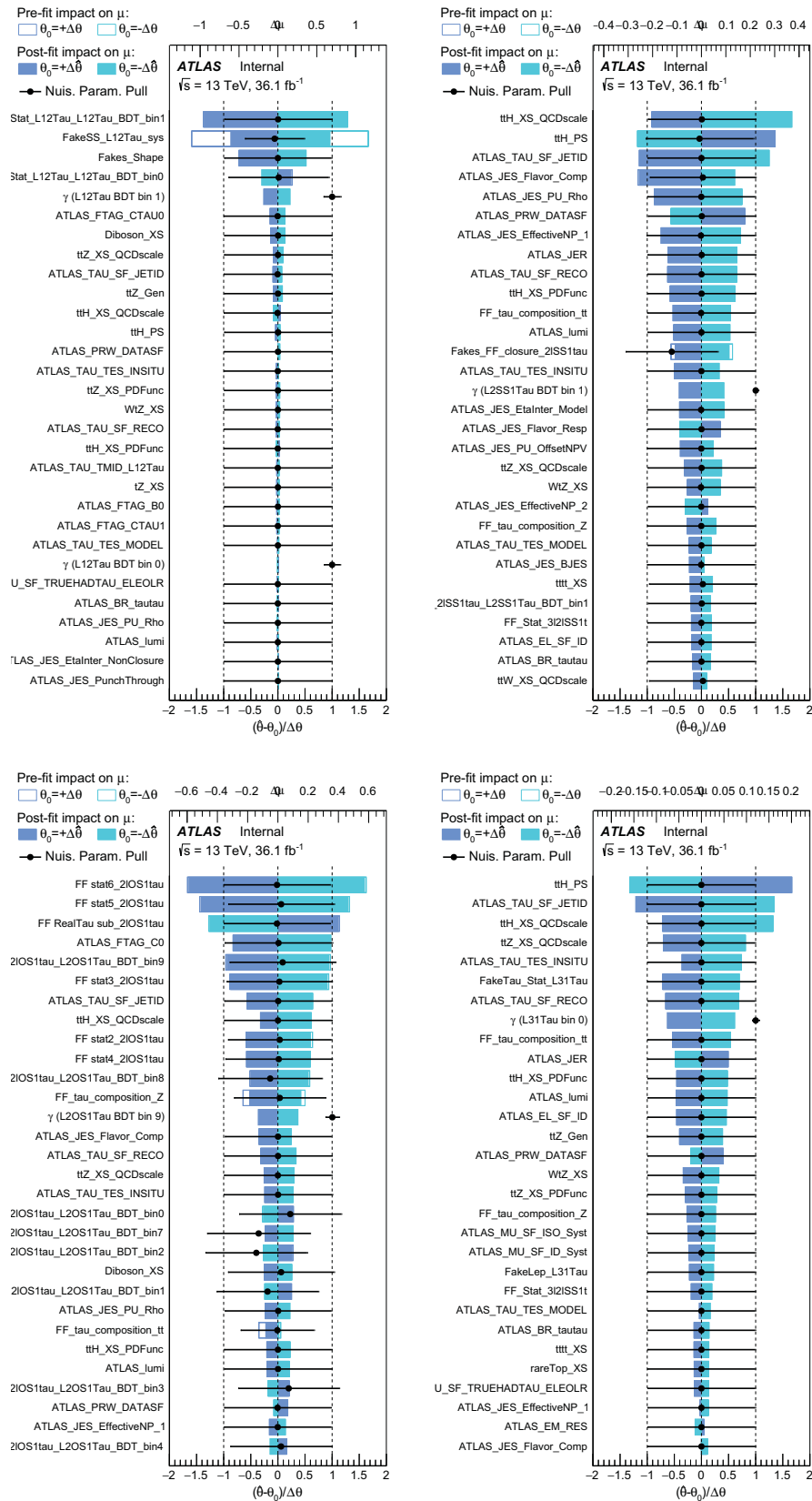


FIGURE 9.10 – Classement des paramètres de nuisance selon leur impact sur la force mesurée du signal lors de l’ajustement aux données observées pour le canal  $1l2\tau$  (haut gauche),  $2lSS1\tau$  (haut droite),  $2lOS1\tau$  (bas gauche) et  $3l1\tau$  (bas droite).



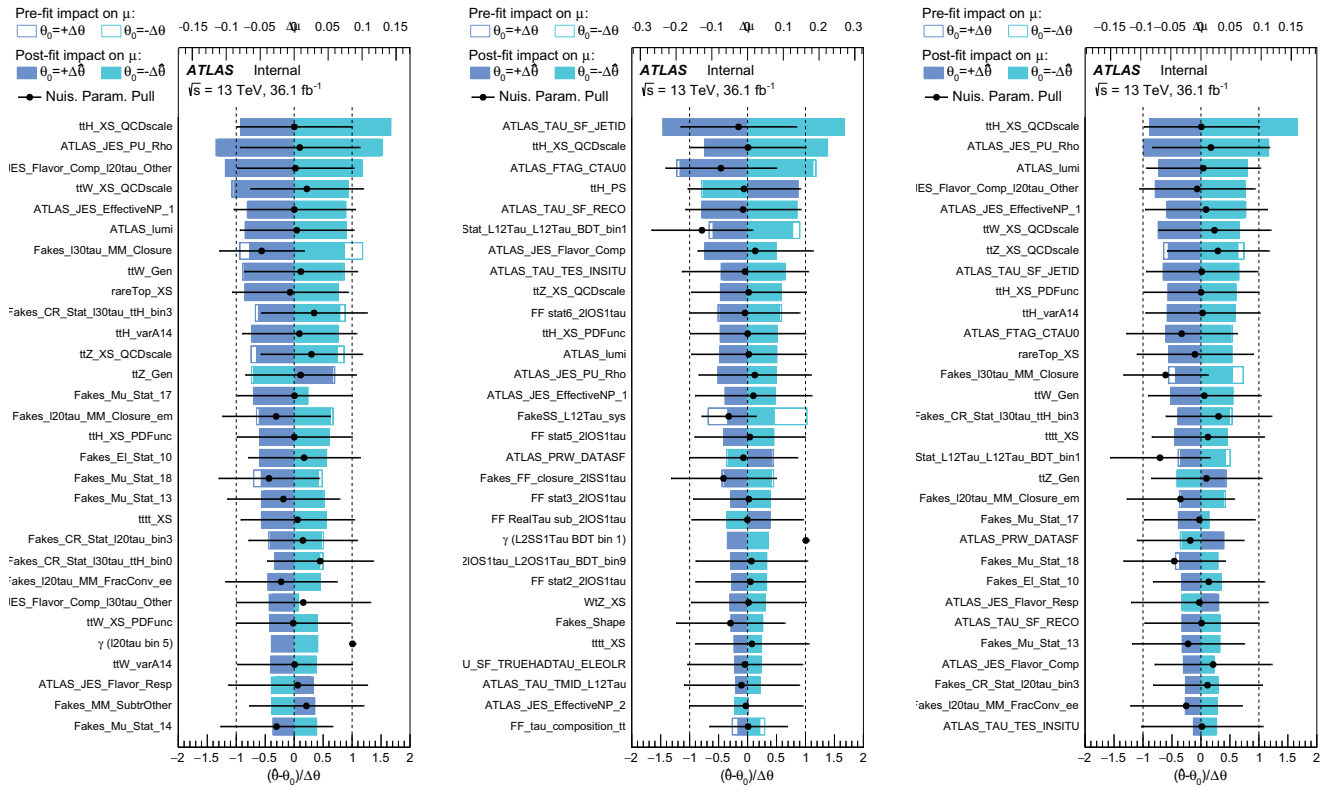


FIGURE 9.11 – Classement des paramètres de nuisance selon leur impact sur la force mesurée du signal lors de l’ajustement aux données observées pour la combinaison des canaux sans  $\tau_{had}$  (gauche), la combinaison des canaux avec  $\tau_{had}$  (centre) et pour la combinaison de tous les canaux (droite).

processus en a été déduite et aboutit à une valeur  $\sigma_{t\bar{t}H} = 790^{+230}_{-210}\text{fb}$ , compatible avec la prédiction du modèle standard de  $\sigma_{t\bar{t}H} = 507^{+35}_{-50}\text{fb}$ .

Incertitude Systématique	Valeur (%)	Type	Comp
Luminosité	3.2	N	1
Pondérage de l'effet d'empilement		SN	1
<b>Objets Physiques</b>			
Electron		SN	6
Muon		SN	15
taus		SN	10
Jet vertex fraction		SN	1
Jet		SN	28
MET		SN	3
étiquetage des quarks $b$		SN	126
TOTAL (Experimental)	–	–	191
<b>Lepton non prompt &amp; Mauvaise reco. de la charge (électron)</b>			
Statistique dans la région de contrôle		SN	38
Efficacité pour lepton prompt et non-prompt		SN	22
Test de fermeture pour lepton non-prompt		N	5
Fraction de conversion de photons		N	5
Estimation des $\tau_{had}$ fakes		N/SN	12
Mauvaise id. charge des électrons		SN	1
TOTAL (DD bdf réductible)	–	–	83
<b>MC ttW</b>			
Section efficace (QCD, PDF)	12.2, 3.4	N	2
Échelles d'énergie QCD		S	3
Generateur		SN	1
Modélisation des gerbes		SN	1
<b>MC ttZ</b>			
Section efficace (QCD, PDF)	10.4, 4.0	N	2
Échelles d'énergie QCD		S	3
Generateur		SN	1
Modélisation des gerbes		SN	1
<b>MC ttH</b>			
Section efficace (QCD, PDF)	7.5, 3.6	N	2
Échelles d'énergie QCD		S	3
Modélisation gerbe partonique		SN	1
Rapport de branchement Higgs		N	4
Modélisation des gerbes		SN	1
<b>MC Diboson</b>			
Section efficace	50	N	1
Échelles d'énergie QCD		SN	1
<b>MC Rare</b>			
Section efficace		N	14
TOTAL (Théorie)	–	–	41
TOTAL	–	–	315

TABLE 9.8 – Liste des incertitudes systématiques considérées pour l'obtention du résultat final. Pour chaque incertitude il est précisé si elle concerne uniquement la normalisation du processus (N), la forme des distributions (S) ou les deux (SN). Dans le cas où la systématique est divisée en plusieurs composantes, leur nombre est indiqué. Les processus "Rare" couvrent les bruits de fond  $ttWW$ ,  $WtZ$ ,  $VVV$ ,  $t\bar{t}$ ,  $tHjb$ ,  $WtH$ ,  $tZ$ ,  $t\bar{t}\bar{t}$ ,  $t\bar{t}t$  et les désintégrations rare du processus  $t\bar{t}$ .

