

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Physique Subatomique et Astroparticules**

Arrêté ministériel : 7 août 2006

Présentée par

**Maud Schwoerer**

Thèse dirigée par **Isabelle Wingerter-Seez**  
codirigée par **Marco Delmastro et Jessica Lévêque**

préparée au sein du **Laboratoire d'Annecy-le-vieux de Physique des Particules**  
et de l'**École doctorale de Physique de Grenoble**

## Étude des états finals diphoton dans l'expérience ATLAS au LHC :

**Mesure de sections efficaces différentielles, découverte d'une nouvelle résonance dans la recherche du boson de Higgs et étude de ses propriétés.**

Thèse soutenue publiquement le **27 Septembre 2013**,  
devant le jury composé de :

**M. Gregorio Bernardi**

LPHNE - Paris, Examinateur

**M. Dave Charlton**

University of Birmingham, Examinateur

**M. Marco Delmastro**

LAPP - Annecy-le-vieux, Co-Directeur de thèse

**M. Jean-Philippe Guillet**

LAPTh - Annecy-le-vieux, Examinateur

**M. Gautier Hamel De Monchenault**

CEA - Saclay, Rapporteur

**M. Yannis Karyotakis**

LAPP - Annecy-le-vieux, Président

**Mme. Jessica Lévêque**

LAPP - Annecy-le-vieux, Co-Directeur de thèse

**M. Guillaume Unal**

CERN, Rapporteur





*A Jean L.*



# Remerciements

Je voudrais en premier lieu remercier mes rapporteurs, Guillaume Unal et Gautier Hamel de Monchenault, merci d'avoir pris le temps de lire ma thèse, vos corrections et remarques m'ont été d'une aide précieuse. Merci également à tous les autres membres du jury Gregorio Bernardi, Dave Charlton, Jean-Philippe Guillet et Yannis Karyotakis, c'était un grand honneur de vous avoir pour juger de la qualité de mon travail.

Merci à Yannis Karyotakis, notre directeur et Jean-Pierre Lees pour m'avoir accueillie au LAPP et merci à l'ensemble du personnel administratif en particulier Marie-Claude Lacombe, Brigitte Putanier et Chantal Vallee pour votre efficacité et votre gentillesse. Un grand merci également à Luis Hervas, qui avec Isabelle, a rendu possible mon stage de fin d'études au CERN, le début de cette aventure.

Merci à l'ensemble du groupe ATLAS LAPP dirigé d'abord par Lucia Di Ciaccio puis par Stéphane Jezequel. Entre réunions de groupe et sorties raquettes (sans raquettes d'ailleurs), cela a été un réel plaisir et très stimulant de faire partie de votre troupe durant trois ans. Je tiens à remercier plus spécifiquement l'équipe Higgs du LAPP avec laquelle j'ai eu la chance de travailler, toujours dans la bonne humeur et souvent avec des croissants. Merci à Nicolas Berger pour avoir apporté des réponses à mes très (trop) nombreuses questions. Merci à Rémi Lafaye pour son soutien et sa disponibilité, j'ai beaucoup appris en travaillant à tes côtés, tu m'a aidé tout au long de ma thèse, aussi bien pour la mesure de section efficace que pour l'analyse  $H \rightarrow \gamma\gamma$ .

Merci également à toutes celles et ceux avec qui j'ai pu travailler, discuter et apprendre, je ne vais malheureusement pas pouvoir citer tout le monde mais un merci particulier à Paolo Inngo, Benjamin Trocmé et Olivier Simard pour le travail sur le détecteur, à Lydia Roos, Sandrine Laplace, Giovanni Marchiori, Jean-Baptiste de Vivie, Hongtao et Olivier pour l'analyse diphoton et pour l'analyse  $H \rightarrow \gamma\gamma$ .

Merci à Eric Pillon et Jean-Philippe Guillet pour nos échanges dans le cadre des Houches, grâce à vous j'ai pu entrevoir le monde merveilleux des théoriciens et cela a été très enrichissant.

Bien entendu ce temps passé au LAPP et au CERN n'aurait pas été le même sans les autres doctorants et post-doctorants. Merci à Louis, Julien, Mathieu, Emmanuel et Olivier c'est vous les premiers, durant mon stage au CERN, qui m'avez confortée dans l'idée de poursuivre une thèse dans ATLAS (l'usage d'un smiley, bien que peu recommandé dans une thèse, serait ici justifié). Un immense merci à Dimitra, Ludovica, Mayuko, Oanh, vous avez été indispensables à ma survie sociale durant ces trois années. Merci également à Elisabeth, cela a été un vrai plaisir de discuter et de travailler avec toi, viel Spass pour la suite. Je remercie également Emma, Hass, Joan, Roger, Guillaume, Laurent et Guilhem, je suis ravie d'avoir croisé votre route. Enfin merci aux "nouveaux" doctorants, Houry et Zuzana, je vous souhaite le meilleur pour la fin de votre thèse.

Merci à Christelle Roy pour m'avoir accueillie à Subatech lors de mon stage au collège, c'est

toi la première qui m'a donné envie de faire une thèse en physique des particules, il y a une douzaine d'années maintenant.

Je voudrais remercier de tout coeur mes directeurs de thèse, Isabelle Wingerter-Seez, Jessica Lévêque et Marco Delmastro. Vous m'avez soutenue, aiguillée et vous m'avez énormément appris, scientifiquement bien-sûr mais humainement aussi. Il y a de toute évidence des moments plus durs que d'autres au cours d'un doctorat et dans cette immense collaboration qu'est ATLAS, mais grâce à vous j'ai réellement l'impression de m'être amusée à faire de la physique durant trois années, et je pense que c'est essentiel. J'ai un grand nombre de "merci" à vous adresser mais parmi eux merci tout d'abord pour votre confiance et votre exigence, merci également pour votre disponibilité, pour votre gentillesse et votre bonne humeur, j'ai eu énormément de chance de vous avoir tous les trois.

Un grand merci à mes amis de longue date, Caro et la team MNT en particulier, merci d'avoir fait le déplacement, je me rappellerai longtemps vos petites têtes dans l'amphithéâtre ce jour-là (et félicitations, aucun d'entre vous n'a piqué du nez).

Merci à ma famille, mes grands frères, mes parents. On ne choisit pas sa famille dit-on, mais dans mon cas je n'aurais pu faire meilleur choix. Dominique et Didier merci pour votre soutien, vos conseils et votre estime, vous avez rempli votre rôle à merveille.

Enfin, Alban, j'aimerais te remercier pour ta constance, ta bienveillance, pour ton inépuisable sourire, pour ton sens des réalités. J'aimerais également m'excuser, m'excuser pour les matins trop matinaux, les soirées trop tardives et les voyages trop longtemps reportés. Tu m'as aidé bien plus que je ne saurais l'avouer.

# Table des matières

<b>Remerciements</b>	<b>iii</b>
<b>1 Introduction théorique</b>	<b>3</b>
1.1 Le Modèle Standard . . . . .	4
1.2 Le boson de Higgs . . . . .	4
1.2.1 Le mécanisme de Higgs . . . . .	5
1.2.2 Contraintes théoriques sur la masse . . . . .	6
1.2.3 Contraintes expérimentales sur la masse . . . . .	7
1.2.3.1 Contraintes directes . . . . .	7
1.2.3.2 Contraintes indirectes . . . . .	8
1.2.4 Les modes de désintégration . . . . .	10
1.2.5 Les modes de production . . . . .	11
1.2.6 Détermination des propriétés . . . . .	13
1.3 Événements diphoton directs au LHC . . . . .	13
1.3.1 Les processus de production . . . . .	14
1.3.2 QCD perturbative . . . . .	15
1.3.3 Théorème de factorisation et fonctions de distribution des partons . . . . .	16
1.3.4 Bruits de fond réductible . . . . .	17
1.3.5 Définition des photons isolés . . . . .	17
1.3.6 Précédentes mesures . . . . .	18
1.4 Prédictions théoriques . . . . .	18
1.4.1 Production du boson de Higgs . . . . .	19
1.4.2 Production de paires de photons . . . . .	21
<b>2 Le LHC et le détecteur ATLAS</b>	<b>23</b>
2.1 Description du LHC . . . . .	24
2.1.1 Fonctionnement nominal . . . . .	24
2.1.2 Luminosité et événements d’empilement . . . . .	25
2.2 Le détecteur ATLAS . . . . .	27
2.2.1 Vue d’ensemble . . . . .	27
2.2.2 Le détecteur interne . . . . .	29
2.2.3 Les calorimètres . . . . .	31
2.2.3.1 Le calorimètre électromagnétique (EM) à argon liquide (LAr) . . . . .	32
2.2.3.2 Le calorimètre hadronique . . . . .	34
2.2.3.3 Le calorimètre avant . . . . .	35
2.2.4 Le système de déclenchement . . . . .	35

<b>3</b>	<b>Identification des cellules bruyantes du calorimètre à argon liquide</b>	<b>37</b>
3.1	Acquisition des données dans ATLAS	38
3.2	Chaîne d'acquisition du calorimètre à argon liquide	38
3.3	Monitoring des cellules bruyantes du calorimètre	41
3.4	Traitement des cellules bruyantes du calorimètre	44
<b>4</b>	<b>Les photons dans ATLAS</b>	<b>47</b>
4.1	Reconstruction des photons	48
4.1.1	Reconstruction des clusters électromagnétiques	48
4.1.2	Photons convertis	48
4.2	Calibration des clusters électromagnétiques	49
4.2.1	Calibration Monte-Carlo	49
4.2.2	Calibration in-situ	50
4.3	Isolation	52
4.3.1	Isolation calorimétrique	52
4.3.2	Isolation des traces	53
4.4	Identification	54
4.4.1	Les variables discriminantes	54
4.4.2	Les sélections LOOSE et TIGHT	57
4.4.3	Efficacité d'identification	57
<b>5</b>	<b>Mesure de sections efficaces différentielles de production de paires de photons</b>	<b>63</b>
5.1	Sélection des événements	64
5.1.1	Les données et le Monte-Carlo	64
5.1.2	La présélection des événements	64
5.1.3	Choix de la paire de photons	65
5.1.3.1	Candidats signal	66
5.1.3.2	Candidats bruit de fond	66
5.2	Estimation du bruit de fond $\gamma$ -jet et jet-jet	67
5.2.1	Description de la méthode $2 \times 2D$ sideband	68
5.2.2	Paramètres d'entrée de la méthode $2 \times 2D$ sideband	70
5.2.3	Incertitudes systématiques	72
5.2.3.1	Définition de la région de contrôle du bruit de fond	72
5.2.3.2	Efficacité d'isolation	73
5.2.3.3	Efficacité d'identification	77
5.2.3.4	Erreurs systématiques en fonction des variables cinématiques	78
5.2.4	Décomposition en événements $\gamma\gamma$ , $\gamma$ -jet, jet- $\gamma$ et jet-jet	81
5.3	Estimation du bruit de fond électron	84
5.3.1	Description de la méthode	84
5.3.2	Extraction de la contamination électron	86
5.4	Déconvolution	87
5.4.1	Description de la méthode	87
5.4.1.1	Incertitudes systématiques	88
5.4.2	Efficacité inclusive	89
5.5	Sections efficaces	90
5.5.1	Sections efficaces mesurées	90
5.5.2	Prédictions théoriques	92
5.5.2.1	Prédictions à ordre fixe	92

5.5.2.2	Générateur Monte-Carlo avec parton shower . . . . .	93
5.5.3	Comparaison entre prédictions théoriques et données . . . . .	95
5.6	Conclusion . . . . .	96
<b>6</b>	<b>Observation d'une résonance dans la recherche du boson de Higgs Modèle Standard dans le canal diphoton</b>	<b>101</b>
6.1	Sélection des données, juillet 2012 . . . . .	102
6.1.1	Sélection inclusive des événements . . . . .	102
6.1.2	Composition du bruit de fond . . . . .	103
6.2	Recherche d'une résonance . . . . .	105
6.2.1	Catégorisation . . . . .	105
6.2.2	Modélisation et paramétrisation du bruit de fond . . . . .	108
6.2.3	Modélisation et paramétrisation du signal . . . . .	110
6.2.4	Erreurs systématiques . . . . .	113
6.2.5	Tests statistiques . . . . .	113
6.3	Observation d'une nouvelle résonance dans le canal diphoton . . . . .	115
<b>7</b>	<b>Etudes des couplages de la nouvelle résonance</b>	<b>117</b>
7.1	Sélection des données, mars 2013 . . . . .	118
7.1.1	Sélection inclusive des événements . . . . .	118
7.1.2	Composition du bruit de fond . . . . .	118
7.2	Catégories VH . . . . .	119
7.2.1	Catégorie leptonique . . . . .	120
7.2.2	Catégorie avec énergie transverse manquante . . . . .	120
7.2.3	Catégorie hadronique . . . . .	120
7.3	Catégorie VBF . . . . .	121
7.3.1	Optimisation des coupures . . . . .	121
7.3.2	Analyse multi-variée . . . . .	123
7.3.3	Comparaison entre l'analyse multivariée et l'analyse avec coupures . . . . .	126
7.4	Autres catégories . . . . .	129
7.5	Extraction du signal . . . . .	130
7.5.1	Modélisation du bruit de fond et du signal . . . . .	130
7.5.2	Incertitudes systématiques . . . . .	131
7.5.2.1	Incertitudes sur le nombre d'événements de signal attendus . . . . .	131
7.5.2.2	Incertitudes sur la résolution en masse du signal . . . . .	133
7.5.2.3	Incertitudes liées aux migrations entre catégories . . . . .	133
7.5.2.4	Incertitudes sur la mesure de la masse . . . . .	135
7.5.3	Procédure statistique . . . . .	137
7.6	Résultats . . . . .	138
7.6.1	$p_0$ et force du signal . . . . .	138
7.6.2	Force du signal par mode de production . . . . .	139
7.6.3	$p_0$ pour le mode de production VBF . . . . .	140
7.7	Extraction des couplages . . . . .	141
	<b>Annexes</b>	<b>147</b>
<b>A</b>	<b>Diagrammes de Feynman pour la production du boson de Higgs par fusion de gluons</b>	<b>149</b>

<b>B</b>	<b>Corrélations des variables d'identification et d'isolation des candidats photons</b>	<b>150</b>
<b>C</b>	<b>Equations utilisées pour la méthode <math>2 \times 2D</math> sideband</b>	<b>153</b>
<b>D</b>	<b>Isolation Frixione, critère discret et continu</b>	<b>156</b>
<b>E</b>	<b>Résumé des incertitudes systématiques pour l'analyse <math>H \rightarrow \gamma\gamma</math> de juillet 2012</b>	<b>158</b>
<b>F</b>	<b>Expressions du nombre d'événements de signal par mode de production et par catégorie pour l'analyse <math>H \rightarrow \gamma\gamma</math> de mars 2013</b>	<b>160</b>
	<b>Bibliographie</b>	<b>167</b>
	<b>Résumé</b>	<b>175</b>
	<b>Abstract</b>	<b>176</b>

# Introduction

Le 4 juillet 2012, après à peine plus de deux ans d'exploitation du LHC, le CERN annonçait la découverte d'une nouvelle particule compatible avec le boson de Higgs prédit par le Modèle Standard.

De la conception à l'analyse des données en passant par la construction du détecteur, sa mise en route, la surveillance de la prise des données et de leur qualité, la compréhension et calibration du détecteur, chaque étape a été indispensable pour permettre à ATLAS d'accéder à cette découverte historique.

Le canal diphoton a joué/joue un rôle essentiel dans la recherche, la découverte et l'étude des propriétés de ce nouveau boson dans ATLAS. Ce canal est largement dominé par la production de paires de photons non résonnantes. L'étude et la compréhension de ce bruit de fond procurent des bases solides qui permettent de construire une analyse rigoureuse dans le canal  $H \rightarrow \gamma\gamma$ . Elles présentent également en elles-mêmes un intérêt tout particulier. La mesure de section efficace de production de paires de photons au LHC avec une statistique conséquente nous permet de tester les prédictions de QCD perturbative et de raffiner les générateurs Monte-Carlo dans des régions de l'espace des phases où les données sont aujourd'hui encore mal comprises.

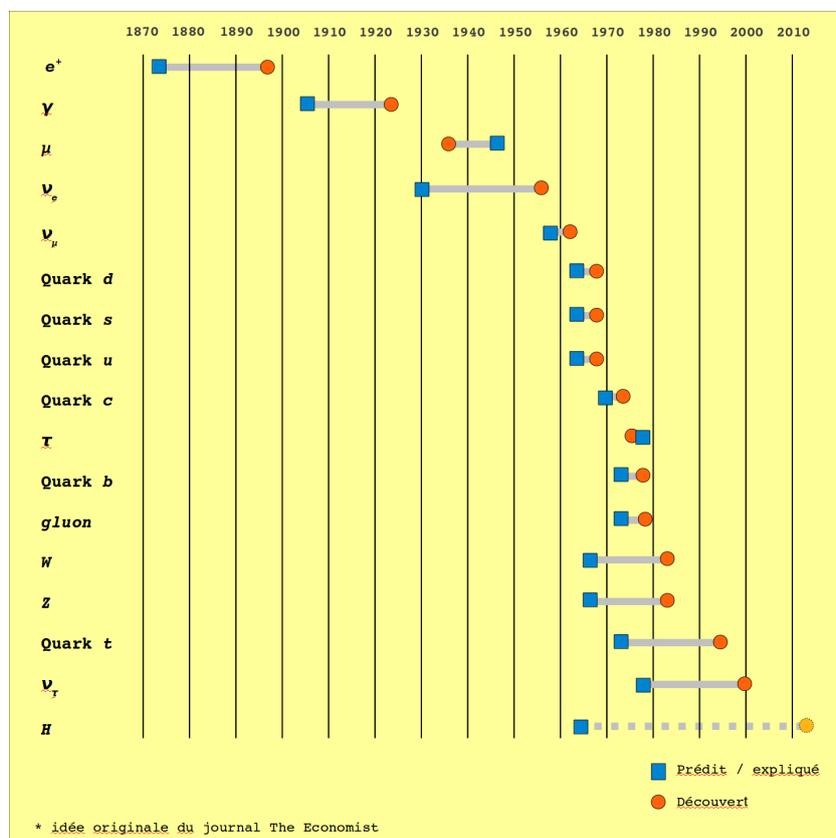
Après avoir posé le cadre théorique au premier chapitre, et expérimental au second chapitre avec la description du LHC et d'ATLAS, nous aborderons la problématique de la qualité des données dans ATLAS à travers la surveillance des cellules bruyantes du calorimètre. Nous détaillerons ensuite la reconstruction, la calibration et l'identification des photons mises en oeuvre dans ATLAS. La qualité des données assurée et la reconstruction des photons validée, il est alors possible de procéder à la mesure de sections efficaces différentielles de paires de photons présentée au chapitre 5. Les coupures appliquées lors de l'identification des photons réduisent considérablement le bruit de fond issu de la reconstruction des jets en photons, mais elles ne sont pas suffisantes pour permettre d'isoler directement le signal diphoton. Nous détaillerons particulièrement l'extraction du nombre d'événements  $\gamma\gamma$  nécessaire à la mesure de section efficace et également utilisée dans la recherche du boson de Higgs. Le chapitre 6 présentera l'observation d'une nouvelle résonance dans le canal diphoton et le dernier chapitre sera dédié à l'étude des propriétés de cette particule. Nous développerons les choix faits pour augmenter la précision sur la mesure des taux de production, nécessaire pour juger du caractère standard ou non de ce boson.



# Chapitre 1

## Introduction théorique

Le Modèle Standard a, jusqu'à aujourd'hui, permis une description très précise des particules élémentaires et de leurs interactions. Basé sur des idées théoriques et des faits expérimentaux, de la découverte de l'électron en 1897 à la découverte d'une particule compatible avec le boson de Higgs en 2012, il représente le travail de milliers de chercheurs sur plus d'un siècle.



Si la particule récemment découverte est bien le boson de Higgs, elle sera le fruit d'une longue quête, et apportera la dernière pierre à l'édifice du Modèle Standard. Malgré sa robustesse expérimentale, ce modèle souffre toutefois de plusieurs lacunes théoriques. Sur les 26 paramètres de ce modèle, 19 doivent être déterminés expérimentalement. Il ne parvient pas à expliquer l'asymétrie baryonique observée dans l'univers, ni la masse des neutrinos, ni pourquoi les masses des quarks s'étalent sur cinq ordres de grandeur. Il n'inclut pas la gravitation. Il n'explique pas le problème de *naturalité*, les corrections radiatives à la masse de Higgs divergent quadratiquement ; pour maintenir la masse du boson de Higgs inférieure au TeV, il est nécessaire d'ajuster la masse nue sur 16 ordres de grandeur (*fine tuning*).

Tous ces points amènent à penser que le Modèle Standard est incomplet et doit être intégré dans une théorie plus fondamentale.

Nous ferons, en premier lieu, une description succincte du Modèle Standard. Nous nous concentrerons ensuite sur la dernière zone d'ombre de ce modèle, le mécanisme de Higgs responsable de la brisure de symétrie électrofaible. Dans la quête du boson de Higgs au LHC, nous verrons pourquoi le canal  $\gamma\gamma$  est particulièrement intéressant. Nous étudierons par la suite le bruit de fond irréductible dans ce canal : les événements diphoton. Nous verrons en quoi la mesure de section efficace de production de paires de photons est de tout premier intérêt pour la recherche du boson de Higgs mais également pour tester les prédictions de QCD perturbative. Enfin nous donnerons un aperçu des différents outils théoriques mis à notre disposition pour la recherche du boson de Higgs et pour les mesures de sections efficaces de production de paires de photons au LHC.

## 1.1 Le Modèle Standard

Le Modèle Standard est une théorie de jauge basée sur le groupe de symétrie  $SU(3)_{couleur} \otimes SU(2)_{gauche} \otimes U(1)_{hypercharge}$ . Il décrit les interactions entre les particules élémentaires de spin demi-entier (leptons et quarks) et les bosons de jauge de spin entier (photons, bosons  $W$  et  $Z$ , et gluons). Les huit gluons n'interagissent qu'avec les particules colorées, ils véhiculent l'interaction forte dont la portée effective est de l'ordre du fermi (la taille d'un noyau) ; cette interaction, basée sur le groupe de symétrie  $SU(3)_{couleur}$ , est décrite par la théorie de la chromodynamique quantique (QCD). Le photon, lui, est le boson médiateur de l'interaction électromagnétique dont la portée est infinie ; à l'échelle microscopique elle est décrite par la théorie de l'électrodynamique quantique (QED). Enfin, les bosons  $W$  et  $Z$  sont porteurs de l'interaction faible dont la portée est de l'ordre de  $10^{-18}\text{m}$ .

A la fin des années 1960, Glashow, Salam et Weinberg unifient, avec la théorie électrofaible, les forces électromagnétique et faible [1, 2, 3]. Cette théorie est basée sur les symétries  $SU(2)_{gauche}$  et  $U(1)_{hypercharge}$  auxquelles correspondent respectivement les constantes de couplage  $g$  et  $g'$ .

L'ensemble des particules du Modèle Standard est présenté figure 1.1, chacune d'entre elles est caractérisée par des nombres quantiques qui correspondent à des charges conservées sous l'effet de transformations de jauge<sup>1</sup>. Les théories de jauge permettent de décrire les interactions entre les particules élémentaires à partir de la connaissance de la nature du groupe de symétrie sous lequel les équations de mouvements du système restent invariantes [4].

Dans le Modèle Standard il n'existe pas de moyen d'introduire la masse des particules sans briser l'invariance de jauge du Lagrangien. Le mécanisme responsable de cette brisure justifie l'existence d'un boson scalaire, le boson de Higgs.

## 1.2 Le boson de Higgs

Le mécanisme de Higgs fait intervenir une brisure spontanée de la symétrie, les équations de mouvement du système sont invariantes sous une symétrie donnée, mais l'état fondamental, lui, ne l'est pas. Un exemple de brisure spontanée de symétrie avait déjà été mis en évidence avec les matériaux ferromagnétiques en 1885 par Pierre Curie et expliqué en 1928 par le modèle d'Heisenberg [6]. En 1950, Ginzburg et Landau développent le concept de brisure spontanée de

---

1. Soit  $\psi(x)$  un champ quantique, une transformation de jauge locale est telle que  $\psi(x) \rightarrow e^{i\alpha(x)}\psi(x)$ , la phase dépend de la position de la particule

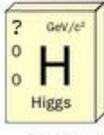
Three generations of matter (fermions)					
	I	II	III		
mass	2.4 MeV/c <sup>2</sup>	1.27 GeV/c <sup>2</sup>	171.2 GeV/c <sup>2</sup>	0	
charge	2/3	2/3	2/3	0	
spin	1/2	1/2	1/2	1	
name	<b>u</b> up	<b>c</b> charm	<b>t</b> top	<b>γ</b> photon	
Quarks	4.8 MeV/c <sup>2</sup>	104 MeV/c <sup>2</sup>	4.2 GeV/c <sup>2</sup>	0	
	-1/3	-1/3	-1/3	0	
	1/2	1/2	1/2	1	
name	<b>d</b> down	<b>s</b> strange	<b>b</b> bottom	<b>g</b> gluon	
Leptons	<2.2 eV/c <sup>2</sup>	<0.17 MeV/c <sup>2</sup>	<15.5 MeV/c <sup>2</sup>	91.2 GeV/c <sup>2</sup>	
	0	0	0	0	
	1/2	1/2	1/2	1	
name	<b>ν<sub>e</sub></b> electron neutrino	<b>ν<sub>μ</sub></b> muon neutrino	<b>ν<sub>τ</sub></b> tau neutrino	<b>Z<sup>0</sup></b> Z boson	
Leptons	0.511 MeV/c <sup>2</sup>	105.7 MeV/c <sup>2</sup>	1.777 GeV/c <sup>2</sup>	80.4 GeV/c <sup>2</sup>	
	-1	-1	-1	±1	
	1/2	1/2	1/2	1	
name	<b>e</b> electron	<b>μ</b> muon	<b>τ</b> tau	<b>W<sup>±</sup></b> W boson	

Figure 1.1 – Liste des particules du Modèle Standard [5].

symétrie dans ce qui deviendra par la suite la théorie de la supraconductivité [7]. C'est dans les années 1960 que Nambu transpose ce concept de la physique de la matière condensée à la physique des particules [8]. Avec Goldstone, ils montrent que l'apparition d'un boson de masse nulle est une conséquence de la brisure spontanée d'une symétrie globale [9]. La prédiction d'un boson de masse nulle n'est cependant pas vérifiée expérimentalement.

Brout, Englert, Higgs, Guralnik, Hagen et Kibble cherchent alors une exception au théorème de Goldstone. Celle-ci semble subvenir lorsque la brisure spontanée de symétrie intervient avec une théorie de jauge locale. En 1964, Brout et Englert découvrent que la rupture de symétrie dans le cadre du théorème de Yang-Mills (dans le cas des groupes non-Abelliens) n'implique pas nécessairement un boson de masse nulle [10]. En 1965, Higgs propose une explication plus complète en introduisant un boson scalaire massif dans le processus proposé par Brout et Englert, le boson de Higgs [11].

### 1.2.1 Le mécanisme de Higgs

L'invariance de jauge  $SU(2)_{\text{gauche}} \otimes U(1)_{\text{hypercharge}}$  implique que les quatre bosons de jauge ainsi que les fermions soient de masse nulle ce qui a été infirmé expérimentalement. Le mécanisme proposé par Higgs, Brout et Englert permet une brisure spontanée de la symétrie  $SU(2)_{\text{gauche}} \otimes U(1)_{\text{hypercharge}}$  et autorise ainsi l'introduction d'un terme de masse dans le lagrangien [10, 11, 12].

On introduit un doublet de champs scalaires complexes  $\phi$  qui se couple aux bosons de jauge massifs :

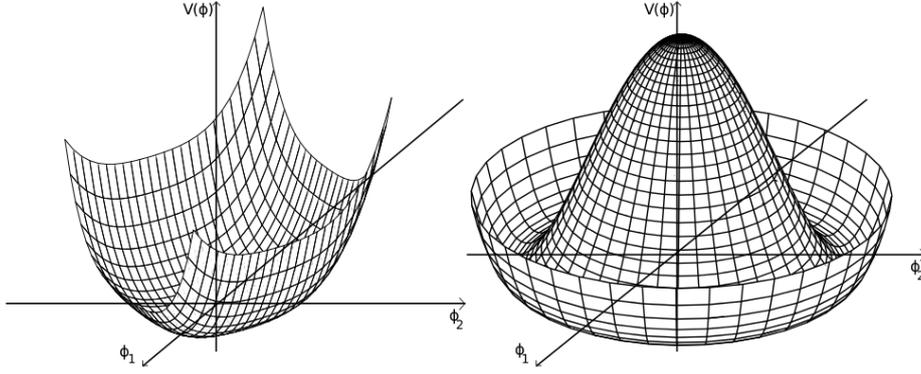
$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+ = \phi_3 + i\phi_4 \\ \phi^0 = \phi_1 + i\phi_2 \end{pmatrix} \quad (1.1)$$

Un terme associé à ce doublet est alors ajouté au lagrangien du Modèle Standard :

$$D_\mu \phi^\dagger D_\mu \phi - V(\phi) \quad (1.2)$$

où  $V$  est le potentiel :

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (1.3)$$



**Figure 1.2** – Forme du potentiel  $V$  à deux dimensions dans le cas  $\mu^2 > 0$  (a) et dans le cas  $\mu^2 < 0$  (b)

$\lambda$  est choisi positif pour que le potentiel soit borné inférieurement, et avec  $\mu^2$  négatif le potentiel est toujours symétrique mais la valeur minimale du potentiel est non nulle (autrement dit, la valeur du champ n'est pas nulle dans son état fondamental) ; la figure 1.2 montre la forme du potentiel selon le signe de  $\mu^2$ . Il y a une infinité de minimum correspondant à  $|\phi|^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$  où  $v$  est la valeur du potentiel dans le vide (ou *vev* pour *vacuum expectation value*). Le choix d'un état fondamental particulier, par exemple  $\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$  mène à la brisure spontanée de symétrie et à l'apparition dans le lagrangien des termes de masse pour les bosons  $W$  et  $Z$ . Des quatre degrés de liberté introduits avec le doublet de champs scalaires complexes, trois donnent donc leur masse au boson  $W^+$ ,  $W^-$ ,  $Z$ , et celui restant correspond à un autre boson massif, le boson de Higgs. La masse du boson de Higgs est donnée par :

$$m_H = \sqrt{2\lambda}v \quad (1.4)$$

Les masses obtenues pour les bosons  $W$  et  $Z$  sont :

$$m_{W^\pm} = \frac{1}{2}gv \quad m_Z = \frac{1}{2}v\sqrt{g'^2 + g^2} \quad (1.5)$$

La symétrie  $U(1)$  reste respectée, les photons conservent une masse nulle. Dans l'expression de la masse du boson de Higgs,  $v$  peut être calculé à partir de la constante de Fermi ( $\frac{G_F}{\sqrt{2}} = \frac{1}{2v^2}$ ) et vaut 246 GeV, mais le paramètre  $\lambda$  est indéterminé.