TABLE 9.9 – Nombre d'événements attendus et observés dans les douze régions de l'analyse pour  $36.1fb^{-1}$  de données à 13 TeV. Les incertitudes reportées prennent en compte à la fois les incertitudes statistiques et systématiques. Les nombres d'événements pour les bruits de fond rares que sont  $ttWW$ ,  $WtZ$ ,  $VVV$ ,  $t\bar{t}$ ,  $tHjb$ ,  $WtH$ ,  $tZ$ ,  $t\bar{t}t$ ,  $t\bar{t}t$  et les désintégrations rare du processus  $t\bar{t}$  ne sont pas donnés séparément mais sont inclus dans la valeur attendue du bruit de fond.

Categorie	lepton non-prompt	$\tau_{had\ fake}$	QMisId	$t\bar{t}W$	$t\bar{t}Z$	Diboson	Bdf attendu	$t\bar{t}H(\mu = 1)$	Données
$2\ell SS0\tau$	$233 \pm 39$	–	$32.6 \pm 11.4$	$123 \pm 18$	$41.4 \pm 5.6$	$25 \pm 15$	$484 \pm 38$	$42.6 \pm 4.2$	514
$3\ell 0\tau$ SR	$14.5 \pm 4.3$	–	–	$5.5 \pm 1.2$	$12.0 \pm 1.8$	$1.2 \pm 1.2$	$39.1 \pm 5.2$	$11.2 \pm 1.6$	61
$3\ell 0\tau$ $t\bar{t}W$ CR	$13.3 \pm 4.3$	–	–	$19.9 \pm 3.1$	$8.7 \pm 1.1$	$< 0.2$	$46.5 \pm 5.4$	$4.18 \pm 0.46$	56
$3\ell 0\tau$ $t\bar{t}Z$ CR	$3.9 \pm 2.5$	–	–	$2.7 \pm 0.6$	$66 \pm 11$	$8.4 \pm 5.3$	$93 \pm 13$	$3.17 \pm 0.41$	107
$3\ell 0\tau$ $VV$ CR	$27.7 \pm 8.7$	–	–	$4.9 \pm 1.0$	$21.3 \pm 3.4$	$51 \pm 30$	$123 \pm 32$	$1.67 \pm 0.25$	109
$3\ell 0\tau$ $t\bar{t}CR$	$70 \pm 17$	–	–	$10.5 \pm 1.5$	$7.9 \pm 1.1$	$7.2 \pm 4.8$	$103 \pm 17$	$4.00 \pm 0.49$	85
$4\ell$ $Z$ -enr.	$0.11 \pm 0.07$	–	–	$< 0.01$	$1.52 \pm 0.23$	$0.43 \pm 0.23$	$2.26 \pm 0.34$	$1.06 \pm 0.14$	2
$4\ell$ $Z$ -dep.	$0.01 \pm 0.01$	–	–	$< 0.01$	$0.04 \pm 0.02$	$< 0.01$	$0.11 \pm 0.03$	$0.20 \pm 0.03$	0
$1\ell 2\tau$	–	$65 \pm 21$	–	$0.09 \pm 0.09$	$3.3 \pm 1.0$	$1.3 \pm 1.0$	$71 \pm 21$	$4.3 \pm 1.0$	67
$2\ell SS1\tau$	$2.4 \pm 1.4$	$1.80 \pm 0.30$	$0.053 \pm 0.023$	$0.88 \pm 0.24$	$1.83 \pm 0.37$	$0.12 \pm 0.18$	$8.2 \pm 1.6$	$3.09 \pm 0.46$	18
$2\ell OS1\tau$	–	$756 \pm 80$	–	$6.5 \pm 1.3$	$11.4 \pm 1.9$	$2.0 \pm 1.3$	$782 \pm 81$	$14.2 \pm 2.0$	807
$3\ell 1\tau$	–	$0.75 \pm 0.15$	–	$0.04 \pm 0.04$	$1.38 \pm 0.24$	$0.002 \pm 0.002$	$2.55 \pm 0.32$	$1.51 \pm 0.23$	5

	$\mu_{t\bar{t}H}$	Signification stat. ( $\sigma$ )
$2\ell SS0\tau$	$1_{-0.4}^{+0.4}(\text{stat})_{-0.4}^{+0.4}(\text{syst})$	1.9
$3\ell 0\tau$	$1_{-0.5}^{+0.6}(\text{stat})_{-0.4}^{+0.5}(\text{syst})$	1.5
$4\ell$	$1_{-1.2}^{+1.7}(\text{stat})_{-0.2}^{+0.4}(\text{syst})$	0.8
$2\ell SS1\tau$	$1_{-0.8}^{+1.1}(\text{stat})_{-0.3}^{+0.5}(\text{syst})$	1.1
$3\ell 1\tau$	$1_{-1.1}^{+1.5}(\text{stat})_{-0.2}^{+0.4}(\text{syst})$	0.9
$2\ell OS1\tau$	$1_{-1.4}^{+1.5}(\text{stat})_{-1.1}^{+1.2}(\text{syst})$	0.5
$1\ell 2\tau$	$1_{-0.9}^{+1.1}(\text{stat})_{-1.1}^{+1.2}(\text{syst})$	0.6
Combinaison	$1_{-0.3}^{+0.3}(\text{stat})_{-0.3}^{+0.3}(\text{syst})$	2.8

TABLE 9.10 – Incertitude sur la force du signal et signification statistique attendue à partir d'un ajustement sur des données Asimov.

Canal	$\mu_{t\bar{t}H}$	Signification statistique ( $\sigma$ )
$2\ell OS1\tau$	$1.7_{-1.5}^{+1.6}(\text{stat.})_{-1.1}^{+1.4}(\text{syst.})$	0.9
$1\ell 2\tau$	$-0.6_{-0.8}^{+1.1}(\text{stat.})_{-1.3}^{+1.1}(\text{syst.})$	–
$4\ell$	$-0.5_{-0.8}^{+1.3}(\text{stat.})_{-0.3}^{+0.2}(\text{syst.})$	–
$3\ell 1\tau$	$1.6_{-1.3}^{+1.7}(\text{stat.})_{-0.2}^{+0.6}(\text{syst.})$	$1.3\sigma$
$2\ell SS1\tau$	$3.5_{-1.2}^{+1.5}(\text{stat.})_{-0.5}^{+0.9}(\text{syst.})$	$3.4\sigma$
$3\ell 0\tau$	$1.8_{-0.6}^{+0.6}(\text{stat.})_{-0.5}^{+0.6}(\text{syst.})$	$2.4\sigma$
$2\ell SS0\tau$	$1.5_{-0.4}^{+0.4}(\text{stat.})_{-0.4}^{+0.5}(\text{syst.})$	$2.7\sigma$
Combinaison	$1.6_{-0.3}^{+0.3}(\text{stat.})_{-0.3}^{+0.4}(\text{syst.})$	$4.1\sigma$

TABLE 9.11 – Force du signal  $\mu$  mesurée dans les données et signification statistique associée dans le cas de l'hypothèse bruit de fond seul.

TABLE 9.12 – Résumé de l'impact des différentes groupes d'incertitudes sur la force du signal.

Source	$\Delta\mu$	
Modélisation $t\bar{t}H$ (section efficace)	+0.20	-0.09
Échelle et résolution sur l'énergie des jets	+0.18	-0.15
Estimations des leptons non-prompts	+0.15	-0.13
Étiquetages de la saveur des jets et identification de $\tau_{had}$	+0.11	-0.09
Modélisation $t\bar{t}W$	+0.10	-0.09
Modélisation $t\bar{t}Z$	+0.08	-0.07
Modélisation des autres bruits de fond	+0.08	-0.07
Luminosité	+0.08	-0.06
Modélisation $t\bar{t}H$ (acceptance)	+0.08	-0.04
Estimations des $\tau_{had}$ fakes	+0.07	-0.07
Autres incertitudes expérimentales	+0.05	-0.04
Taille des simulations	+0.04	-0.04
QMisid	+0.01	-0.01
Total	+0.39	-0.30



# Combinaison des analyses $t\bar{t}H$

---

## 10.1 Introduction

L'analyse de la production associée  $t\bar{t}H$  à travers un état final multilepton n'est qu'une des options considérées pour observer ce processus. D'autres désintégrations du boson de Higgs sont utilisées pour cette recherche :  $H \rightarrow b\bar{b}$  [76] et  $H \rightarrow \gamma\gamma$  [77]. Avec les données collectées au LHC jusqu'en 2016, aucune des analyses correspondant à ces états finals n'a une sensibilité suffisante pour observer le processus  $t\bar{t}H$  donc une combinaison est faite pour tenter d'observer ce processus avec les données 2015 et les premières données 2016 pour une luminosité intégrée de  $13.2fb^{-1}$  [78]. Une seconde combinaison est ensuite réalisée avec l'ensemble des données 2015 et 2016 pour une luminosité intégrée de  $36.1fb^{-1}$  [79]. J'ai participé à ces combinaison en fournissant les résultats de l'analyse multilepton et en travaillant à l'harmonisation des processus de bruits de fond et des incertitudes systématiques avec les groupes des autres analyses.

## 10.2 Recherche du processus $t\bar{t}H$ avec $H \rightarrow b\bar{b}$

Deux catégories sont définies dans l'étude du processus  $t\bar{t}H$  à travers la désintégration du boson de Higgs en paire  $b\bar{b}$  :

- au moins un lepton et quatre (cinq pour l'analyse avec toutes les données) jets dont deux sont étiquetés  $b$ .
- au moins deux leptons de charge opposées et trois jets dont deux sont étiquetés  $b$ .

Des catégories supplémentaires dans les deux cas en fonction du nombre de jets et de jets étiquetés  $b$ . Pour les régions les plus enrichies en signal qui sont des régions avec 3 ou 4 jets étiquetés  $b$ , des analyses multivariées sont utilisées pour séparer le signal  $t\bar{t}H$  des bruits de fond  $t\bar{t}$ .