La masse des fermions  $m_f$  s'obtient ensuite en introduisant un couplage de Yukawa  $\lambda_Y$  entre les champs de quarks et le champs de Higgs :

$$m_f = \frac{\lambda_Y v}{\sqrt{2}} \quad (1.6)$$

### 1.2.2 Contraintes théoriques sur la masse

La masse du boson de Higgs n'est pas prédite par le Modèle Standard, il existe cependant des arguments théoriques qui la contraignent et permettent de prédire une valeur inférieure à l'échelle du TeV.

Le premier d'entre eux est l'unitarité. Si seuls les bosons  $Z$  et  $\gamma$  peuvent être échangés dans la diffusion du boson  $W$ , l'amplitude de ce processus est proportionnelle à l'énergie disponible dans le centre de masse  $s$  au carré<sup>2</sup> [13] :  $\mathcal{M} \simeq \sqrt{2}G_F(s+t)$ . L'unitarité, qui

2. On considère le processus  $ab \rightarrow cd$ ,  $s = (p_a + p_b)^2$  et  $t = (p_a - p_c)^2$ , avec  $p_i$  le quadri-vecteur de la particule  $i$ .

traduit simplement la conservation des probabilités, n'est donc pas respectée à haute énergie. Cette amplitude peut être régularisée si on considère la contribution du boson de Higgs :  $\mathcal{M} \simeq -\sqrt{2}G_F m_h^2 (s/(s - m_h^2) + t/(t - m_h^2))$ . Mais le respect de l'unitarité quelle que soit l'énergie implique une limite supérieure sur la masse du boson de Higgs,  $m_H \lesssim 700$  GeV.

Le second argument avancé est celui de la trivialité. Dans l'expression du potentiel de Higgs, le terme de couplage  $\lambda$  dépend de l'échelle d'énergie  $\mu$ . Cette dépendance peut être calculée en résolvant les équations du groupe de renormalisation [14]. Pour que la théorie soit valide (perturbative) quel que soit  $\mu$ , cela implique que pour l'échelle d'énergie de référence  $\mu_0$ , l'on ait :  $\lambda(\mu_0) \rightarrow 0$ . La théorie devient triviale, sans interaction (théorie des champs libres). Il est donc nécessaire d'introduire une limite d'énergie  $\Lambda$  jusqu'à laquelle la théorie est valide. On peut alors poser une limite sur  $m_H$  en fonction de  $\Lambda$  :

$$m_H < v^2 \frac{16\pi^2}{3 \ln(\Lambda/\mu_0)} \quad (1.7)$$

Si  $\Lambda$  est de l'ordre de la masse de Planck ( $10^{19}$ ), le boson de Higgs doit être tel que  $m_H \lesssim 145$  GeV. Si par contre l'échelle d'énergie est faible  $\sim 10^3$  GeV, la masse du boson de Higgs est limitée par  $m_H \lesssim 750$  GeV.

Enfin la stabilité du vide apporte la dernière contrainte théorique sur la masse du boson de Higgs. La condition  $\lambda > 0$  est nécessaire pour préserver la stabilité du vide [15]. Or si on considère la contribution des fermions dans l'expression de  $\lambda$  celui-ci peut devenir négatif en raison de la masse du top. Ceci impliquerait un potentiel de Higgs non borné et instable. Il est donc nécessaire d'imposer une condition sur le couplage  $\lambda$  qui se traduit par une limite inférieure sur la masse du Higgs :

$$m_H > 70 \text{ GeV pour } \Lambda = 10^3 \text{ GeV} \quad (1.8)$$

$$m_H > 130 \text{ GeV pour } \Lambda = 10^{19} \text{ GeV} \quad (1.9)$$

### 1.2.3 Contraintes expérimentales sur la masse

La recherche du boson de Higgs a débuté il y a plus de 35 ans ; aux contraintes théoriques, se sont ajoutées, au fil des expériences, des contraintes expérimentales directes et indirectes. Nous faisons ici un bref récapitulatif de l'état des recherches en juin 2012 avant l'annonce du CERN.

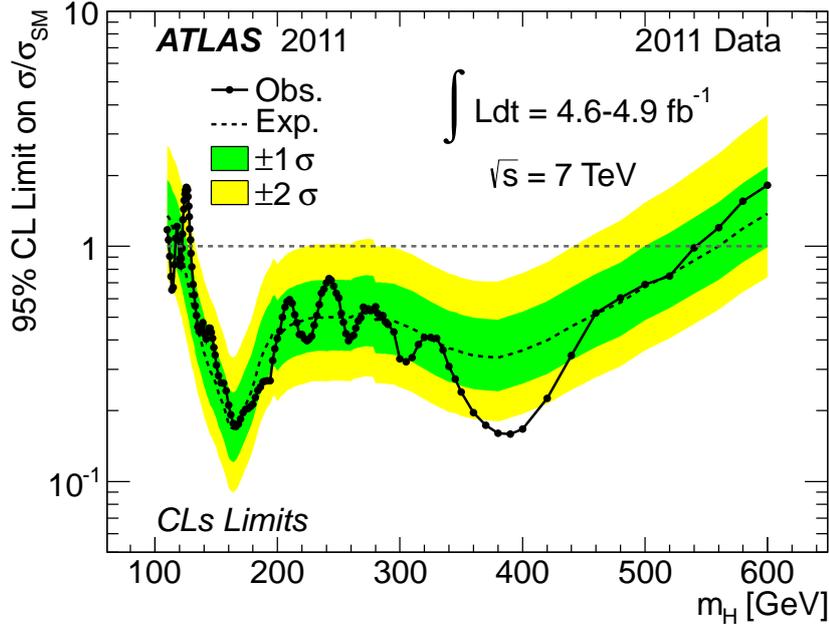
#### 1.2.3.1 Contraintes directes

La première recherche directe dans une large zone de masse a été faite au LEP (Large Electron Positron collider) avec une énergie au centre de masse comprise entre 90 et 209 GeV. Le principal mode de production était le processus de Bjorken  $e^+e^- \rightarrow Z^*H$ , avec le boson de Higgs se désintégrant ensuite en  $b\bar{b}$ . Les mesures combinées des quatre expériences au LEP (ALEPH, DELPHI, L3 et OPAL) ont exclu à 95% de niveau de confiance un boson de Higgs de masse inférieure à 114.4 GeV [16].

Les recherches ont ensuite continué au Tevatron avec les expériences D0 et CDF et des collisions protons-antiprotons ( $10 \text{ fb}^{-1}$  de luminosité intégrée). Les résultats de juin 2012 ont exclu à 95% de niveau de confiance le boson de Higgs dans les intervalles de masse 100-106 GeV et 147-179 GeV [17]. L'exclusion n'a pas pu être établi entre 115 et 140 GeV, un excès supérieur à  $2\sigma$  a été observé pour  $m_H \approx 120$  GeV.

Au LHC, en décembre 2011 avec  $5 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$ , l'expérience ATLAS a restreint les intervalles autorisés pour le boson de Higgs à 117.5-118.5 GeV, 122.5-129 GeV et au dessus de 539 GeV [18](figure 1.3). CMS, quant à elle, autorisait les intervalles en dessous de 127.5 GeV et au dessus de 600 GeV [19]. Les régions exclues le sont à 95% de niveau de confiance. ATLAS et CMS ont toutes deux reporté un excès de respectivement  $2.8\sigma$  à 126.5 GeV et  $3.1\sigma$  à 124 GeV en décembre 2011.

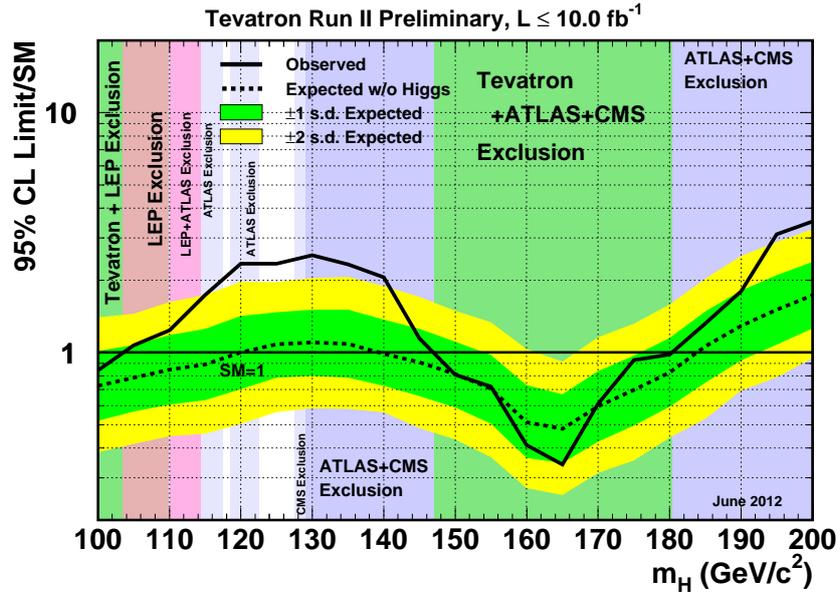
La figure 1.4 montre les résultats combinés pour les expériences CDF et D0 et résume l'ensemble des zones exclues par les expériences du LEP, du Tevatron et du LHC jusqu'en juin 2012.



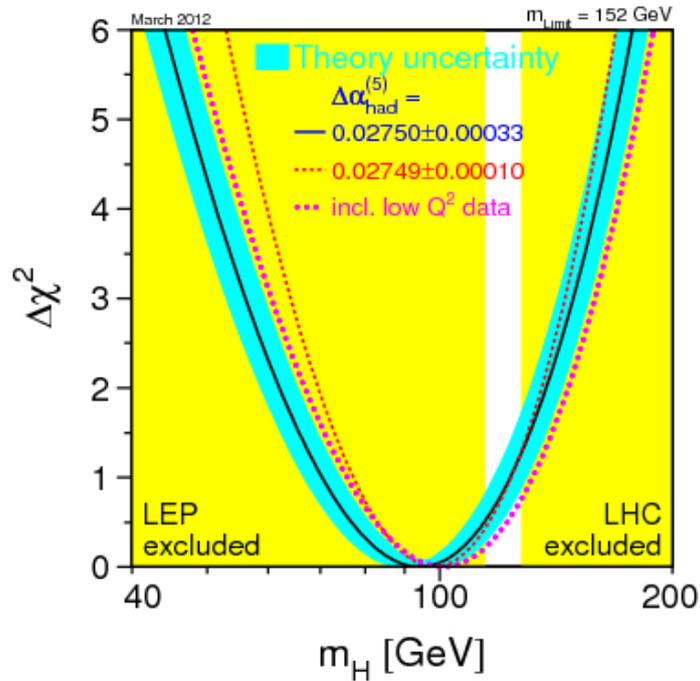
**Figure 1.3** – Limites sur la section efficace de production du boson de Higgs Modèle Standard obtenues par l'expérience ATLAS. La ligne pleine correspond à la limite observée et la ligne pointillée à la limite attendue (en unité du Modèle Standard). Les zones exclues (les zones pour lesquelles la limite observée est inférieure à 1) le sont à 95% de niveau de confiance [18].

### 1.2.3.2 Contraintes indirectes

La masse du boson de Higgs intervient dans le calcul des corrections radiatives de plusieurs observables électrofaibles comme la masse des bosons  $W$  et  $Z$ . Les mesures de précision de ces observables permettent de contraindre la masse du boson de Higgs. Cela a été fait par le *LEP Electroweak Working Group* qui combine les résultats de mesures au LEP, Tevatron et SLAC. Le  $\chi^2$  du fit sur les données en fonction de la masse du boson de Higgs est présenté figure 1.5. La meilleure valeur du fit trouvée permet de contraindre  $m_H$  à 68% de niveau de confiance telle que  $m_H = 94^{+29}_{-24}$  GeV et à 95% telle que  $m_H < 152$  GeV. Des résultats similaires sont obtenus par la collaboration *Gfitter* [20].



**Figure 1.4** – Limites sur la section efficace de production du boson de Higgs Modèle Standard obtenues par les expériences du Tevatron. La ligne pleine correspond à la limite observée et la ligne pointillée à la limite attendue (en unité du Modèle Standard). Les zones exclues (les zones pour lesquelles la limite observée est inférieure à 1) le sont à 95% de niveau de confiance [17].



**Figure 1.5** –  $\Delta\chi^2 = \chi^2 - \chi_{\text{min}}$  de l'ajustement global de la masse du boson de Higgs sur les mesures de précision électrofaibles, obtenu par le groupe LEP-EWWG. La partie jaune à gauche est la partie exclue par les expériences au LEP, la partie jaune à droite est la zone exclue par les expériences du LHC [21].

### 1.2.4 Les modes de désintégration

Les couplages du boson de Higgs sont proportionnels à la masse des particules.