## 10.3 Recherche du processus $t\bar{t}H$ avec $H \rightarrow \gamma\gamma$

Cette analyse recherche le processus  $t\bar{t}H$  à travers un ajustement aux données de la masse invariante de deux photons simultanément pour le signal et le bruit de fond. Pour cette analyse, deux catégories existent :

- leptonique : au moins un lepton et deux jets dont l'un est étiqueté  $b$ .
- hadronique : au moins cinq (quatre pour l'analyse de toutes les données) jets dont l'un est étiqueté  $b$ .

Par ailleurs un veto est fait sur les événements comprenant une paire de photons dont la masse invariante est proche de la masse du boson  $Z$ . L'avantage de cette analyse est d'avoir une région de signal très pure en  $t\bar{t}H$  mais au prix d'un nombre d'événements attendus faible.

## 10.4 Résultat de la combinaison

### 10.4.1 Orthogonalité des analyses

Chaque analyse considère spécifiquement certaines voies de désintégration du boson de Higgs, toutefois les analyses multilepton et  $b\bar{b}$  ont certains événements dans leurs régions de signal corres-

pondant à des voies de désintégration non choisie. La Table 10.1 présente pour chaque analyse et région de signal les désintégrations du boson de Higgs des événements  $t\bar{t}H$  pour l'analyse à  $13.2fb^{-1}$ .

Canal	Region	$WW$	$\tau\tau$	$ZZ$	$b\bar{b}$	$\gamma\gamma$
$H \rightarrow \gamma\gamma$	hadronique	–	–	–	–	100%
	leptonique	–	–	–	–	100%
$H \rightarrow (WW, \tau\tau, ZZ)$	$2\ell SS ee$	76%	17%	2%	4%	–
	$2\ell SS e\mu$	77%	17%	3%	3%	–
	$2\ell SS \mu\mu$	79%	17%	3%	1%	–
	$2\ell SS + \tau_{had}$	46%	51%	2%	1%	–
	$3\ell$	74%	20%	4%	1%	–
	$4\ell$	72%	18%	9%	–	–
$H \rightarrow b\bar{b}$	$\ell + jets (\geq 6j, 3bj)$	5%	1%	1%	90%	–
	$\ell + jets (5j, \geq 4bj)$	–	–	–	99%	–
	$\ell + jets (\geq 6j, \geq 4bj)$	1%	–	1%	97%	–
	dilepton ( $\geq 4j, 3bj$ )	6%	1%	1%	90%	–
	dilepton ( $\geq 4j, \geq 4bj$ )	–	–	–	98%	–

TABLE 10.1 – Contribution de chaque désintégration du boson de Higgs pour les régions de signal des trois analyses. Seules les contributions supérieures au pourcent sont indiquées.

Pour avoir l'orthogonalité des analyses, la sélection des objets utilisés pour la catégorisation et la procédure d'*Overlap Removal* a été harmonisée entre les trois analyses. Ainsi la même définition est utilisée pour sélectionner les leptons *Loose*. Par ailleurs l'analyse  $t\bar{t}(H \rightarrow b\bar{b})$  ajoute un veto sur la présence d'un lepton supplémentaire. Cet ajout représente une diminution du nombre d'événements de signal de l'ordre de 2% dans les régions les plus sensibles.

Aucun veto n'est fait sur les photons dans les analyses multilepton et  $b\bar{b}$  car le nombre d'événements de signal en commun pour l'une de ces analyses avec l'analyse  $H \rightarrow \gamma\gamma$  a été estimée inférieur au pourcent dans les deux cas.

#### 10.4.2 Schéma de corrélation

Sur les quelques 200 paramètres de nuisances présents lors de la combinaison des trois analyses pour les deux luminosités intégrées considérées, ceux associés à la même source dans les différentes analyses sont corrélés, en particulier les incertitudes théoriques sur la section efficace, l'acceptance ou les valeurs des rapports de branchements du signal sont corrélés.

Les principaux processus de bruits de fond diffèrent d'une analyse à l'autre, aussi les incertitudes associées sont décorrélées, excepté pour les processus  $t\bar{t}V$  et Diboson dont les incertitudes sont corrélées entre les analyses multilepton et  $b\bar{b}$ .

#### 10.4.3 Résultats pour l'analyse à $13.2fb^{-1}$

La combinaison des trois analyses après ajustement aux données aboutit à :

$$\mu_{t\bar{t}H} = 1.8^{+0.4}_{-0.4}(stat)^{+0.6}_{-0.5}(syst) = 1.8^{+0.7}_{-0.7}. \quad (10.1)$$

La Figure 10.1 regroupe les valeurs de la force du signal pour chaque analyse individuelle ainsi que le résultat de la combinaison des mêmes analyses effectuée au Run1 du LHC.



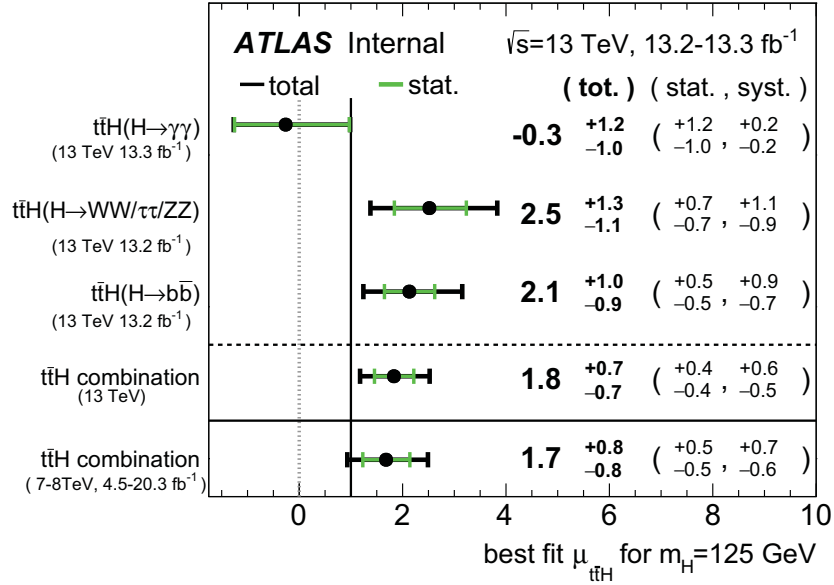


FIGURE 10.1 – Résumé de la valeur de la force observée du signal dans les analyses individuelles ainsi que pour la combinaison, en supposant un boson de Higgs de masse de 125 GeV. La valeur de la force observée du signal au Run 1 du LHC est aussi indiquée pour comparaison.

Les différentes sources d'incertitudes pour cette combinaison et leur impact sur la force mesurée du signal sont présentées dans la Table 10.2. La source d'incertitude la plus importante porte sur l'estimation du processus  $t\bar{t}$  produits avec des quarks  $b$  supplémentaires.

Source	$\Delta\mu$	
Modélisation $t\bar{t}+ \geq 1b$	+0.34	-0.33
Etiquetage saveur des jets	+0.19	-0.19
Background Model statistics	+0.18	-0.18
Modélisation $t\bar{t}+ \geq 1c$	+0.17	-0.17
Échelle et résolution en énergie des jets	+0.18	-0.18
Modélisation $t\bar{t}H$	+0.20	-0.13
Modélisation $t\bar{t}+light$	+0.14	-0.14
Modélisation autres bruits de fond	+0.16	-0.15
Incertitudes Fake lepton	+0.11	-0.12
Association Jet-vertex, modélisation du pileup	+0.09	-0.09
Luminosité	+0.09	-0.09
Modélisation $t\bar{t}Z$	+0.08	-0.07
ID, Isolation, trigger des leptons légers ( $e, \mu$ ), photons, et $\tau$	+0.04	-0.04
Incertitude systématique totale	+0.57	-0.54
Normalisation $t\bar{t}+ \geq 1b$	+0.24	-0.24
Normalisation $t\bar{t}+ \geq 1c$	+0.11	-0.11
Incertitude statistique	+0.38	-0.38
Incertitude totale	+0.69	-0.66

TABLE 10.2 – Résumé des incertitudes présentes dans la combinaison des canaux individuels  $t\bar{t}H$ . Les incertitudes sont classées selon leur impact sur la force observée du signal.

Finalement la Table 10.3 regroupe les significations statistiques des canaux individuels et de la combinaison. Le résultat final est d'une signification statistique observée (attendue) de  $2.8\sigma$  ( $1.8\sigma$ ).

Canal	Signification statistique	
	Observée [ $\sigma$ ]	Attendue [ $\sigma$ ]
$t\bar{t}H, H \rightarrow \gamma\gamma$	-0.2	0.9
$t\bar{t}H, H \rightarrow (WW, ZZ, \tau\tau)$	2.2	1.0
$t\bar{t}H, H \rightarrow b\bar{b}$	2.4	1.2
Combinaison $t\bar{t}H$	2.8	1.8

TABLE 10.3 – Signification observée et attendue par rapport à l'hypothèse de bruit de fond seul pour les canaux individuels et la combinaison.

#### 10.4.4 Résultats pour l'analyse à $36.1 fb^{-1}$

La combinaison finale de l'analyse des différentes voies de désintégrations du boson de Higgs aboutissent à  $\mu_{t\bar{t}H} = 1.2^{+0.3}_{-0.3}$  compatible avec la prédiction du modèle standard. Les valeurs pour chaque analyse individuelle sont reportées dans la table 10.4.

Dans ce cas une observation était attendue avec une signification statistique de  $3.8\sigma$  et confirmée par les données et une signification statistique finale de  $4.2\sigma$ . Une mesure de la section efficace du processus  $t\bar{t}H$  a été extraite à  $\sigma(t\bar{t}H) = 590^{+160}_{-150} \text{fb}$ , compatible avec la prédiction du modèle standard ( $\sigma(t\bar{t}H) = 507^{+35}_{-50} \text{fb}$ ).

Chaque analyse visant des désintégrations du boson de Higgs spécifiques, il est possible d'extraire séparément des forces du signal pour les désintégrations en bosons vecteurs, en quarks  $b$  ou en leptons  $\tau$ . Les Figures 10.2 présentent en deux dimensions les valeurs de ces forces du signal comparées à la valeur du modèle standard (par définition  $\mu_{VV} = \mu_{b\bar{b}} = \mu_{\tau\bar{\tau}} = 1$ ).