Pour les fermions :

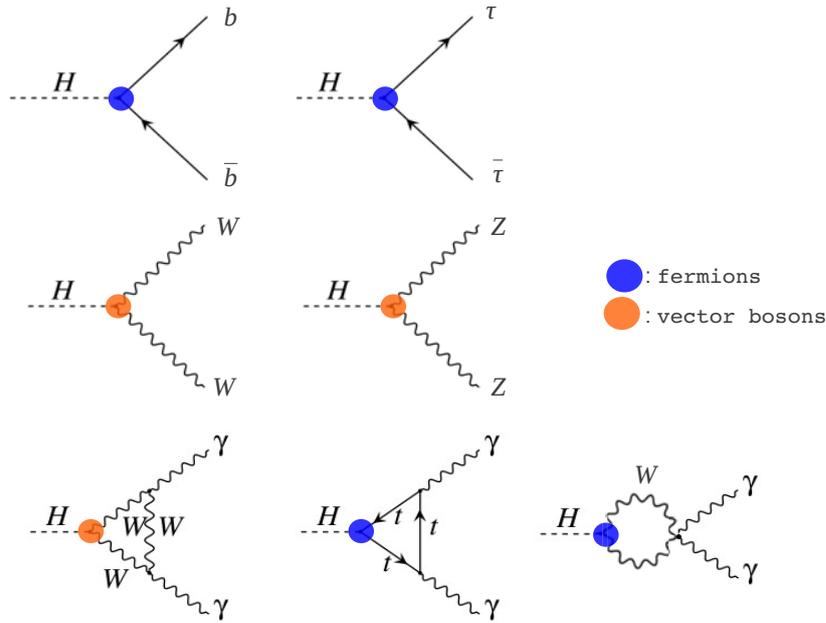
$$g_{Hff} = i \frac{m_f}{v} \quad (1.10)$$

Pour les bosons  $Z$  et  $W$  :

$$g_{HWW} = 2i \frac{m_W^2}{v} \quad (1.11)$$

$$g_{HZZ} = 2i \frac{m_Z^2}{v} \quad (1.12)$$

Ceci implique que le boson de Higgs se désintègre préférentiellement en la paire de particules la plus lourde permise et les photons et gluons, dépourvus de masse, ne peuvent se coupler au boson de Higgs que par l'intermédiaire de boucles faisant intervenir des particules lourdes. Par ailleurs, les masses des fermions et bosons vecteurs étant connues, une fois la masse du boson de Higgs établie, il devient possible de calculer les largeurs de désintégrations partielles du boson de Higgs.



**Figure 1.6** – Diagrammes de Feynman pour les principaux modes de désintégration du boson de Higgs.

Les principaux modes de désintégration du boson de Higgs sont présentés figure 1.6 ; la désintégration du boson de Higgs en deux photons se fait par l'intermédiaire du boson  $W$  et du quark top dont le signe de la contribution est opposé et la contribution du boson  $W$  domine largement [14]. Les rapports d'embranchement du boson de Higgs sont résumés figure 1.7. Jusqu'à  $\sim 140$  GeV, le principal mode de désintégration est  $H \rightarrow b\bar{b}$ , ensuite le mode  $H \rightarrow WW$  devient dominant.

La région de masse 122.5 - 127.5 GeV nous intéresse tout particulièrement en raison des contraintes expérimentales sur la masse du Higgs détaillées section 1.2.2, nous concentrerons donc sur cette région par la suite.

Dans cette région, la désintégration en  $b\bar{b}$  est dominante ( $\sim 80\%$ ) mais l'importance du bruit

de fond QCD au LHC rend ce canal difficilement exploitable expérimentalement. Le canal  $H \rightarrow \tau^+\tau^-$  présente un rapport d'embranchement plus faible que  $H \rightarrow b\bar{b}$  ( $\sim 8\%$ ) mais sa signature le rend plus discriminant pour le bruit de fond QCD. Le rapport d'embranchement pour la désintégration  $H \rightarrow WW^{(*)}$  est de l'ordre de 4%. Le canal  $H \rightarrow WW^{(*)} \rightarrow \nu\nu$  est particulièrement intéressant car les spécificités des produits de la désintégration des bosons  $W$  permettent d'isoler le signal du bruit de fond. La désintégration  $H \rightarrow ZZ^{(*)}$  est rare ( $\sim 0.4\%$ ) et le rapport d'embranchement  $Z \rightarrow l^+l^-$  est faible ( $\sim 3\%$ ) mais le canal  $H \rightarrow ZZ^{(*)} \rightarrow llll$  est très propre avec un faible bruit de fond.

Malgré son faible rapport d'embranchement, de l'ordre de 0.2%,  $H \rightarrow \gamma\gamma$  est l'un des plus intéressants au LHC. Sa signature consiste en deux photons isolés de grande impulsion transverse dont les paramètres cinématiques peuvent être mesurés précisément avec le calorimètre électromagnétique. Ceci permet une reconstruction précise de la masse invariante  $m_{\gamma\gamma}$ , et confère à ce canal une grande sensibilité expérimentale. Le bruit de fond irréductible dans ce canal correspond à la production non résonante de paires de photons (détaillée section 1.3.1) qui est très largement supérieure au signal. Le signal dans le canal  $H \rightarrow \gamma\gamma$  consistera en un pic étroit (la largeur de boson de Higgs est de  $\sim 4$  MeV à  $\sim 125$  GeV) au dessus d'un bruit de fond continu.

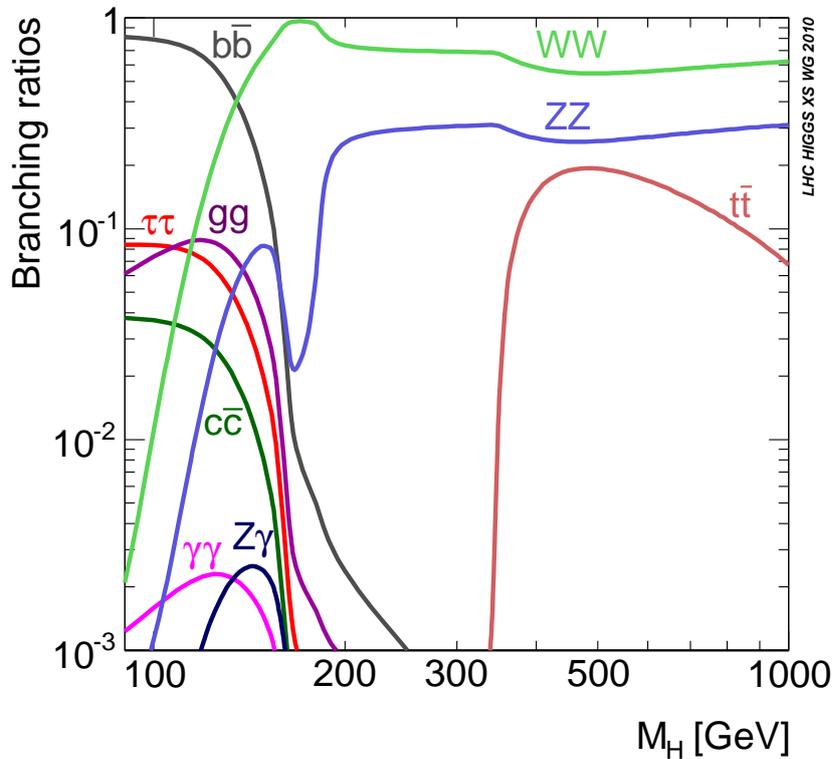
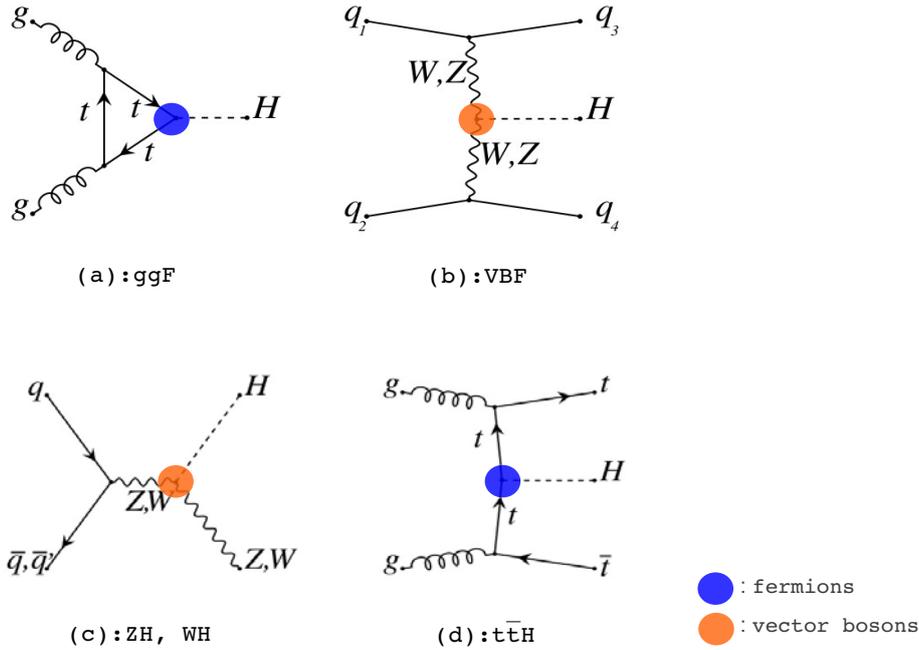


Figure 1.7 – Rapports d'embranchement du boson de Higgs en fonction de sa masse [22].

### 1.2.5 Les modes de production

Les quatre modes de production dominants au LHC sont : la fusion de gluons  $ggF$  ( $\sim 87\%$  pour  $m_H \sim 125$  GeV), la fusion de bosons vecteurs  $VBF$  ( $\sim 7\%$ ), la production associée aux bosons  $W$  et  $Z$   $VH$  ( $\sim 5\%$ ) et la production associée à des quarks top  $ttH$  ( $\sim 1\%$ ) [23]. Les diagrammes de Feynman correspondants sont présentés figure 1.8.

La fusion de gluons est le processus principal de production du boson de Higgs au LHC. Le couplage du boson de Higgs avec les gluons se fait par l'intermédiaire d'une boucle de quarks



**Figure 1.8** – Diagrammes de Feynman pour la production du boson de Higgs : (a) fusion de gluons, (b) fusion de bosons vecteurs, (c) production associée au boson  $W/Z$ , (d) production associée à des quarks top

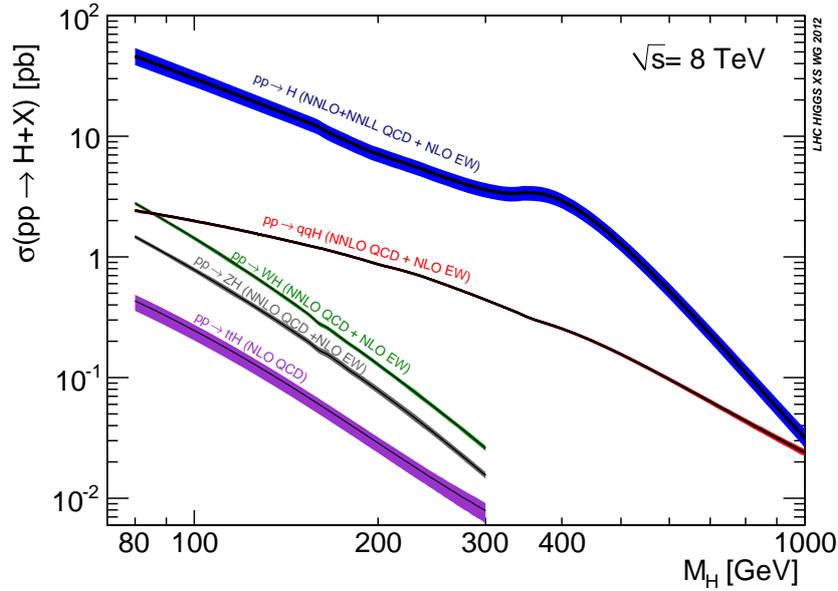
top. Dans la mesure où ce mode de production fait intervenir l'interaction forte, les corrections de QCD radiative sont importantes. La principale contribution vient de l'émission d'un gluon dans l'état final, le boson de Higgs produit par fusion de gluons est donc souvent accompagné d'un jet (plus de détails section 1.4.1).

Dans le processus VBF, le boson de Higgs est produit à partir de bosons  $W$  ou  $Z$ , mais la contribution principale vient de la fusion de  $W$  car le couplage est plus fort. Le boson de Higgs est produit avec deux jets de grande énergie avec une large séparation en pseudo-rapacité (dans les régions avant du détecteur). Les bosons vecteurs ne portant qu'une faible fraction de l'énergie des partons initiaux ( $\simeq m_h/2$  chacun), les deux quarks dans l'état final ont une grande énergie et un faible  $p_T$  (petit angle de diffusion). Les produits de la désintégration du Higgs, quant-à-eux se situent dans la région plus centrale du détecteur. Ce processus possède donc des caractéristiques clairement identifiables et présente ainsi un grand pouvoir discriminant par rapport aux processus QCD qui constituent notre bruit de fond.

La production d'un boson de Higgs en association avec des bosons vecteurs peut être traitée comme un processus Drell-Yann. Un boson virtuel est produit par processus Drell-Yann, ce boson virtuel donne ensuite un boson vecteur réel et un boson de Higgs.

Enfin, l'état final de la production associée à des quarks top est composé de trois particules massives, c'est le mode de production le plus compliqué. Contrairement au processus de fusion de gluon, la production associée à des quarks top ne fait pas intervenir de boucle, elle nous renseigne donc directement sur le couplage de Yukawa entre le boson de Higgs et le quark top.

Les sections efficaces de production du boson de Higgs au LHC pour  $\sqrt{s} = 8$  TeV sont présentées figure 1.9 (entre 7 et 8 TeV et pour une masse du Higgs  $\sim 125$  GeV, la section efficace de production a été multipliée par  $\sim 1.3$ ).



**Figure 1.9** – Section efficace de production du boson de Higgs par canal au LHC à  $\sqrt{s} = 8$  TeV. Les bandes de couleurs représentent l'incertitude théorique [23].

### 1.2.6 Détermination des propriétés

La découverte d'une nouvelle particule au LHC dans les canaux de recherche cité auparavant n'est toutefois pas suffisante pour établir qu'il s'agit bien de la particule prédite par le Modèle Standard, il est également nécessaire d'étudier ses propriétés comme le spin et la masse, ainsi que ses couplages.

Dans le canal  $H \rightarrow \gamma\gamma$ , le signal apparaît comme un pic au dessus du bruit de fond que l'on peut soustraire. La mesure de la masse dans ce canal est donc relativement aisée mais dépend de la connaissance de l'échelle d'énergie des photons. La largeur de désintégration totale, elle, n'est pas accessible, elle est prédite à  $\sim 4$  MeV (pour  $m_H=125$  GeV), quand la résolution des détecteurs est de l'ordre du GeV. Par ailleurs, selon le théorème de Landau et Yang [24, 25], la désintégration d'une particule de spin 1 en deux particules de spin 1 identiques est interdite. L'observation dans le canal  $H \rightarrow \gamma\gamma$  permet donc de rejeter l'hypothèse de spin 1.

La mesure du taux de production du boson de Higgs dans un canal fournit une mesure de la section efficace de production fois le rapport d'embranchement. Les mesures des taux de production nous renseignent surtout sur le caractère "standard" ou non de la particule. Les couplages du boson de Higgs aux fermions et bosons vecteurs déterminent les sections efficaces et les rapports d'embranchement ; en inversant la démarche et en faisant un certain nombre d'hypothèses théoriques il est donc possible de remonter aux couplages (plus de détails section 7.7). Toutefois, la largeur totale de désintégration n'étant pas accessible, il n'est possible de mesurer que des rapports de constantes de couplages. Dans ces mesures, le canal  $H \rightarrow \gamma\gamma$  est particulièrement intéressant car il est sensible au signe relatif des couplages du boson de Higgs au top et au boson  $W$  en raison de la contribution opposée de leur boucle dans la désintégration du boson de Higgs en deux photons.