Dans le cadre de cette combinaison, les facteurs de couplages avec les bosons vecteurs  $\kappa_V$  et avec les fermions  $\kappa_f$  ont aussi été extraits. Dans ce cas les rapports de branchement du boson de Higgs sont gardés constants. Les facteurs de couplages obtenus, représentés sur la Figure 10.3, sont compatibles avec la prédiction du modèle standard ( $\kappa_V = \kappa_f = 1$ ).

TABLE 10.4 – Résumé des valeurs de  $\mu$  attendues et observées et de la signification statistique de chaque analyse des différentes voies de désintégrations du boson de Higgs et de la combinaison. Dans le cas de l'analyse  $H \rightarrow 4\ell$ , une limite à 68% de niveau de confiance sur  $\mu$  est extraite suite à l'absence d'événements observés.

Canal	$\mu$		Signification statistique	
	Observée	Attendue	Observée	Attendue
Multilepton	$1.6^{+0.5}_{-0.4}$	$1.0^{+0.4}_{-0.4}$	$4.1\sigma$	$2.8\sigma$
$H \rightarrow b\bar{b}$	$0.8^{+0.6}_{-0.6}$	$1.0^{+0.6}_{-0.6}$	$1.4\sigma$	$1.6\sigma$
$H \rightarrow \gamma\gamma$	$0.6^{+0.7}_{-0.6}$	$1.0^{+0.8}_{-0.6}$	$0.9\sigma$	$1.7\sigma$
$H \rightarrow 4\ell$	$< 1.9$	$1.0^{+3.2}_{-1.0}$	—	$0.6\sigma$
Combinaison	$1.2^{+0.3}_{-0.3}$	$1.0^{+0.3}_{-0.3}$	$4.2\sigma$	$3.8\sigma$

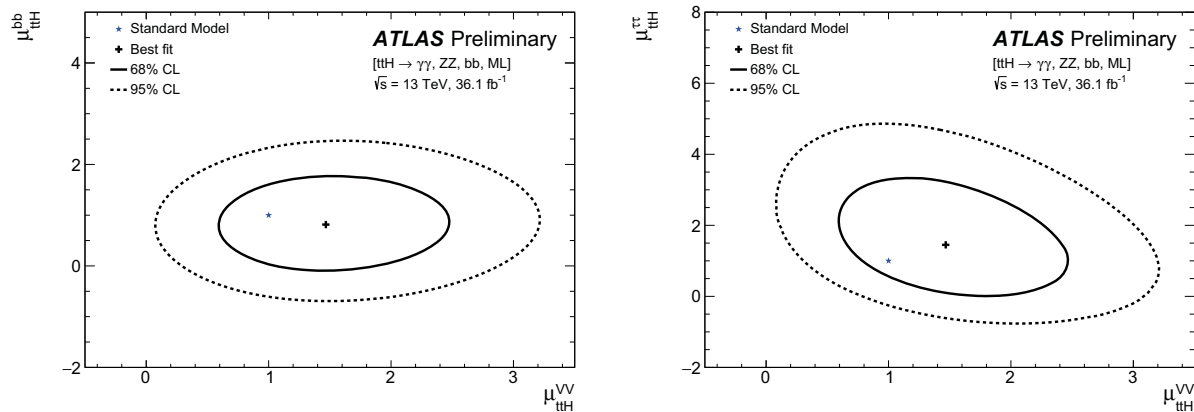


FIGURE 10.2 – Valeurs des force du signal séparées pour les différentes désintégrations du boson de Higgs pour les processus  $t\bar{t}H, H \rightarrow b\bar{b}$  vs  $t\bar{t}H, H \rightarrow WW, ZZ$  (gauche) et  $t\bar{t}H, H \rightarrow \tau\bar{\tau}$  vs  $t\bar{t}H, H \rightarrow WW, ZZ$  (droite).

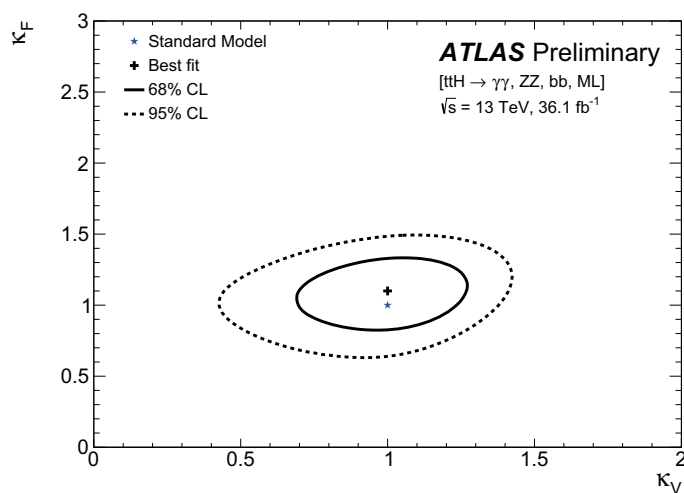


FIGURE 10.3 – Région disponible pour un niveau de confiance de 68% et 95% dans le plan  $\kappa_V - \kappa_F$ . Il est supposé que le boson de Higgs ne se couple qu'à des particules du Modèle Standard.

## 10.5 Conclusion

La combinaison des différentes analyses  $t\bar{t}H$  dans ATLAS a été effectuée pour  $13.2\text{ fb}^{-1}$  de données aboutissant à une signification statistique observée (attendue) de  $2.8$  ( $1.8$ )  $\sigma$ . De nombreuses améliorations ont été faites dans les différentes analyses et la combinaison des différentes analyses pour une luminosité finale de  $36.1\text{ fb}^{-1}$  aboutit à une signification statistique observée (attendue) de  $4.2\sigma$  ( $3.8\sigma$ ) aboutissant à l'observation du processus  $t\bar{t}H$ . Une première extraction des facteurs de couplage  $\kappa_V$  et  $\kappa_f$  aboutissent à des valeurs compatibles avec le modèle standard.

# Conclusion

Suite à l'annonce de la découverte du boson de Higgs en juillet 2012, un champ d'étude nouveau s'est établi avec la mesure des propriétés de cette particule pour tester le modèle standard. L'une de ces propriétés est le couplage du boson de Higgs aux autres particules et plus particulièrement aux fermions. Au cours de cette thèse, l'étude du couplage de Yukawa du boson de Higgs au quark *top* a été menée par la recherche du processus de production d'un boson de Higgs associé à une paire  $t\bar{t}$  dans un état final avec plusieurs leptons et en particulier deux leptons de même charge électrique.

La première partie de ce document a été consacrée à l'étalonnage du calorimètre hadronique à tuiles d'ATLAS par un système Laser. Ce système permet avec les autres systèmes d'étalonnage du calorimètre de s'assurer d'une réponse en énergie stable du calorimètre au cours du temps. Dans ce cadre une estimation de la précision du nouveau système Laser d'ATLAS au Run 2 a été effectuée. Pour évaluer la validité des variations observées par le système Laser, une comparaison a été effectuée avec les variations observées par un autre système d'étalonnage du *Tilecal* qu'est le système Césium. Enfin un algorithme a été mis en place pour automatiser l'extraction des constantes d'étalonnage du calorimètre. En effet avec l'augmentation de la luminosité depuis le Run 1, le nombre de canaux ayant une variation de gain non négligeable est de plus en plus important. Cet algorithme qui est employé depuis la prise de données en 2016 réduit le temps nécessaire pour extraire les constantes d'étalonnages et la liste de canaux défectueux en effectuant des tests auparavant fait manuellement.

La seconde partie de cette thèse a porté sur la recherche du processus  $t\bar{t}H$  dans un état final avec plusieurs leptons. Deux réalisations de cette analyse ont été présentées. Une première portait sur les données de 2015 et une portion des données de 2016 pour un total de  $13.2 \text{ fb}^{-1}$  pour une énergie dans le centre de masse de 13 TeV. Cette analyse reprenait l'analyse effectuée au Run 1 avec un comptage des événements dans des régions enrichies en signal. Le travail effectué ici a porté sur l'estimation du nombre d'événements avec une mauvaise identification de la charge des électrons et des leptons non-prompts dans les régions avec deux leptons de même charge ainsi que sur le traitement statistique des données. Cette analyse n'a pas permis une mise en évidence du processus  $t\bar{t}H$  conduisant à la recherche d'amélioration pour une seconde analyse. Suite à cette analyse une combinaison de l'ensemble des analyses portant sur le processus  $t\bar{t}H$  dans ATLAS a été réalisée, avec le travail réalisé dans le cadre de cette thèse portant sur l'harmonisation des incertitudes du point de vue de l'analyse multilepton avec les autres groupes.

Une seconde analyse utilisant l'ensemble des données du LHC pour 2015 et 2016 est en cours. Cette analyse présente une évolution importante dans la stratégie d'analyse avec l'utilisation d'analyses multivariées et l'ajout d'états finals supplémentaires. Le travail réalisé pour cette analyse a consisté à modifier l'estimation de la mauvaise identification de la charge des électrons en ajoutant une dimension supplémentaire nécessaire pour l'estimation des leptons non-prompts. Par ailleurs une analyse similaire à la précédente consistant à définir des catégories avant de compter les événements dans ces catégories a été menée en parallèle de l'analyse principale utilisant des BDTs pour le canal avec deux leptons de même charge électrique. Finalement l'ajustement aux données a été effectué pour aboutir à une signification statistique observée de  $4.1\sigma$ .

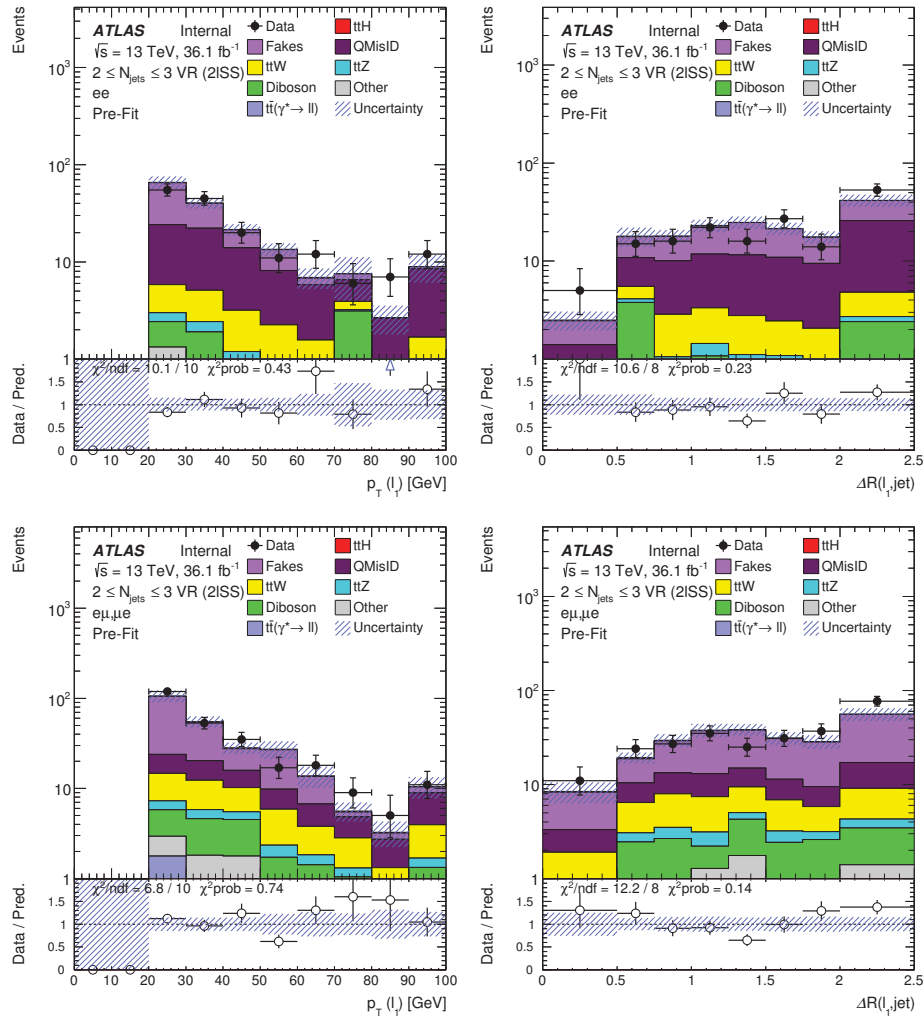
Une nouvelle combinaison des analyses portant sur le processus  $t\bar{t}H$  a été effectuée et la signification statistique observée (attendue) finale est de  $4.2\sigma$  ( $3.8\sigma$ ). Ainsi ces analyses ont permis l'observation du processus  $t\bar{t}H$ . La prochaine étape importante pour cette analyse est d'atteindre une signification statistique de  $5\sigma$  conduisant à se prononcer sur la découverte du processus  $t\bar{t}H$ . À partir de là, une mesure du couplage de Yukawa du quark top au boson de Higgs pourra être menée. Pour aboutir

à ce résultat en améliorant la sensibilité de l'analyse multilepton, un premier travail a été effectué sur la reconstruction du processus  $t\bar{t}$  semileptonique par ajustement cinématique dans le but de rejeter ce bruit de fond.

# Validation des bruits de fond l'analyse à $36.1 \text{ fb}^{-1}$

## A.1 Validation des leptons non-prompts dans les canaux $2\ell SS0\tau$ et $3\ell 0\tau$

Les comparaisons des prédictions et des données dans des régions enrichies en leptons non-prompts sont présentées sur les Figures A.1 et A.1 respectivement pour les canaux  $2\ell SS0\tau$  et  $3\ell 0\tau$ .



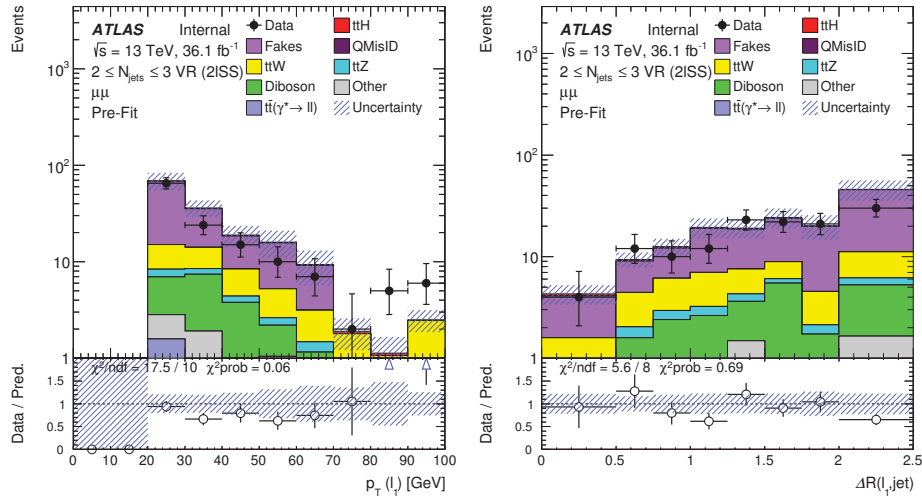
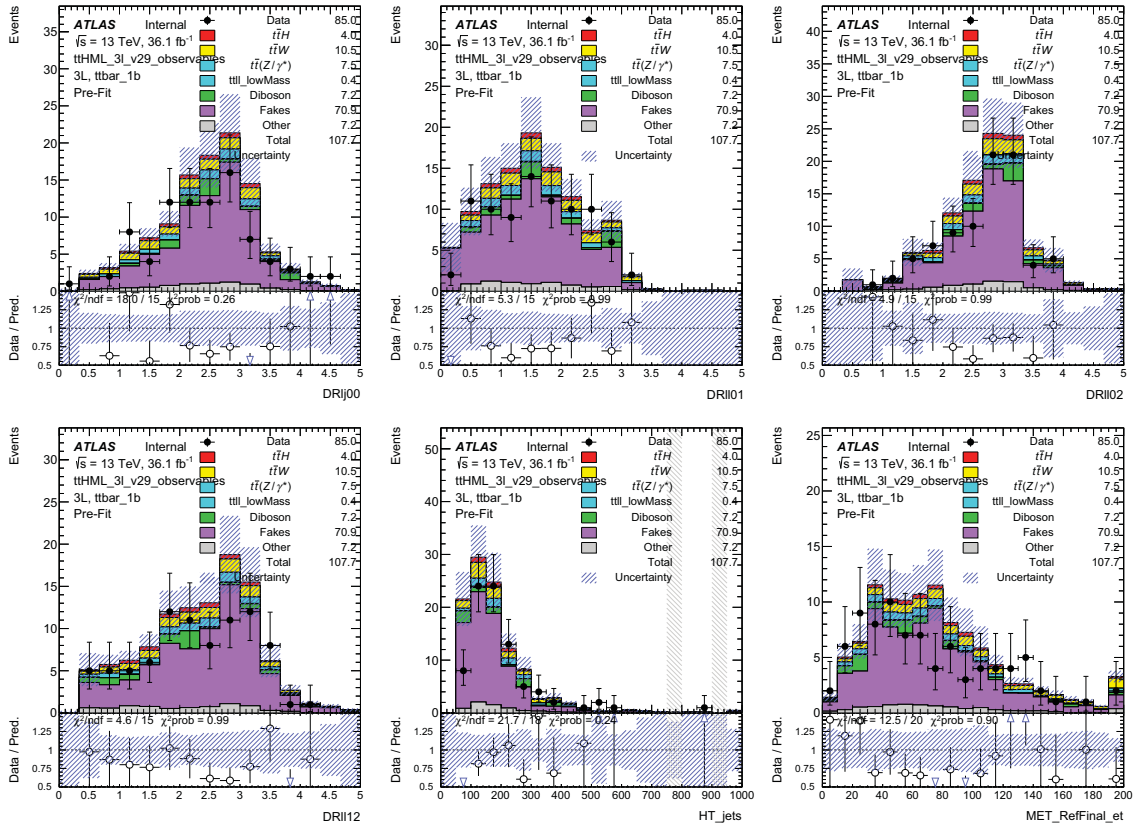


FIGURE A.1 – Comparaison prédiction et données dans des régions à basses multiplicités en nombre de jets correspondant à des régions enrichies en lepton non-prompts dans les canaux avec deux leptons de même charge.





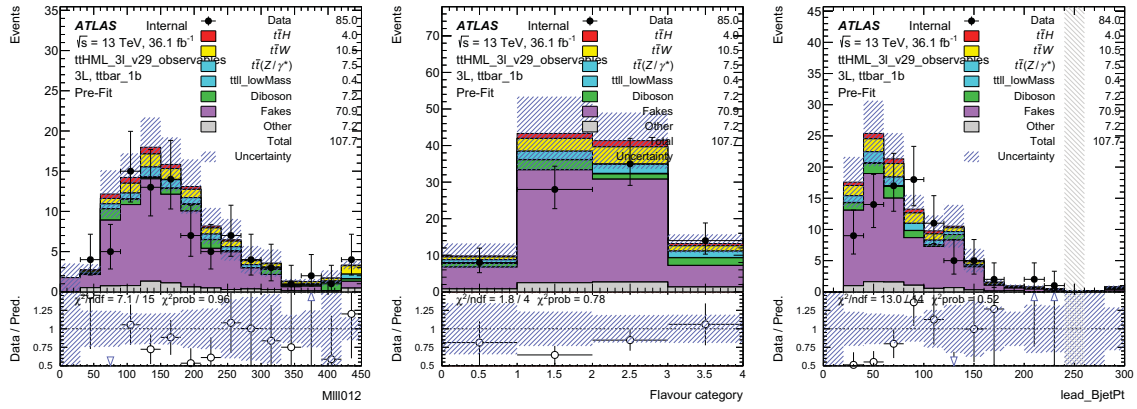


FIGURE A.2 – Comparaison prédiction et données dans une région enrichie en lepton non-prompts définie par BDT pour trois leptons.

## A.2 Validation du processus $t\bar{t}W$ dans les canaux $2\ell SS0\tau$ et $3\ell 0\tau$

Les comparaisons des prédictions et des données dans des régions enrichies en leptons non-prompts sont présentées sur les Figures A.3 pour les canaux  $2\ell SS0\tau$  et  $3\ell 0\tau$ .