## 1.3 Evénements diphoton directs au LHC

Le canal  $H \rightarrow \gamma\gamma$  se manifeste par un pic de masse invariante étroit au dessus d'un bruit de fond continu. Ce bruit de fond est constitué de deux composantes :

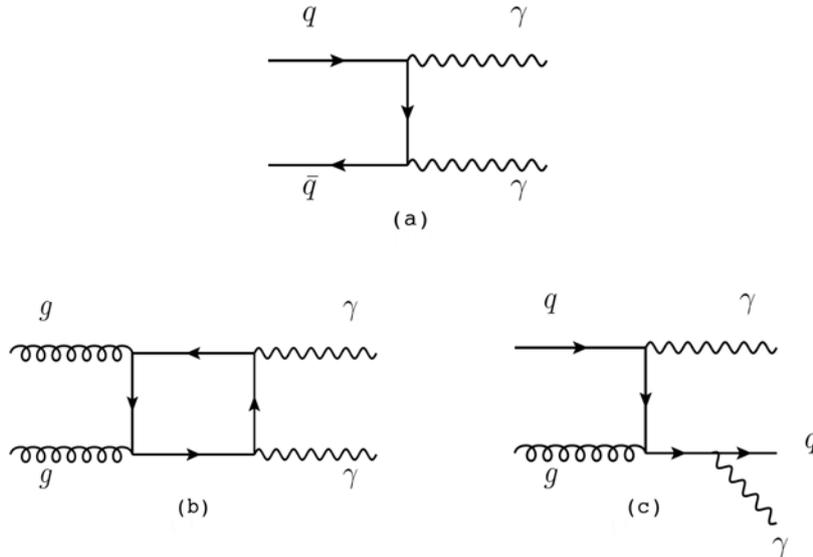
- Un bruit de fond irréductible : les événements à deux photons directs.
- Un bruit de fond réductible : les événements impliquant un photon et un jet, ou plusieurs jets, du fait de la mauvaise identification des jets en photons.

Oltre la compréhension du bruit de fond pour l'analyse  $H \rightarrow \gamma\gamma$ , l'étude des paires de photons directs constitue un excellent test pour les prédictions de QCD perturbative dans des régions cinématiques inexplorées ou mal comprises jusqu'alors et permet ainsi de raffiner les générateurs Monte-Carlo. Elle nous amène à aborder le problème de la fragmentation des partons et de la resommation des gluons mous. Enfin, elle nous permet de vérifier notre compréhension du traitement expérimental des photons (isolation et identification).

### 1.3.1 Les processus de production

Il existe deux modes principaux de production de paires de photons directs au LHC, les photons peuvent résulter d'un processus dur (production directe) ou de la fragmentation d'un parton. Dans le cas de la production directe à l'ordre dominant, on distingue trois processus principaux dont les diagrammes de Feynman sont présentés figure 1.10 :

- Le processus de Born : l'annihilation  $q\bar{q} \rightarrow \gamma\gamma$ , d'ordre  $\alpha_{QED}^2$
- Le processus de boîte (box) :  $gg \rightarrow \gamma\gamma$ , d'ordre  $\alpha_S^2\alpha_{QED}^2$ . Même si ce processus est d'ordre supérieur, la très forte densité de gluons au LHC rend sa section efficace du même ordre de grandeur que celle du processus de Born [26].
- Le processus de rayonnement (bremsstrahlung) : un photon est rayonné par un quark  $qg \rightarrow q\gamma\gamma$ , d'ordre  $\alpha_S\alpha_{QED}^2$ . Ce processus est également d'ordre supérieur mais compensé par l'effet de densité des gluons.

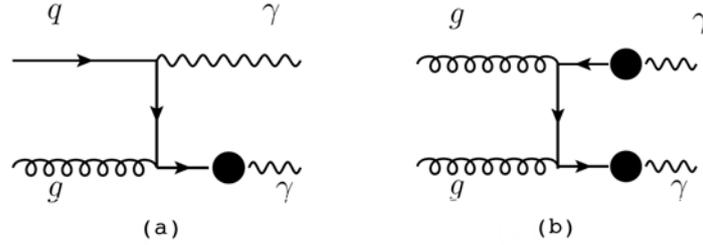


**Figure 1.10** – Diagrammes de Feynman à l'ordre dominant des processus de Born (a), boîte (b) et bremstrahlung (c)

Les corrections d'ordre supérieur correspondent à des émissions de partons et à des corrections virtuelles, lorsque le parton n'est pas émis, mais réabsorbé à l'aide d'une boucle. Dans le

cas du processus de Born elles augmentent la section efficace d'environ 30%.

Dans le cas du bremsstrahlung, quand l'angle de séparation entre le parton et le photon émis tend vers 0, une singularité dans le calcul du processus  $qg \rightarrow q\gamma\gamma$  apparaît. Les divergences sont factorisées à tous les ordres  $\alpha_s$  et absorbées dans une fonction de fragmentation (voir section 1.3.3). L'échelle arbitraire de fragmentation détermine la distinction entre les deux processus (bremsstrahlung ou fragmentation) et n'est pas un paramètre physique. Les deux contributions, directe et fragmentation doivent être combinées pour avoir une description complète de l'émission d'un photon par un parton ; seule la somme est un observable physique [26]. Les processus de fragmentation sont présentés figure 1.11.



**Figure 1.11** – Diagrammes de Feynman à l'ordre dominant pour la simple (a), et double (b) fragmentation

### 1.3.2 QCD perturbative

La compréhension de la production de paires de photons repose sur la description des interactions entre partons. Les gluons, en raison de leur charge de couleur, peuvent se coupler entre eux. C'est ce qui confère à l'interaction forte un caractère non linéaire. Le potentiel de l'interaction forte entre deux quarks peut s'écrire [27] :

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} + \lambda r \quad (1.13)$$

Avec  $\lambda$  une constante de l'ordre de  $1 \text{ GeV fm}^{-1}$ , elle représente la densité d'énergie du lien formé entre deux quarks qui s'éloignent, et  $\alpha_s$  est la constante de couplage de l'interaction forte, elle dépend de l'échelle d'énergie  $Q$  du processus. La figure 1.12 montre la variation de  $\alpha_s$  en fonction de  $Q$  ;  $\alpha_s$  est une fonction décroissante (croissante) de l'énergie (la distance). On distingue deux régimes. A haute énergie (ou petite distance,  $r < 0.1 \text{ fm}$ ),  $\alpha_s \ll 1$  amène la propriété de liberté asymptotique, les quarks sont considérés libres ce qui autorise les calculs perturbatifs. A l'inverse, à basse énergie (ou longue distance  $r > 0.1 \text{ fm}$ ),  $\alpha_s \sim 1$  explique le confinement des quarks dans les hadrons et rend invalide l'approche perturbative.

Dans le régime perturbatif, à haute énergie,  $\alpha_s$  peut s'exprimer en fonction de l'impulsion échangée  $\alpha_s(Q^2) \propto \frac{1}{\ln(Q^2/\Lambda_{QCD}^2)}$ , avec  $\Lambda_{QCD} \sim 200 \text{ GeV}$ , le paramètre d'échelle de QCD. La section efficace partonique admet alors un développement en puissance de  $\alpha_s$ . Si l'ordre de grandeur obtenu au premier ordre (LO pour *Leading Order*) est correct, les calculs aux ordres supérieurs (NLO pour *Next to Leading Order* et éventuellement NNLO) sont nécessaires pour avoir une estimation précise. Les divergences apparaissant dans les calculs à ordre fixe (diagrammes de Feynman en boucles) sont absorbées en renormalisant la théorie [28]. Ceci implique l'introduction d'une échelle arbitraire de renormalisation  $\mu$ .

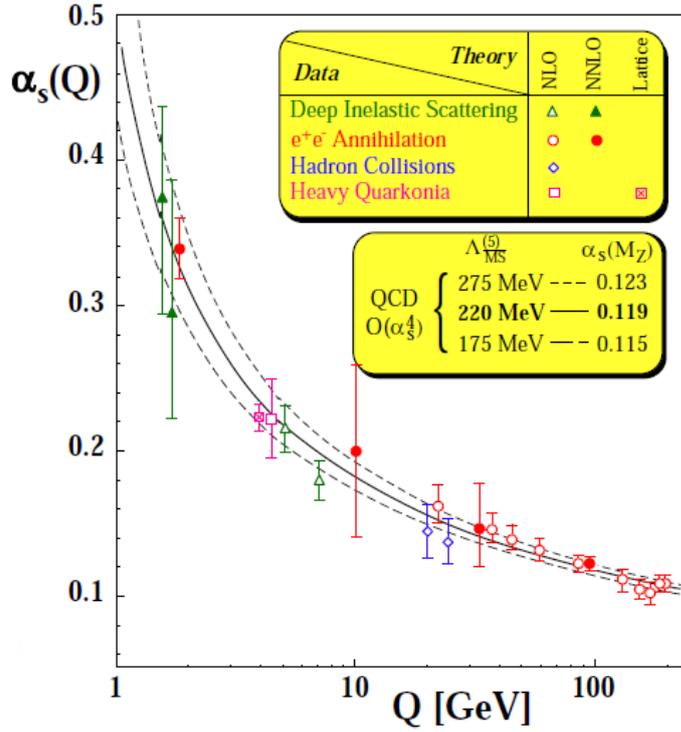


Figure 1.12 – Variation de la constante de couplage forte en fonction de l'échelle d'énergie [29].

### 1.3.3 Théorème de factorisation et fonctions de distribution des partons

Nous avons vu qu'à haute énergie, les interactions entre partons peuvent être correctement approximées grâce à une approche perturbative. Toutefois, cette approche n'est applicable qu'à haute énergie et il est également nécessaire de décrire les interactions entre partons à basse énergie.

Le théorème de factorisation [30] permet de décomposer le calcul de la section efficace  $\sigma_{pp \rightarrow \gamma\gamma+X}$  en un produit de convolution :

$$\begin{aligned} \sigma_{pp \rightarrow \gamma\gamma+X} &= PDFs \otimes \sigma_{partonique} \otimes FFs \\ \sigma_{pp \rightarrow \gamma\gamma+X} &= [f_p(x_1, Q^2) \otimes f_p(x_2, Q^2)] \otimes \sigma(x_1, x_2, Q^2) \otimes D(z, Q^2) \end{aligned} \quad (1.14)$$

Le premier terme correspond aux PDFs (pour *Parton Distribution Fonction*) ; ce sont les densités de probabilité de trouver un parton dans un proton avec une certaine fraction  $x$  de quantité de mouvement pour une impulsion transférée  $Q^2$ . Le deuxième terme représente la section partonique et comprend uniquement de la physique perturbative. Enfin le dernier terme correspond aux fonctions de fragmentation, elles donnent la probabilité qu'un parton s'hadronise en émettant un photon portant une fraction  $z$  d'impulsion. La décomposition permise par le théorème de factorisation implique le choix d'une échelle de factorisation  $M$ , qui sépare le régime perturbatif (section partonique) du régime non-perturbatif (PDFs), mais aussi le choix d'une échelle de fragmentation  $\mu_f$  qui fixe la séparation entre processus durs et processus de fragmentation.

Les PDFs ne peuvent pas être calculées [28] et sont extraites des données expérimentales collectées principalement à HERA et au Tevatron. Toutefois, la QCD perturbative prédit l'évolution des PDFs en fonction de l'échelle d'énergie (équations DGLAP [31]), deux principaux groupes CTEQ [32] et MRST [33] fournissent des paramétrisations des PDFs.

De même que pour les PDFs, les fonctions de fragmentation ne peuvent pas être déterminées de façon perturbative, elles sont dérivées des données expérimentales (la désintégration hadronique du  $Z$  au LEP par la collaboration ALEPH [34], [35]). Leur évolution en fonction de  $\mu_f$  est calculée perturbativement.

### 1.3.4 Bruits de fond réductible

Le bruit de fond provient de la mauvaise identification des jets comme des photons directs. Tous les événements impliquant la production d'un photon et un jet, ou de plusieurs jets constituent donc un potentiel bruit de fond. Dans le cas du bruit de fond  $\gamma$ -jet, les deux processus principaux à l'ordre  $\alpha_{QED}\alpha_s$  sont présentés figure 1.13. Le processus Compton  $qg$  représente  $\sim 95\%$  des contributions, contre  $\sim 5\%$  pour le processus  $q\bar{q}$ .

Pour le bruit de fond di-jets, la figure 1.14 montre des exemples à l'ordre dominant. Comme pour les événements  $\gamma$ -jet, des corrections d'ordre supérieur en  $\alpha_S$  viennent d'émissions de gluons et de quarks dans l'état final ou dans l'état initial.

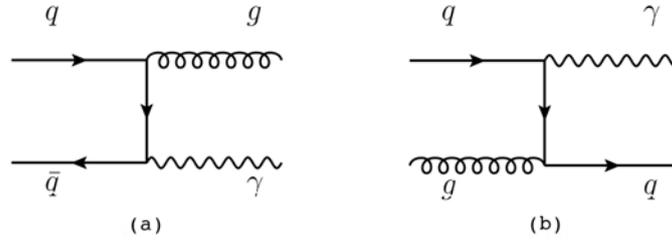


Figure 1.13 – Diagrammes de Feynman pour des événements  $\gamma$ -jet

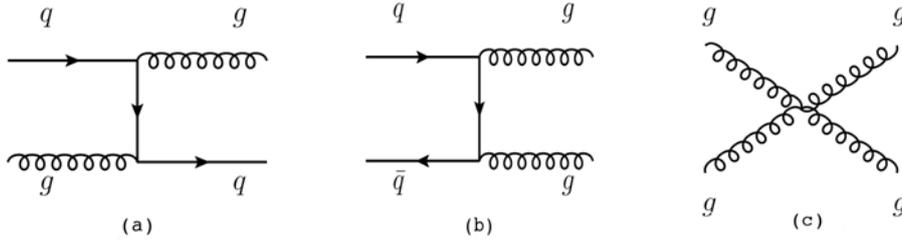


Figure 1.14 – Diagrammes de Feynman pour des événements di-jets

Ces événements de bruit de fond sont particulièrement difficiles à distinguer lorsqu'un quark ou un gluon donne dans l'état final un  $\pi^0$  très énergétique qui se désintègre en deux photons très proches. Le détecteur électromagnétique est optimisé pour identifier ce type d'événements et un certain nombre de coupures dites d'identification permettent de rejeter ce bruit de fond.