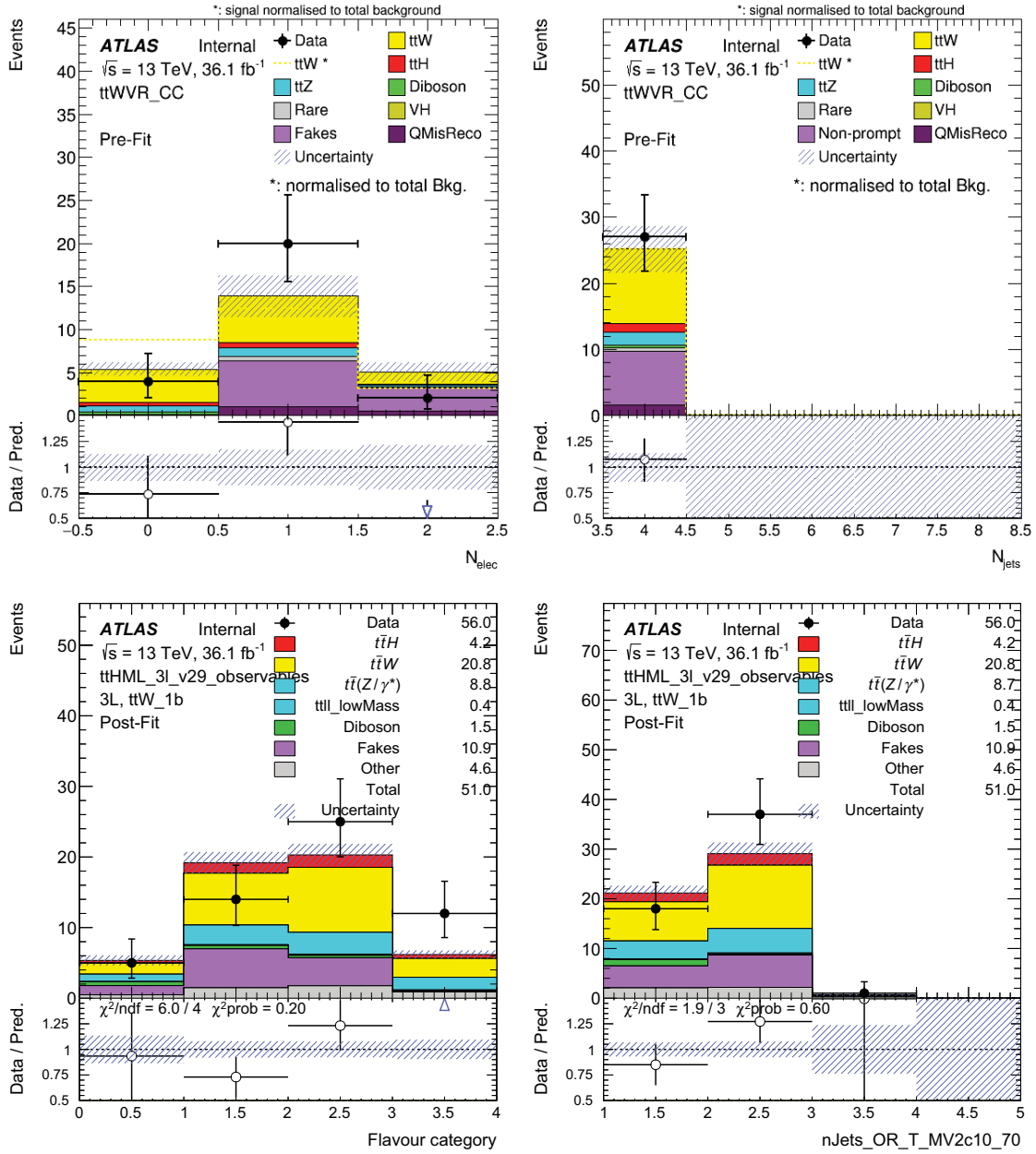
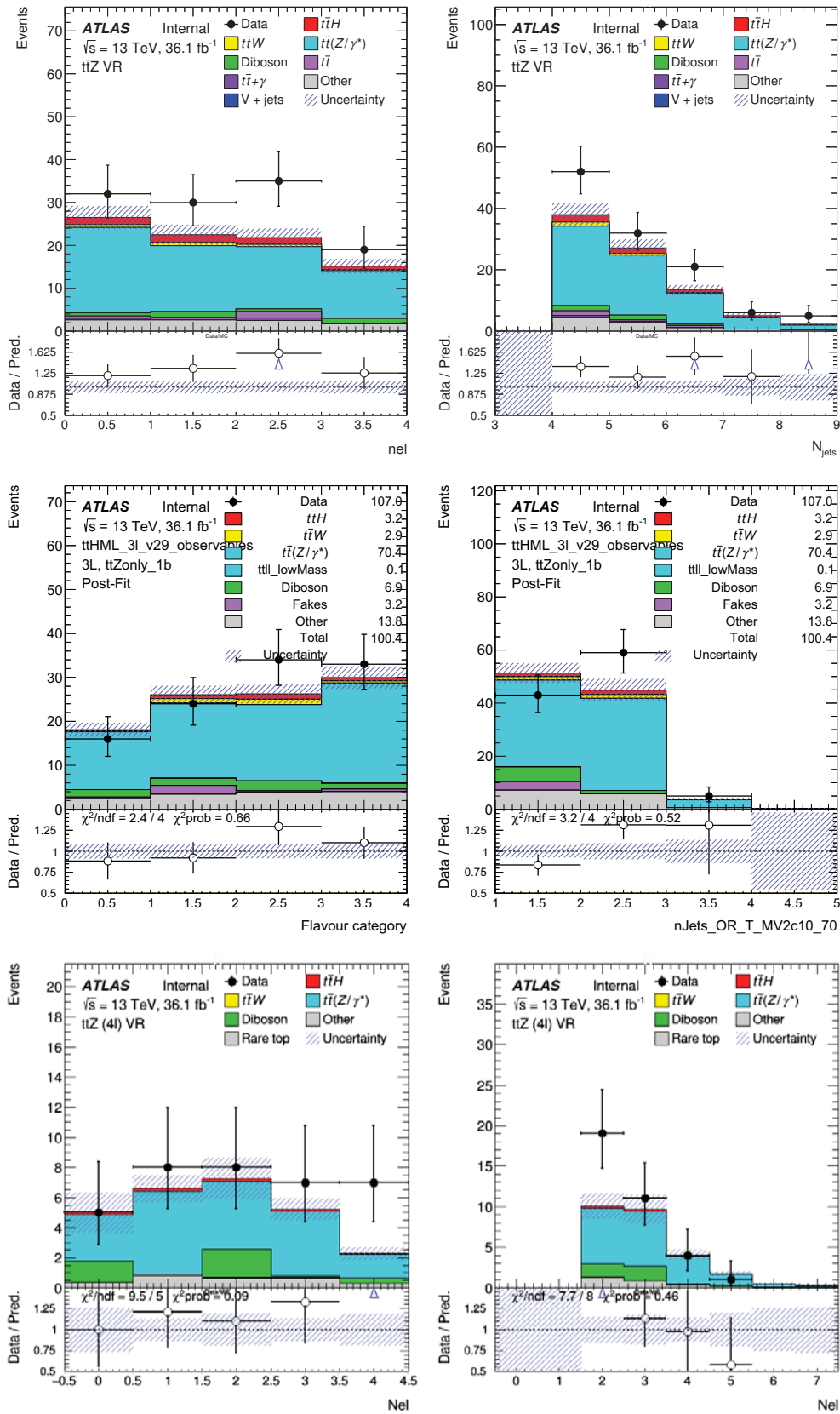


FIGURE A.3 – Comparaison prédiction et données dans des régions enrichie en  $t\bar{t}W$  par sélections rectangulaires dans le canal  $2\ell SS0\tau$  (haut) et par BDT dans le canal  $3\ell 0\tau$  (bas).

### A.3 Validation du processus $t\bar{t}Z$ dans les canaux $3\ell 0\tau$ et $4\ell$

Les comparaisons des prédictions et des données dans des régions enrichies en leptons non-prompt sont présentées sur les Figures A.4 pour les canaux  $3\ell 0\tau$  et  $4\ell$ .


 FIGURE A.4 – Nombre d'électrons (gauche) et de jets (droite) pour une région enrichie en  $t\bar{t}Z$  définie par une sélection rectangulaire avec 3 leptons (haut), par BDT avec 3 leptons (centre) et avec 4 leptons (bas).



# Incertitudes systématiques expérimentales appliquées sur les objets dans l'analyse à $36.1 \text{ fb}^{-1}$

---

Les incertitudes instrumentales sont appliquées sur tous les objets reconstruits de l'analyse. Toutes les incertitudes sont évaluées événement par événement. Elles portent sur les efficacités du système de déclenchement et de l'étiquetage des jets de  $b$ , la reconstruction et l'identification des leptons, la calibration de l'énergie des jets.

Ces incertitudes peuvent être appliquées de deux manières, soit par repondération du poids appliqué à l'événement, soit par un renormalisation de l'énergie ou de l'impulsion de l'objet ( $p_T$  Correction). Les incertitudes instrumentales présentes dans cette analyse sont regroupées dans les Tables B.1, B.2 et B.3. Pour chacune de ces systématiques, la méthode d'application est indiquée.

Une autre incertitude instrumentale prise en compte est celle sur la mesure de la luminosité intégrée lors des collisions. Elle est de 2.1% et est appliquée à tous les processus simulés par Monte-Carlo y compris le signal.

Une dernière incertitude porte sur la repondération des événements simulés afin que la modélisation l'effet d'empilement reproduise l'effet observé. Cette incertitude, `PRW_DATASF`, est elle aussi appliquée sur tous les processus simulés.

<b>Systématiques expérimentales sur les leptons</b>			
Type	Description	Nom de la systématique	Application
<b>Système de déclenchement</b>			
Facteurs d'échelle	Efficacité du système	lepSFTrigTight_MU(EL)_SF_Trigger_STAT(SYST)	Repondération
<b>Muons</b>			
Efficacités	Reconstruction et Identification	lepSFObjTight_MU_SF_ID_STAT(SYST)	Repondération
	Isolation	lepSFObjTight_MU_SF_Isol_STAT(SYST)	Repondération
	Track To Vertex Association	lepSFObjTight_MU_SF_TTVA_STAT(SYST)	Repondération
Échelle sur $p_T$	Échelle sur $p_T$	MUONS_SCALE	$p_T$ Correction
Résolution	Détecteur Interne	MUONS_ID	$p_T$ Correction
	Résolution en énergie Spectromètre à muon	MUONS_MS	$p_T$ Correction
<b>Electrons</b>			
Efficacités	Reconstruction	lepSFObjTight_EL_SF_ID	Repondération
	Identification	lepSFObjTight_EL_SF_Reco	Repondération
	Isolation	lepSFObjTight_EL_SF_Isol	Repondération
Facteur d'échelle	Échelle d'énergie	EG_SCALE_ALL	Ener. Correction
Résolution	Résolution en énergie	EG_RESOLUTION_ALL	Ener. Correction
<b>Taus Hadroniques</b>			
Efficacités	Reconstruction	tauSFLoose_TAU_SF_RECO_TOTAL	Repondération
	Identification BDT	tauSFTight_TAU_SF_JETID_TOTAL	Repondération
	Electron Veto BDT	tauSFTight_TAU_SF_ELEOLR_TOTAL	Repondération
Facteur d'échelle	Échelle en $p_T$	TAUS_TRUEHADTAU_SME_TES_MODEL	$p_T$ Correction
		TAUS_TRUEHADTAU_SME_TES_DETECTOR	$p_T$ Correction
		TAUS_TRUEHADTAU_SME_TES_INSITU	$p_T$ Correction

TABLE B.1 – Incertitudes systématiques instrumentales appliquées aux leptons (tau hadronique compris).