### 1.3.5 Définition des photons isolés

Les sections efficaces de production des événements  $\gamma$ -jet et jet-jet sont respectivement de l'ordre de  $10^3$  et  $10^7$  fois supérieures à la section efficace de production  $\gamma\gamma$ . Il est donc nécessaire d'avoir une excellente réjection des jets pour ne pas être dominé par le bruit de fond et les coupures d'identification expérimentales mentionnées auparavant ne sont pas suffisantes. On utilise donc également le fait que, à l'inverse des photons, les jets sont entourés d'une intense activité hadronique. On définit l'énergie d'isolation transverse  $E_T^{iso}$  comme la

somme de l'énergie transverse déposée autour du candidat photon dans un cône de rayon fini  $R$ . En coupant sur  $E_T^{iso}$  on peut donc faire la distinction entre les jets identifiés comme des photons et les photons directs produits par le processus de Born ou de boîte. Dans le cas des photons issus de la fragmentation d'un parton la situation est un peu plus complexe. Lorsque l'angle entre le parton et le photon émis est petit,  $E_T^{iso}$  augmente. En conséquence, la coupure sur  $E_T^{iso}$  soustrait au signal une partie de la composante de fragmentation. Entre 2 à 5% de la fragmentation simple subsiste et la fragmentation double est presque totalement supprimée pour les critères d'isolation habituellement utilisés [26]. Si l'on considère la situation à NLO, l'émission de gluons mous vient s'ajouter à la fragmentation et contribue à l'augmentation de  $E_T^{iso}$ . La coupure sur cette variable limitera donc l'espace des phases disponible pour l'émission de gluons mous.

Expérimentalement cette coupure sur  $E_T^{iso}$ , combinée avec les coupures d'identification, est nécessaire pour réduire significativement le bruit de fond, mais une partie du signal est également supprimée par cette coupure. La comparaison entre sections efficaces mesurées et prédites n'a donc de sens que si la coupure sur l'énergie d'isolation est appliquée dans les deux cas. Théoriquement cette coupure peut être implémentée dans les calculs ([36] et [37]), une énergie hadronique minimum autour du photon est cependant nécessaire pour supprimer toutes les divergences infrarouges et justifie le choix des coupures expérimentales détaillées section 4.3. Deux principaux critères d'isolation sont utilisés dans les prédictions théoriques : l'isolation avec un cône, équivalente à l'isolation expérimentale, et l'isolation Frixione décrite section 5.5.2.1 qui supprime complètement les composantes de fragmentation.

### 1.3.6 Précédentes mesures

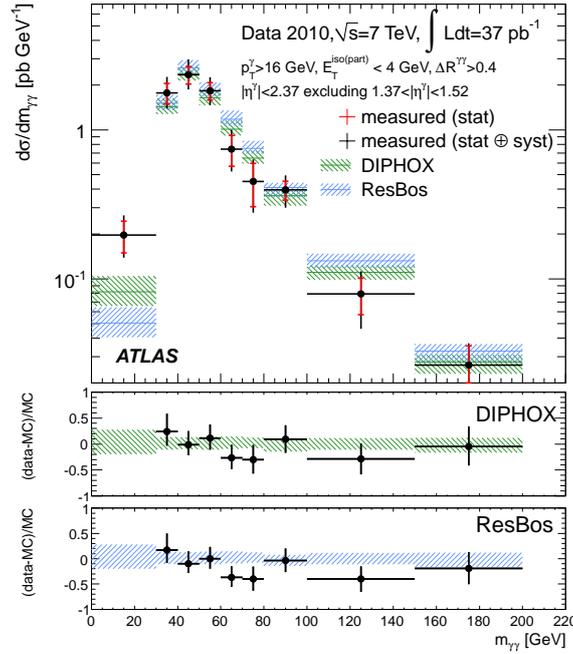
Des mesures de sections efficaces de production de paires de photons dans des collisionneurs hadroniques ont déjà été effectuées par les collaborations DØ [38] et CDF [39] pour  $\sqrt{s} = 1.96$  TeV au Tevatron et par les collaborations ATLAS [40] et CMS [41] pour  $\sqrt{s} = 7$  TeV au LHC. Les figures 1.15 et 1.16 présentent les mesures de sections efficaces différentielles en fonction de la masse invariante et de l'angle azimutal entre les deux photons  $\Delta\phi_{\gamma\gamma}$  effectuées par l'expérience ATLAS en 2010, ainsi que la comparaison avec les prédictions à NLO des générateurs DIPHOX [26] et ResBos [42].

Dans les deux cas, des désaccords entre les données et les prédictions à NLO sont observés à basse masse invariante et à bas  $\Delta\phi$ . Ces différences peuvent résulter d'une sous-estimation des corrections NNLO. Une amélioration de la précision de cette mesure, dominée par l'erreur statistique, et une comparaison avec des calculs à NNLO permettront de comprendre ce désaccord. Il est également nécessaire d'améliorer la description des données dans les régions sensibles à l'émission des gluons mous, comme la région à haut  $\Delta\phi$ , où les prédictions théoriques à ordre fixe sont surestimées.

## 1.4 Prédictions théoriques

Schématiquement on peut séparer les prédictions théoriques en deux types, les générateurs Monte-Carlo avec gerbe partonique ou *parton shower* [43, 44, 45] et les générateurs spécialisés avec des prédictions à ordre fixe.

Dans le cas des générateurs Monte-Carlo avec parton shower (comme SHERPA, PYTHIA...), les éléments de matrice sont généralement calculés à LO. Les contributions d'ordre supérieur et les radiations initiales (ISR) et finales (FSR) sont prises en charge par la parton shower. L'un des problèmes majeurs de ces générateurs est de combiner les calculs perturbatifs avec l'approche parton shower sans double comptage [46] [43] [47]. Après la parton shower, dès lors que la distance entre deux quarks augmente, le couplage des particules colorées augmente (équation 1.13). Le parton commence à émettre des gluons qui se décomposent ensuite en une paire quark anti-quark qui eux-mêmes émettent des gluons qui se décomposent en paire quark



**Figure 1.15** – Mesure de la section-efficace différentielle de production de paire de photons en fonction de la masse invariante  $m_{\gamma\gamma}$  par ATLAS [40].

anti-quark etc, créant ainsi une gerbes de partons. Ces partons vont ensuite s'assembler pour former des hadrons (état physique neutre de couleur), c'est le processus d'hadronisation. Ce processus est inclus dans les générateurs Monte-Carlo avec parton shower, et permet le passage du niveau partonique au niveau particule (appelée aussi "la vérité"). Les événements générés peuvent alors être injectés dans la simulation du détecteur.

Les prédictions à ordre fixe des générateurs spécialisés sont des prédictions au niveau partonique, elles reposent uniquement sur des calculs de QCD perturbative. Les éléments de matrice sont généralement calculés à un ordre plus grand (NLO ou NNLO) que pour les générateurs Monte-Carlo avec parton shower. Ces prédictions fournissent le plus souvent des distributions des variables cinématiques, et ne permettent pas de simuler des événements individuels, on les appelle également des intégrateurs. Il existe toutefois des exceptions comme POWHEG qui permet de générer des événements interfaçables ensuite avec un processus de parton shower.

### 1.4.1 Production du boson de Higgs

Pour la production de bosons de Higgs par fusion de gluons (ggF) et par fusion de boson vecteurs (VBF) les événements sont générés par le générateur POWHEG [48, 49], dans lequel les éléments de matrice sont calculés à NLO, et sont ensuite interfacés avec PYTHIA pour la parton shower et l'hadronisation. Pour la production du boson de Higgs en association avec des bosons vecteurs (VH) ou avec une paire de quark top (ttH), les événements sont générés par PYTHIA (éléments de matrice à LO et parton shower).

Les sections efficaces de production du boson de Higgs sont calculées à l'ordre supérieur (NLO ou NNLO) avec les générateurs à ordre fixe et un facteur de correction est ensuite appliqué aux prédictions de POWHEG et PYTHIA.

La section efficace du processus ggF est calculée à NNLO (corrections à trois boucles) [50, 51, 52, 53, 54, 55]. Les corrections à NLO font intervenir des gluons virtuels, l'émission de gluons dans l'état final et les diagrammes correspondants à une collision gluon quark et quark



figure 1.9, les bandes de couleurs entourant les lignes continues correspondent à l'incertitude théorique.

Les rapports d'embranchement sont calculés dans [77, 78, 79, 80] et leurs incertitudes sont estimées dans [81, 23]. La figure 1.17 représente les rapports d'embranchement avec leur incertitude en fonction de la masse du boson de Higgs dans l'intervalle 80-200 GeV, elle varie entre 2 à 6% pour  $m_H$  entre 110 et 150 GeV.

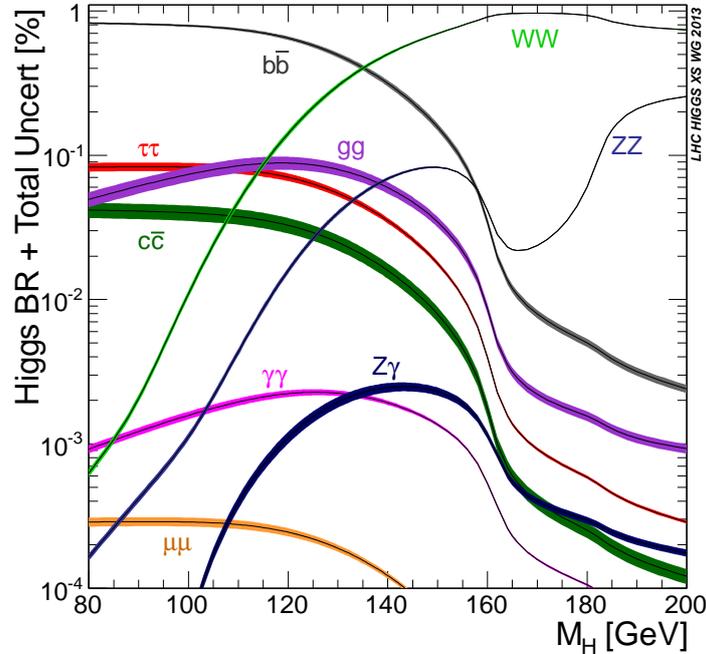


Figure 1.17 – Rapports d'embranchement du boson de Higgs en fonction de sa masse [81, 23].

### 1.4.2 Production de paires de photons

Pour la production d'événements diphoton, dans le cas des prédictions à ordre fixe, la composante de fragmentation est incluse dans les calculs via une fonction de fragmentation (section 1.3.3) tandis que pour les générateurs Monte-Carlo avec parton shower elle est prise en compte par la parton shower.

Les outils théoriques considérés dans la mesure de sections efficaces de paires de photons présentée au chapitre 5 sont résumés dans le tableau 1.1 pour les prédictions à ordre fixe et dans le tableau 1.2 pour les générateurs avec parton-shower.

Nom	Etat final	Diagramme / Ordre				Isolation	
		Born	Box	1 frag	2 frag	Cône	Frixione
DIPHOX	$\gamma\gamma$	NLO	LO	NLO	NLO	✓	✓
GAMMA2MC	$\gamma\gamma$	NLO	NLO			✓	✓
MCFM	$\gamma\gamma$	NLO	NLO	LO	LO	✓	✓
2 $\gamma$ NNLO	$\gamma\gamma$	NNLO	NLO				✓

**Tableau 1.1** – Résumé des prédictions à ordre fixe considérés dans la mesure de sections efficaces de paires de photons présentée au chapitre 5.

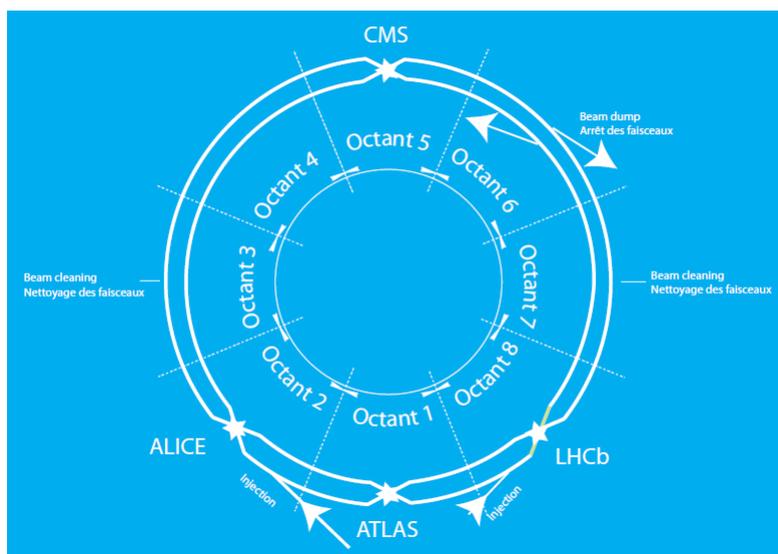
Nom	Etat final	Diagramme / Ordre				Isolation	
		Born	Box	1 frag	2 frag	Cône	Frixione
PYTHIA/ HERWIG	$\gamma\gamma$	LO	LO	PS	PS	✓	✓
SHERPA	$\gamma\gamma + 2$ jets	LO	LO	PS	PS	✓	✓
ALPGEN	$\gamma\gamma + 5$ jets	LO		PS	PS	✓	✓

**Tableau 1.2** – Résumé des générateurs Monte-Carlo avec parton-shower considérés dans la mesure de sections efficaces de paires de photons présentée au chapitre 5.

## Chapitre 2

# Le LHC et le détecteur ATLAS

Le *Large Hadron Collider* est un accélérateur de particules circulaire situé au CERN (Organisation Européenne pour la Recherche Nucléaire) à la frontière Franco-Suisse près de Genève. Il constitue à ce jour le plus grand outil jamais imaginé et construit pour la recherche scientifique fondamentale. Imaginé à partir de 1984, sa construction débute en 2000 après l'arrêt du LEP (*Large Electron-Positron collider*). Il a été conçu afin d'atteindre, en fonctionnement nominal, une énergie au centre de masse de 14 TeV, soit sept fois celle du Tevatron et une luminosité cent fois supérieure à celle du LEP ou du Tevatron.



Les protons sont plus faciles à accélérer que les électrons dans un accélérateur circulaire, de par leur plus grande masse ils perdent beaucoup moins d'énergie par rayonnement synchrotron. En outre, les collisions de protons mettent en jeu des particules composites dans l'état initial ; pour chaque collision l'énergie disponible pour le processus dur n'est pas connue car l'impulsion du proton est répartie entre les partons. L'énergie en jeu n'est donc pas constante, ce qui est un avantage pour l'une des principales recherches au LHC, le boson de Higgs dont la masse n'est pas déterminée théoriquement. Le choix d'un collisionneur hadronique semble donc tout à fait justifié mais les précédents collisionneurs hadroniques réalisaient des collisions proton-antiproton (découverte des bosons  $W$  et  $Z$  au *Super Proton Synchrotron*, découverte du quark top au Tevatron...), et ce sont les collisions proton-proton qui ont lieu au LHC. En réalité, le taux de production d'antiprotons est un facteur limitant ; la luminosité au LHC aurait été 100 fois plus faible avec des collisions proton-antiproton.