Incertitudes instrumentales sur les Jets et la MET			
Type	Origine	Nom	Application
<b>Jets</b>			
Jet Vertex Tagger		JVT	Repondération
Échelle d'énergie	méthode de calibration ( <i>e.g.</i> In-Situ)	JET_21NP_	$p_T$ Correction
		JET_EffectiveNP_1,2,3,4,5,6,7,8restTerm	$p_T$ Correction
$\eta$ inter-calibration		JET_EtaIntercalibration_Modelling	$p_T$ Correction
		JET_EtaIntercalibration_NonClosure	$p_T$ Correction
		JET_EtaIntercalibration_TotalStat	$p_T$ Correction
Jets à haut $p_T$		JET_SingleParticle_HighPt	$p_T$ Correction
Effet d'empilement		JET_Pileup_OffsetNPV	$p_T$ Correction
		JET_Pileup_OffsetMu	$p_T$ Correction
		JET_Pileup_PtTerm	$p_T$ Correction
		JET_Pileup_RhoTopology	$p_T$ Correction
Test de fermeture		JET_PunchThrough_MC15	$p_T$ Correction
Saveur		JET_Flavor_Response	$p_T$ Correction
		JET_BJES_Response	$p_T$ Correction
		JET_Flavor_Composition	$p_T$ Correction
Résolution		JET_JER_SINGLE_NP	Repondération
<b>MET</b>			
Soft Tracks Terms	Résolution	MET_SoftTrk_ResoPerp	$p_T$ Correction
	Résolution	MET_SoftTrk_ResoPara	$p_T$ Correction
	Échelle	MET_SoftTrk_ScaleUp	$p_T$ Correction
	Échelle	MET_SoftTrk_ScaleDown	$p_T$ Correction

TABLE B.2 – Incertitudes systématiques appliquées sur la calibration de l'énergie des jets et de l'énergie transverse manquante.

Systématiques instrumentales sur l'étiquetage des jets de $b$		
Type	Origine	Nom
	b-tags	
Facteurs d'échelle	Efficacité sur continuuous b-tagger sur les jets étiquetés $b$ (bins en $\eta$ )	ATLAS_FTAG_B0-29
	Efficacité sur continuuous b-tagger sur les jets étiquetés $c$ (bins en $\eta$ )	ATLAS_FTAG_C0-14
	Efficacité sur continuuous b-tagger sur les jets issus de quarks légers dans des bins en $\eta$ et $p_T$	ATLAS_FTAG_Light0-79
	Continuous b-tagger Efficacité d'extrapolation	ATLAS_FTAG_EXTRAP_C

TABLE B.3 – Incertitudes systématiques appliquées sur la reconstruction et l'étiquetage des jets provenant de hadrons beaux.



# Bibliographie

- [1] C. Patrignani *et al.*, “Review of Particle Physics,” *Chin. Phys.*, vol. C40, no. 10, p. 100001, 2016. (Cité en pages 4, 10 et 14.)
- [2] A. Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,” *Phys.Lett.*, vol. B716, pp. 1–29, 2012. (Cité en page 4.)
- [3] C. collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,” *Phys.Lett.*, vol. B716, pp. 30–61, 2012. (Cité en page 4.)
- [4] S.Weinberg, *The Quantum Theory of Fields*. vol. 1. Cambridge University Press, 1995. (Cité en page 4.)
- [5] E.Noether, “Invariante variationsprobleme,” *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, pp. 235–257, 1918. (Cité en page 5.)
- [6] R. H. D. H. R. H. C.S. Wu, E. Ambler, “Experimental Test of Parity Conservation in Beta Decay,” *Phys. Rev.*, vol. 105, pp. 1413–1415, 1957. (Cité en page 6.)
- [7] S. Glashow, “Partial symmetries of weak interactions,” *Nuclear Physics*, vol. 22, pp. 579–588, 1961. (Cité en page 6.)
- [8] S. W. J. Goldstone, A. Salam, “Broken symmetries,” *Phys. Rev.*, vol. 127, pp. 965–970, 1962. (Cité en page 6.)
- [9] A. M. F. Halzen, *Quarks and Leptons : An Introduction Course in Modern Particle Physics*. John Wiley & Sons, 1984. (Cité en page 6.)
- [10] U. Collaboration, “Experimental observation of lepton pairs of invariant mass around 95  $\text{gev}/c^2$  at the cern sps collider,” *Physics Letters B*, vol. 126, no. 5, pp. 398–410, 1983. (Cité en page 7.)
- [11] U. Collaboration, “Experimental observation of events with large missing transverse energy accompanied by a jet or a photon (s) in  $p\bar{p}$  collisions at  $\sqrt{s} = 540 \text{ gev}$ ,” *Physics Letters B*, vol. 139, no. 1-2, pp. 115–125, 1984. (Cité en page 7.)
- [12] U. Collaboration, “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern  $p\bar{p}$  collider,” *Physics Letters B*, vol. 122, no. 5-6, pp. 476–485, 1983. (Cité en page 7.)
- [13] U. Collaboration, “Evidence for  $z^0 \rightarrow e^+e^-$  at the cern  $p\bar{p}$  collider,” *Physics Letters B*, vol. 129, no. 1-2, pp. 130–140, 1983. (Cité en page 7.)
- [14] R. B. F. Englert, “Broken symmetry and the mass of gauge vector mesons,” *Phys. Rev. Lett*, vol. 13, pp. 321–323, 1964. (Cité en page 7.)
- [15] P. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys. Rev. Lett*, vol. 13, pp. 508–509, 1964. (Cité en page 7.)
- [16] N. Cabibbo, “Unitary symmetry and leptonic decays,” *Phys. Rev. Lett*, vol. 10, pp. 531–533, 1963. (Cité en page 9.)
- [17] T. M. M. Kobayashi, “Cp-violation in the renormalizable theory of weak interaction,” *Progress of Theoretical Physics*, vol. 49, no. 2, pp. 652–657, 1973. (Cité en page 9.)
- [18] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, “Evidence for the  $2\pi$  decay of the  $k_2^0$  meson,” *Phys. Rev. Lett.*, vol. 13, pp. 138–140, 1964. (Cité en page 9.)
- [19] C. collaboration, “Study of the Mass and Spin-Parity of the Higgs Boson Candidate via its Decay to Z Boson Pairs,” *Phys. Rev. Lett.*, vol. 110, p. 081803, 2013. (Cité en page 10.)

- [20] A. Collaboration, “Evidence for the spin-0 nature of the Higgs boson using ATLAS data,” *Physics Letters B*, vol. 726, no. 1–3, pp. 120–144, 2013. (Cit  en page 10.)
- [21] A. collaboration, “Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at  $\sqrt{s} = 7$  and 8 TeV in the ATLAS experiment,” *The European Physical Journal C*, vol. 76, no. 1, p. 6, 2016. (Cit  en pages 11 et 12.)
- [22] M. S. F. Bezrukov, “Why should we care about the top quark Yukawa coupling,” *Journal of Experimental and Theoretical Physics*, vol. 120, no. 3, pp. 335–443, 2015. (Cit  en pages 12 et 13.)
- [23] P. B. L. Evans, “LHC Machine,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08001, 2008. (Cit  en page 17.)
- [24] *LEP design report*. Geneva : CERN, 1984. (Cit  en page 18.)
- [25] A. Collaboration, “The ALICE experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08002, 2008. (Cit  en page 18.)
- [26] A. Collaboration, “The ATLAS experiment at the CERN Large Hadron Collider,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008. (Cit  en pages 18, 25, 26, 27 et 29.)
- [27] C. Collaboration, “The CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008. (Cit  en page 18.)
- [28] L. Collaboration, “The LHCb experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008. (Cit  en page 18.)
- [29] C. De Melis, “The CERN accelerator complex. Complexe des acc el erateurs du CERN,” 2016. (Cit  en page 18.)
- [30] “Atlas luminosity results public web page.” (Cit  en pages 19 et 20.)
- [31] *ATLAS inner detector : Technical Design Report, 1*. Geneva : CERN, 1997. (Cit  en page 22.)
- [32] “Atlas inner detector illustration web page.” (Cit  en page 23.)
- [33] *ATLAS pixel detector : Technical Design Report*. Geneva : CERN, 1998. (Cit  en page 22.)
- [34] *ATLAS central solenoid : Technical Design Report*. Geneva : CERN, 1997. (Cit  en page 24.)
- [35] “Atlas calorimeter illustration web page.” <http://atlasexperiment.org/photos/calorimeters-combined-barrel.html>. (Cit  en page 24.)
- [36] *ATLAS liquid-argon calorimeter : Technical Design Report*. Geneva : CERN, 1996. (Cit  en page 24.)
- [37] *ATLAS tile calorimeter : Technical Design Report*. Geneva : CERN, 1996. (Cit  en page 25.)
- [38] A. Collaboration, “Readiness of the ATLAS Tile Calorimeter for LHC collisions,” *Eur.Phys.J.*, vol. C70, pp. 1191–1236, 2010. (Cit  en pages 27 et 28.)
- [39] *Performance of the ATLAS Trigger System in 2015*. Geneva : CERN, 2016. (Cit  en page 30.)
- [40] K. A. et al., “Design of the front-end analog electronics for the atlas tile calorimeter,” *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 551, no. 2, pp. 469 – 476, 2005. (Cit  en page 32.)
- [41] E. Starchenko *et al.*, “Cesium monitoring system for ATLAS Tile Hadron Calorimeter,” *Nucl. Instrum. Meth.*, vol. A494, pp. 381–384, 2002. (Cit  en page 32.)
- [42] J. Abdallah *et al.*, “The Laser calibration of the Atlas Tile Calorimeter during the LHC run 1,” *JINST*, vol. 11, no. 10, p. T10005, 2016. (Cit  en page 33.)
- [43] P. Adragna *et al.*, “Testbeam studies of production modules of the ATLAS tile calorimeter,” *Nucl. Instrum. Meth.*, vol. A606, pp. 362–394, 2009. (Cit  en page 33.)
- [44] S. Physics, “Navigator laser j40-bl6s-532q.” (Cit  en page 35.)