Le bilan du LHC après deux années de fonctionnement est excellent mais la mise en route n'a pas été chose évidente. Quelques jours après la première circulation de protons dans le

faisceau du LHC en septembre 2008, un incident majeur dû à un problème de connection dans le système cryogénique survient. Les réparations durent un an, le programme d'exploitation du LHC doit être modifié ; afin de protéger l'accélérateur, l'énergie est augmentée progressivement. Durant les années 2010 et 2011, la phase d'exploitation se fait à 7 TeV et c'est après un an et demi de prises de données que le LHC atteint 8 TeV. Après l'arrêt technique de 2013/2014, le LHC fonctionnera à  $\sqrt{s} \sim 13$  TeV.

## 2.1 Description du LHC

### 2.1.1 Fonctionnement nominal

Le LHC est situé dans le tunnel du LEP, entre 50 et 170 mètres sous la surface de la terre et sa circonférence est de 27 km. Avant d'être injectés dans cet anneau, les protons sont produits, groupés en paquets puis accélérés par une chaîne d'accélérateurs représentée figure 2.1.

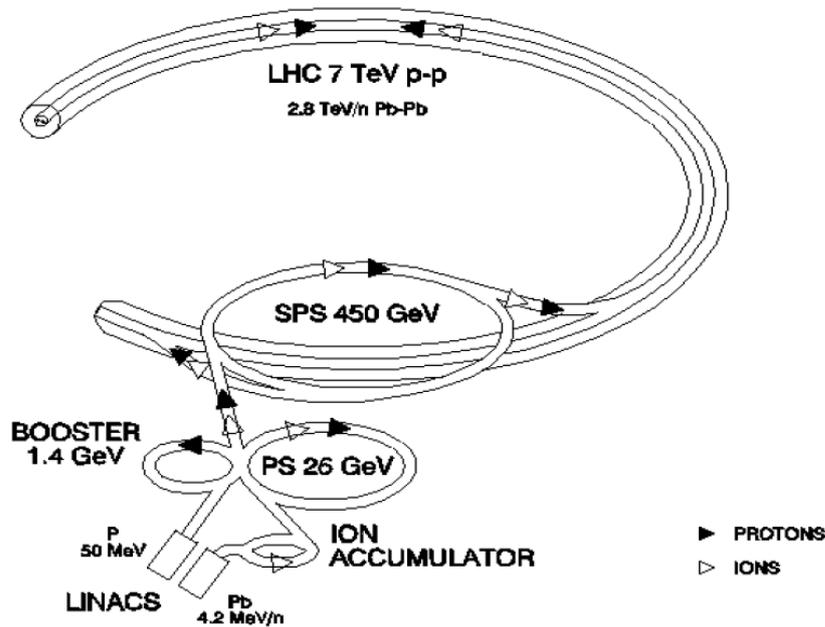


Figure 2.1 – L'ensemble d'injection et d'accélération du LHC [82].

Les protons sont obtenus à partir d'hydrogène gazeux ionisé et sont accélérés dans un premier temps à 750 keV puis 50 MeV par un accélérateur linéaire (LINAC). C'est ensuite l'entrée dans le Booster où ils atteignent une énergie de 1.4 GeV. Ils sont alors injectés dans un premier synchrotron, le PS (Proton Synchrotron) où les paquets ou *bunch* ( $10^{11}$  protons) espacés de 25 ns, en fonctionnement nominal, sont formés et accélérés à 26 GeV pour être introduits dans le SPS qui les accélère à 450 GeV. Les paquets sont finalement injectés dans le LHC (2808 paquets en fonctionnement nominal) et accélérés grâce à des cavités radiofréquences à raison de 485 keV par tour afin d'atteindre leur énergie finale.

La relation entre l'impulsion  $p$  d'une particule chargée et le champ magnétique  $B$  est :

$$p(\text{TeV}/c) = 0.3 \times B(\text{T}) \times R(\text{km}) \quad (2.1)$$

Dans le cas du LHC, avec  $R \simeq 4.3$  km (circonférence de 27 km), il faudrait un champ magnétique de 5.4 T pour obtenir un faisceau de 7 TeV. En réalité, seuls 17.6 km sont équipés d'aimants, ce qui nécessite un champs de 8.4 T. Les 1232 dipôles supraconducteurs de 14.3 m sont refroidis à 2 K et soumis à un courant de 11.9 kA pour générer un tel champ.

On compte quatre points de collisions sur l’anneau du LHC. A chacun d’entre eux est installée une expérience :

- ATLAS (A Toroidal LCH ApparatuS) et CMS (Compact Muon Solenoid) sont deux détecteurs polyvalents dédiés à la recherche du boson de Higgs et de nouvelle physique mais aussi à des mesures de précision.
- LHCb (Large Hadron Collider beauty) explore les caractéristiques du quark b pour étudier l’asymétrie matière et anti-matière.
- ALICE (A Large Ion Collider Experiment) est un détecteur conçu pour l’étude du plasma quark-gluon dans les collisions d’ions lourds.
- LHCf (Large Hadron Collider forward) utilise des particules à très petits angles créées lors des collisions pour simuler des rayons cosmiques en conditions de laboratoire et ainsi aider à l’étalonnage des détecteurs de rayons cosmiques.
- TOTEM (TOTal cross section, Elastic scattering and diffraction Measurement at the LHC) s’intéresse à la mesure de la section efficace totale proton-proton et permet une mesure précise de la luminosité du LHC.

### 2.1.2 Luminosité et événements d’empilement

Une des caractéristiques principales du LHC est sa luminosité. Pour un processus physique donné, on peut exprimer le nombre d’événements par seconde sous la forme :

$$n = \mathcal{L}\sigma \quad (2.2)$$

$\sigma$  est la section efficace du processus et  $\mathcal{L}$  la luminosité instantanée de la machine, c’est-à-dire le nombre de collisions entre les faisceaux par unité de temps et de surface ( $\text{cm}^{-2}\text{s}^{-1}$ ). On définit également la luminosité intégrée sur une période par :

$$L = \int \mathcal{L}dt \quad (2.3)$$

Dans le cas du LHC la luminosité instantanée est donnée par :

$$\mathcal{L} = \frac{N_p n_b^2 f_{rev} \gamma}{4\epsilon_n \beta^*} \times F \quad (2.4)$$

- $N_p$  est le nombre de protons par paquet ( $10^{11}$  en fonctionnement nominal, il a atteint  $\sim 1.5 \times 10^{11}$  en 2012).
- $n_b$  est le nombre de paquets par faisceau.
- $f_{rev}$  est la fréquence de révolution des paquets.
- $\gamma$  est le facteur de Lorentz.
- $\epsilon_n = \gamma\pi\sigma\sigma'$  est l’émittance transverse normalisée des faisceaux.  $\sigma$  et  $\sigma'$  correspondent respectivement à la dispersion des faisceaux en  $x$  et  $x'$  avec  $x' = p_x/p_z$ . L’émittance représente le volume occupé par le faisceau. L’émittance normalisée est un invariant de Lorentz.
- $\beta^*$  est la valeur de la fonction d’amplitude  $\beta = \sigma/\sigma'$  au point de collision. Elle doit être minimisée pour augmenter la luminosité [83].
- $F$  est le facteur de réduction de luminosité géométrique due à l’angle de croisement entre les faisceaux au point de collision. La réduction de luminosité est une conséquence du caractère non frontal des collisions entre les paquets.

L’incertitude sur la mesure de la luminosité délivrée est souvent l’une des incertitudes dominantes dans les mesures de sections efficaces et de rapports d’embranchement, il est donc

primordial de la minimiser. La calibration absolue de la luminosité dans ATLAS est basée sur la méthode des scans de Van Der Meer [84]. Les dernières mesures de luminosité [85], [86] et [87] permettent d'atteindre une incertitude sur la luminosité de 1.8% pour les données à  $\sqrt{s} = 7$  TeV ; pour les données à  $\sqrt{s} = 8$  TeV, une mesure préliminaire donne une incertitude de 3.6 % [88].

Le LHC n'a pas encore atteint son fonctionnement nominal, l'espacement des paquets est de 50 ns mais ses performances sont déjà remarquables. Durant la prise de données à 7 TeV (2010 et 2011), la luminosité instantanée a dépassé  $10^{33} \text{ cm}^{-2}\text{s}^{-1}$  (environ 10 bosons de Higgs Modèle Standard produits par minute) et la luminosité intégrée a atteint  $\sim 5 \text{ fb}^{-1}$ . En 2012 avec les collisions à 8 TeV, le pic de luminosité instantanée était de  $7.7 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ , alors que la luminosité nominale prévue est de  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , et la quantité de données délivrées avoisine les  $25 \text{ fb}^{-1}$ . La figure 2.2 présente la luminosité accumulée en fonction du temps pour les années 2010, 2011 et 2012.

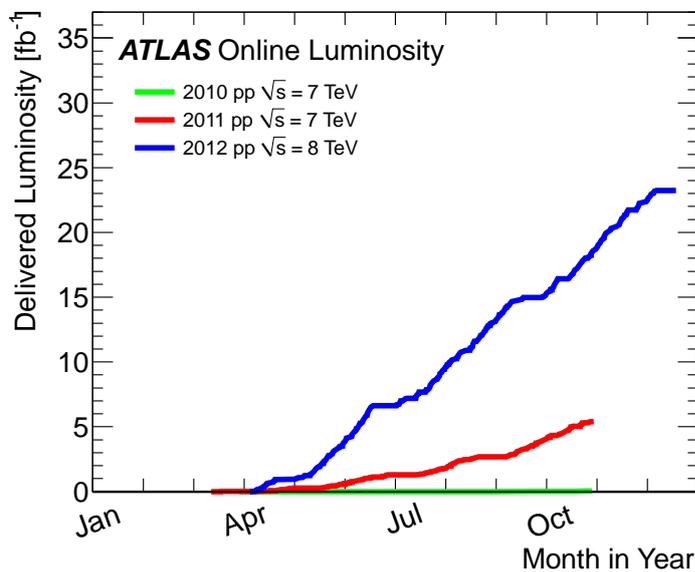


Figure 2.2 – Luminosité cumulée au cours du temps [89].

L'augmentation de la luminosité instantanée a été obtenue en particulier par l'augmentation du nombre de protons par paquet. En 2012 la valeur typique était de  $1.5 \times 10^{11}$  protons par paquets. Avec cette densité de protons dans les paquets, lors du croisement des paquets, plusieurs protons interagissent simultanément. On parle d'événements d'empilement ou *pile-up*. On distingue deux types d'événements d'empilement, les événements provenant des croisements de paquets successifs ou *out of time pile-up* et les événements provenant du même croisement de paquets, il s'agit du *in time pile-up*. Ce dernier est le plus problématique, il est plus difficile de distinguer l'événement d'intérêt des autres événements dans la mesure où ils se produisent tous en même temps. La figure 2.3 représente l'évolution du pic de luminosité instantanée au cours des années 2010, 2011 et 2012, tandis que la figure 2.4 représente l'évolution du nombre maximum d'événements d'empilement par croisement de paquets sur la même échelle de temps, autrement dit le pic d'empilement. La figure 2.5 présente quant à elle la distribution, pondérée en fonction de la luminosité, du nombre moyen d'événements d'empilement par croisement de paquets  $\mu_{BCID}$ . La valeur moyenne de  $\mu_{BCID}$  est de 9.1 pour l'année 2011 et de 20.7 pour l'année 2012, la valeur moyenne attendue en fonctionnement nominal était de 25. Au niveau de la simulation, des interactions de biais minimum sont ajoutées au processus d'interaction dur pour prendre en compte ce phénomène, le nombre d'interactions ajoutées est pondéré en

fonction de la distribution  $\mu_{BCID}$  mesurée dans les données.

Par ailleurs, dans les collisions dures, seuls deux partons interagissent, mais les restes des protons peuvent également contribuer à l'activité dans le détecteur. C'est ce qui est appelé événements sous-jacent ou *underlying events*, ils viennent s'ajouter au phénomène d'empilement.

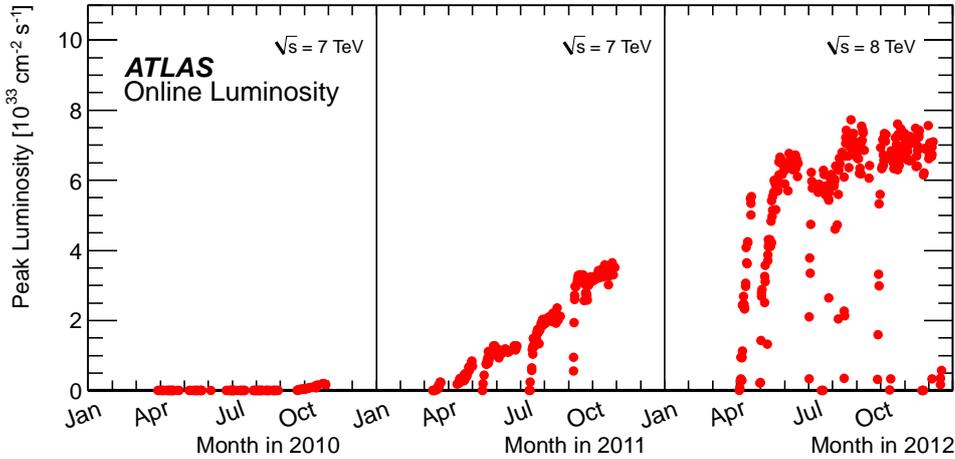


Figure 2.3 – Evolution du pic de luminosité instantanée au cours du temps [89].

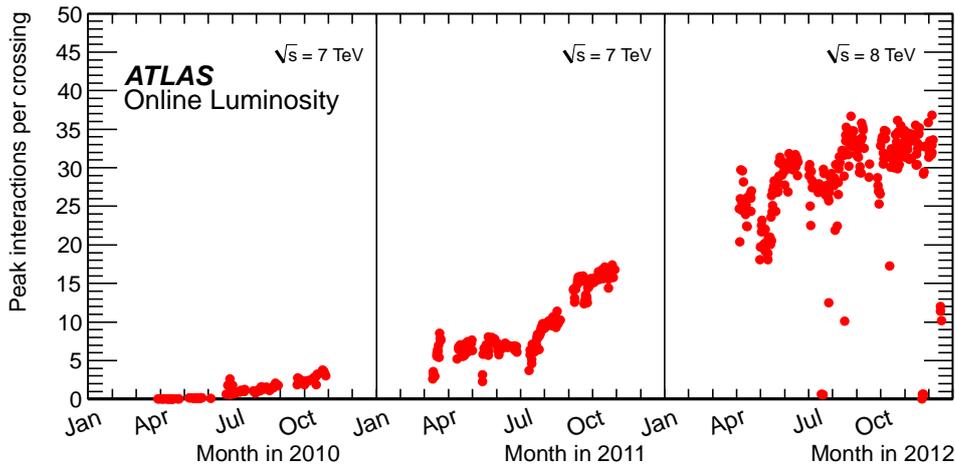


Figure 2.4 – Nombre maximum d'événements d'empilement au cours du temps [89].