- [45] G. Sabato and M. Elsing, “ATLAS Fast Physics Monitoring : TADA,” Tech. Rep. ATL-SOFT-PROC-2017-038, Feb 2017. (Cit  en page 47.)
- [46] P. Calafiura and al., “The ATHENA control framework in production, new developments and lessons learned.,” *Computing in High-Energy Physics (CHEP’04)*, pp. 456–458, 2005. (Cit  en page 49.)
- [47] A. B. et al., “General-purpose event generators for LHC physics,” *Phys.Rept.*, vol. 504, pp. 145–233, 2011. (Cit  en page 50.)
- [48] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, “Parton Fragmentation and String Dynamics,” *Phys. Rept.*, vol. 97, pp. 31–145, 1983. (Cit  en page 50.)
- [49] B. R. Webber, “Fragmentation and hadronization,” *Int. J. Mod. Phys.*, vol. A15S1, pp. 577–606, 2000. (Cit  en page 50.)
- [50] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP*, vol. 05, p. 026, 2006. (Cit  en page 51.)
- [51] M. Bahr *et al.*, “Herwig++ Physics and Manual,” *Eur. Phys. J.*, vol. C58, pp. 639–707, 2008. (Cit  en page 51.)
- [52] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1,” *JHEP*, vol. 02, p. 007, 2009. (Cit  en pages 51 et 86.)
- [53] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, “MadGraph 5 : Going Beyond,” *JHEP*, vol. 06, p. 128, 2011. (Cit  en page 51.)
- [54] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, “The MCaNL0 4.0 Event Generator,” 2010. (Cit  en pages 51 et 85.)
- [55] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations : the POWHEG method,” *JHEP*, vol. 11, p. 070, 2007. (Cit  en pages 51 et 86.)
- [56] A. collaboration, “The ATLAS Simulation Infrastructure,” *Eur. Phys. J.*, vol. C70, pp. 823–874, 2010. (Cit  en page 51.)
- [57] S. Agostinelli *et al.*, “GEANT4 : A Simulation toolkit,” *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003. (Cit  en page 51.)
- [58] E. Richter-Was, D. Froidevaux, and L. Poggioli, “ATLFAST 2.0 a fast simulation package for ATLAS,” 1998. (Cit  en page 51.)
- [59] “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data,” Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun 2016. (Cit  en page 52.)
- [60] G. Aad *et al.*, “Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data,” *Eur. Phys. J.*, vol. C74, no. 11, p. 3130, 2014. (Cit  en page 54.)
- [61] “Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at  $\sqrt{s} = 13$  TeV,” Tech. Rep. ATL-PHYS-PUB-2015-015, CERN, Geneva, Jul 2015. (Cit  en page 54.)
- [62] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm,” *JHEP*, vol. 04, p. 063, 2008. (Cit  en page 56.)
- [63] “Monte Carlo Calibration and Combination of In-situ measurements of Jet Energy Scale, Jet Energy Resolutoin and Jet Mass in ATLAS,” Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, 2015. (Cit  en page 57.)
- [64] “Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data,” Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, 2011. (Cit  en page 58.)

- [65] “Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC,” Tech. Rep. ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov 2015. (Cité en page 58.)
- [66] “Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at  $\sqrt{s}=13$  TeV,” Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul 2015. (Cité en page 58.)
- [67] D. Paredes Hernández, *Search for New Physics in events with 4 top quarks in the ATLAS detector at the LHC*. PhD thesis, 2013. (Cité en page 61.)
- [68] G. Cowan, *Statistical Data Analysis*. Oxford science publications. Clarendon Press, 1998. (Cité en page 77.)
- [69] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, “HistFactory : A tool for creating statistical models for use with RooFit and RooStats,” tech. rep., New York U., New York, 2012. (Cité en page 77.)
- [70] R. J. Barlow, *Statistics : A Guide to the Use of Statistical Methods in the Physical Sciences*. Manchester Physics Series. John Wiley & Sons, West Sussex, 1989. (Cité en page 78.)
- [71] A. L. Read, “Presentation of search results : The CL(s) technique,” *J. Phys.*, pp. 2693–2704, 2002. (Cité en page 80.)
- [72] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1,” *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008. (Cité en page 85.)
- [73] D. e. a. de Florian, “Handbook of LHC Higgs Cross Sections : 4. Deciphering the Nature of the Higgs Sector,” Tech. Rep. FERMILAB-FN-1025-T, Oct 2016. (Cité en page 93.)
- [74] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics,” *Eur. Phys. J.*, vol. C71, p. 1554, 2011. (Cité en page 101.)
- [75] D. Cinca, *Etude de la production de paires de quark TOP avec ATLAS au LHC, mesure de la masse du quark TOP*. PhD thesis, 2011. (Cité en page 107.)
- [76] “Search for the Standard Model Higgs boson produced in association with top quarks and decaying into  $b\bar{b}$  in pp collisions at  $\sqrt{s}=13$  TeV with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2016-080, 2016. (Cité en page 141.)
- [77] “Measurement of fiducial, differential and production cross section in the  $H \rightarrow \gamma\gamma$  decay channel with  $13.3 \text{ fb}^{-1}$  of 13 TeV proton-proton collision data, collaboration =,” tech. rep. (Cité en page 141.)
- [78] “Combination of the searches for Higgs boson production in association with top quarks in the  $\gamma\gamma$ , multilepton and  $b\bar{b}$  decay channels at  $\sqrt{s}=13$  TeV with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2016-068, Aug 2016. (Cité en page 141.)
- [79] “Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector,” Tech. Rep. ATLAS-CONF-2017-077, CERN, Nov 2017. (Cité en page 141.)

## Résumé

La découverte du boson de Higgs au LHC en 2012 est la plus récente confirmation de la validité du modèle standard, théorie décrivant les particules élémentaires et leurs interactions fondamentales. De nombreuses études ont été mises en place pour étudier les différentes caractéristiques de cette particule récemment découverte. Ce travail se concentre sur la recherche du processus  $t\bar{t}H$  dans l'expérience ATLAS, pour réaliser une première mesure directe du couplage de Yukawa du boson de Higgs au quark top, paramètre important dans les modèles de physique au-delà du modèle standard.

La première partie de ce travail a porté sur l'automatisation de l'étalonnage du calorimètre hadronique à tuiles du détecteur ATLAS (*TileCal*) par un système laser. Une description de ce système qui permet un étalonnage régulier de toutes les cellules du calorimètre est tout d'abord présentée. Ensuite, l'algorithme d'automatisation de ce processus d'étalonnage, qui a été écrit lors de ce travail, est décrit. Le but final de cet algorithme est de faciliter et accélérer la correction des canaux dont le gain dérive.

Une deuxième partie concerne la recherche du processus  $t\bar{t}H$  par l'étude des états finals multi-leptons et particulièrement avec deux leptons de même charge électrique et au moins 4 jets dont 1 étiqueté b. L'estimation des bruits de fond instrumentaux et le traitement statistique réalisés sont décrits en détail dans le document pour deux versions de l'analyse. Une première version correspondant aux données de l'année 2015 et du début 2016, soit une luminosité intégrée de  $13.2\text{fb}^{-1}$  de données, aboutit à une précision insuffisante pour conclure sur la présence du processus  $t\bar{t}H$ . Une deuxième version de l'analyse, optimisée avec l'utilisation d'analyses multivariées sur l'ensemble des données 2015 et 2016, se conclut par une observation du processus  $t\bar{t}H$  lors de la combinaison de l'ensemble des états finals  $t\bar{t}H$ . La signification statistique observée est alors de  $4.2\sigma$  pour une signification statistique attendue de  $3.8\sigma$ . Ce résultat est donc en accord avec la prédiction du modèle standard.

Mots-clés : modèle standard, couplage de Yukawa au quark top,  $t\bar{t}H$ , ATLAS, LHC, calorimètre hadronique à tuiles, étalonnage par système laser, leptons de même charge

## Abstract

Discovery of Higgs boson at LHC in 2012 is the most recent confirmation of the validity of Standard Model, theory describing elementary particles and their interactions. Many analysis now target the extraction of properties of the newly-discovered particle. A direct measurement in the ATLAS experiment of the top Yukawa coupling, one of these properties, is targeted in this work through  $t\bar{t}H$  process. This coupling is of particular interest because of its strong sensitivity to New Physics.

The first part of the work is about the automation of the calibration of the hadronic tile calorimeter of the ATLAS detector. A detailed description of the laser system used for a regular calibration of the calorimeter is done as well as of the calibration itself. Then more details on the algorithm written for the automation of the calibration are given. The final goal of this algorithm is to ease and fasten the calibration of channels with gain variation.

The second part is dedicated to the search for  $t\bar{t}H$  process through multilepton final states with emphasis on final state with two same-sign leptons, at least four jets and at least 1 b-tagged jet. Estimation of reducible backgrounds and statistical treatment of the analysis are detailed. A first version of the analysis with  $13.2fb^{-1}$ , corresponding to 2015 and mid-2016 LHC data, ends with a final precision too low to extract any conclusion on the  $t\bar{t}H$  process. In a second version of the analysis, improvements are made using multivariate analysis and adding more signal regions. The results from the multilepton analysis is combined with results from other  $t\bar{t}H$  analysis, targeting other Higgs decay, to attain a final observed sensitivity of  $4.2\sigma$ . Thus an evidence for  $t\bar{t}H$  production can be claimed from this combination. The final results give good agreement with Standard Model prediction.

Keywords : Standard Model, top Yukawa coupling,  $t\bar{t}H$ , ATLAS, LHC, tile hadronic calorimeter, laser calibration, same-sign leptons