## 2.2 Le détecteur ATLAS

Afin de répondre aux objectifs de découverte ainsi qu'aux conditions expérimentales imposées par le LHC (haute luminosité et grand nombre d'événements d'empilement), le détecteur ATLAS a été conçu et construit avec les critères suivant : réponse rapide, fine granularité, large couverture angulaire, système de déclenchement efficace, tenue aux radiations.

### 2.2.1 Vue d'ensemble

Le détecteur ATLAS est présenté figure 2.6. Il fait 44 m de longueur et 25 m de hauteur pour un poids total de 7000 tonnes. Il est constitué de trois ensembles de sous-détecteurs, or-

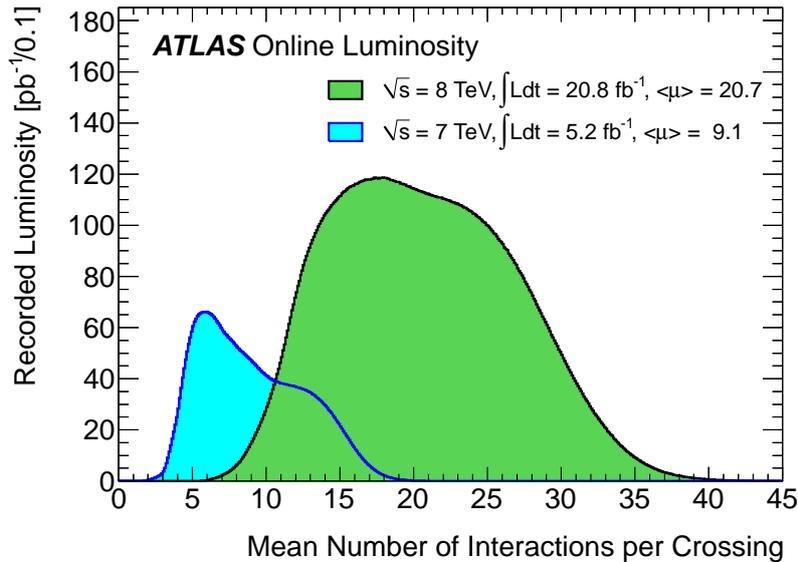


Figure 2.5 – Nombre moyen d'événements d'empilement pour les données 2011 et 2012 [89].

ganisés autour du tube à vide en "couches d'oignon" :

- Le trajectographe est le plus proche du point d'interaction. Baigné dans un champ magnétique solénoïdal, il doit assurer une mesure des trajectoires, impulsions et vertex (d'interaction et de désintégration secondaire) des particules chargées, afin de contribuer, en particulier, à l'identification des électrons, essentielle à l'étude des photons convertis.
- Le calorimètre est composé d'une partie électromagnétique (EM) à argon liquide (ou LAr) qui détermine avec précision l'énergie et la position des objets électromagnétiques (électrons et photons) et d'une partie hadronique qui assure la détermination de l'énergie des jets et de l'énergie transverse manquante.
- Le spectromètre à muons permet l'identification des muons et la mesure de leur impulsion, pour se faire, les chambres à muons sont immergées dans un intense champs magnétique toroïdal.

Chacun des sous-détecteurs est constitué d'une partie centrale en forme cylindrique autour de l'axe du faisceau, appelée tonneau (ou *barrel*) et de deux parties avant perpendiculaires à l'axe du faisceau appelées bouchons (ou *end-cap*).

Dans la suite on s'intéressera plus particulièrement aux systèmes qui sont primordiaux dans le traitement des photons, c'est-à-dire le trajectographe et le calorimètre EM.

Le système de coordonnées d'ATLAS est un système cartésien direct avec pour origine le point d'interaction au centre du détecteur. L'axe  $z$  est confondu avec l'axe du faisceau tandis que les axes  $x$  et  $y$  sont dans un plan transverse à la direction du faisceau. L'axe  $x$  pointe positivement vers le centre de l'anneau du LHC et l'axe  $y$  pointe vers la surface. Par convention, on appelle le côté A la partie du détecteur située en  $z > 0$  (en direction de Genève) et côté C la partie en  $z < 0$  (en direction du Jura). On mesure l'angle azimutal  $\phi$  dans le plan transverse, et l'angle polaire  $\theta$  entre la direction de la particule et l'axe  $z$ , mais la rapidité, invariante de Lorentz, est plus communément utilisée auprès des collisionneurs hadroniques. Elle est définie par :

$$y = \frac{1}{2} \ln \left[ \frac{E + p_z}{E - p_z} \right] \quad (2.5)$$

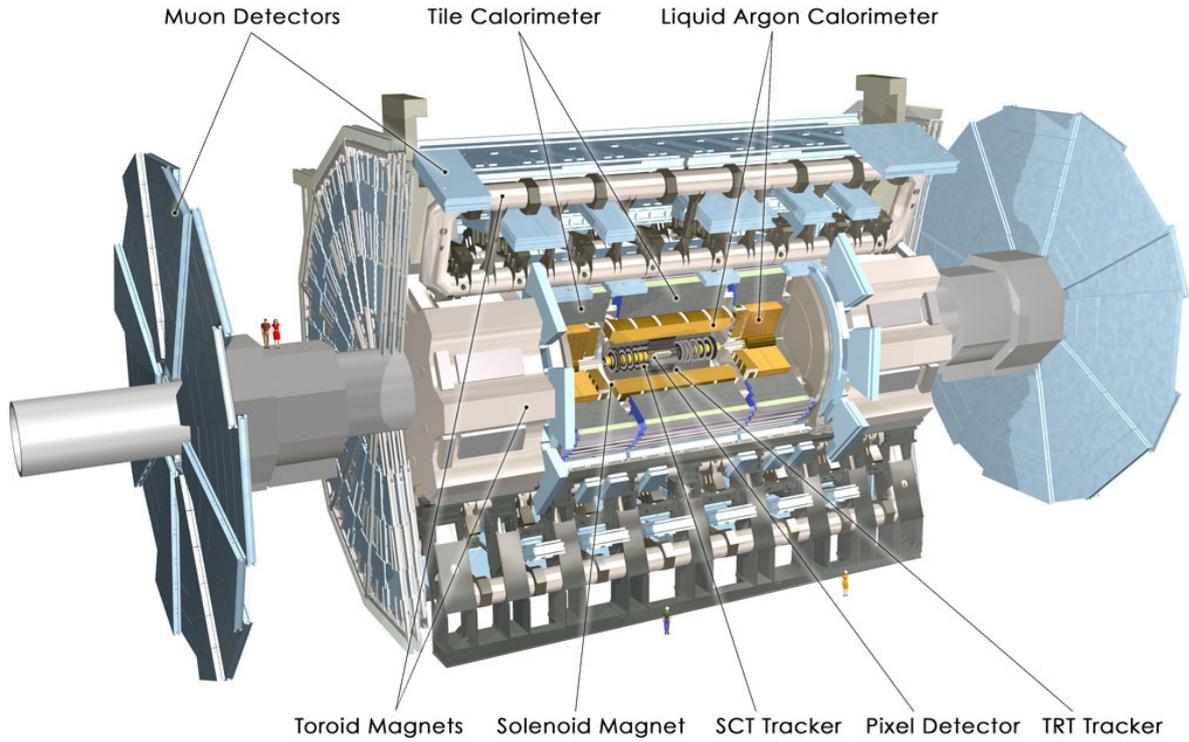


Figure 2.6 – Vue d'ensemble du détecteur ATLAS [90].

$E$  et  $p_z$  sont respectivement l'énergie et la composante selon  $z$  de l'impulsion. Dans le cas d'une particule relativiste la pseudo-rapidité est une bonne approximation de la rapidité :

$$\eta = -\ln \left| \tan \frac{\theta}{2} \right| \quad (2.6)$$

Le flux de particules produites dans les collisions est constant par unité de pseudo-rapidité ; c'est pourquoi les détecteurs sont en général segmentés en canaux de lecture avec une largeur constante en pseudo-rapidité.

### 2.2.2 Le détecteur interne

Le détecteur interne d'ATLAS est un trajectographe de forme cylindrique de 5.5 m de long avec un rayon de 1.15 m. Il a pour but de mesurer l'impulsion des particules chargées dans un large domaine en pseudo-rapidité  $|\eta| < 2.5$  et de permettre la reconstruction des vertex primaires et secondaires. Pour courber la trajectoire des particules chargées, il est immergé dans un champ magnétique solénoïdal de 2 T. Il est constitué de trois sous-détecteurs : le détecteur à pixels, le détecteur à micro-pistes de silicium (ou SCT pour *Semi-Conductor-Tracker*) et le détecteur à rayonnement de transition (ou TRT pour *Transition-Radiation-Tracker*). La figure 2.7 présente une vue schématisée de l'ensemble du trajectographe.

Le détecteur à pixels (silicium) a pour fonction principale de reconstruire les vertex d'interactions ; il fournit trois points de mesure de la trajectoire à proximité du faisceau. Dans la partie tonneau, il est constitué de trois couches cylindriques et concentriques autour de l'axe du faisceau ( $R = 50.5, 85.5$  et  $122.5$  mm). La première couche, la plus proche du faisceau, est appelée couche-b (ou *b-layer*), elle est essentielle à la reconstruction des vertex, en particulier ceux issus de la désintégration d'un quark b (ou méson B) et donc à l'étiquetage des jets de b. Elle amène également une discrimination entre électrons, généralement reconstruits avec au moins un "hit" dans la couche-b, et photons convertis qui, la plupart du temps, ne laissent

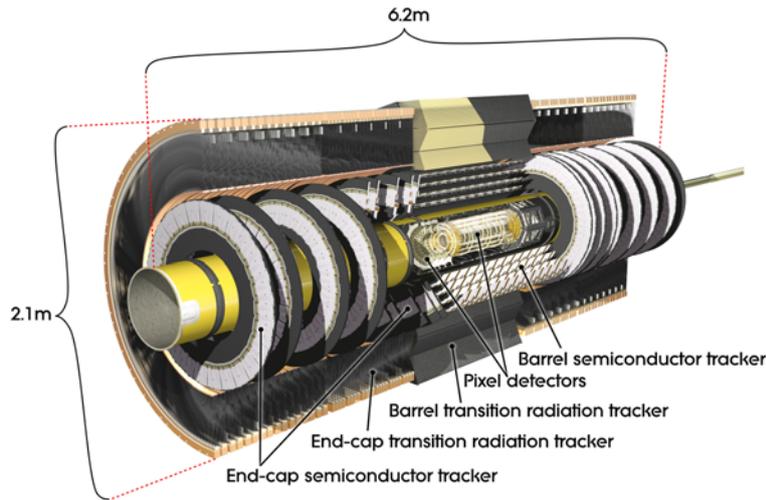


Figure 2.7 – Le détecteur interne d'ATLAS [90].

pas de hit dans cette couche (sauf si la conversion a lieu avant ou dans la couche-b). Dans chaque bouchon, on trouve deux séries de trois disques perpendiculaires à l'axe du faisceau ( $z = \pm 495, \pm 580, \pm 650$  mm), ils permettent de couvrir la zone  $|\eta| < 2.5$ . Le détecteur à pixels compte 1744 capteurs avec chacun 46080 voies de lecture, soit environ 80 millions de voies de lecture au total. La taille d'un pixel est de  $50 \times 400 \mu\text{m}^2$ . La précision attendue pour la position est de l'ordre de  $14 \mu\text{m}$  dans le plan  $R - \phi$  et de  $115 \mu\text{m}$  dans la direction longitudinale (resp. radiale) pour le tonneau (resp. les bouchons).

Une quatrième couche, appelée IBL (*Insertable Blayer*), qui permettra d'améliorer l'identification des quarks b est en construction et sera installée au printemps 2014.

Le détecteur à semi-conducteurs ou SCT est formé de micro-pistes en silicium, il participe à la reconstruction des trajectoires des particules chargées et des vertex, en complément du détecteur à pixels. La partie tonneau est constituée de quatre couches cylindriques concentriques ( $R = 299, 371, 443, 514$  mm) et couvre la partie centrale, jusqu'à  $|\eta| < 1.1$ . Chaque bouchon comporte neuf disques et assure une couverture jusqu'à  $|\eta| < 2.5$ . Chaque module contient deux plaques de détecteurs dos-à-dos, fournissant un double point d'impact pour les traces le traversant. Le SCT compte environ 6.3 millions de voies de lecture. La précision attendue pour la position est de l'ordre de  $17 \mu\text{m}$  dans le plan  $R - \phi$  et de  $580 \mu\text{m}$  dans la direction longitudinale (resp. radiale) pour le tonneau (resp. les bouchons).

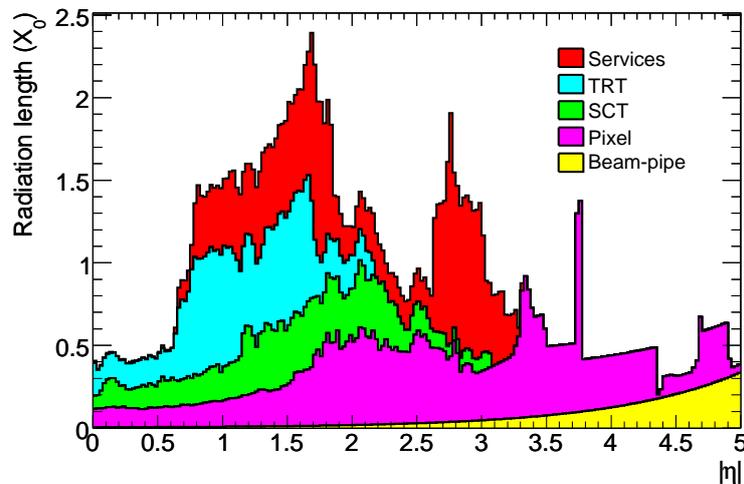
Le détecteur à rayonnement de transition ou TRT est constitué de tubes (ou pailles) à dérive de 4 mm de diamètre contenant un mélange gazeux Xe-CO<sub>2</sub> (70% Xe, 27% CO<sub>2</sub> et 3%O<sub>2</sub>) et un fil de tungstène. Quand une particule chargée traverse le tube à dérive, elle ionise le mélange gazeux et délivre un signal. Le fil de tungstène, au centre du tube, constitue l'anode et est connecté aux voies de lecture. La partie centrale est constituée de 53000 tubes répartis sur trois cylindres de 160 cm de long et chaque bouchon compte 125000 pailles sur trois roues. La zone de couverture du TRT s'étend jusqu'à  $|\eta| = 2.0$  et le nombre total de voies de lecture est d'environ 350000. Le TRT fournit un grand nombre de points d'impact ( $\sim 36$ ) par trace, et il sert également à distinguer les électrons des pions chargés. En effet, les tubes sont entourés de polypropylène/polyéthylène, et lorsqu'une particule chargée arrive à une interface entre deux matériaux dont les constantes diélectriques sont différentes, elle émet un rayonnement de transition qui dépend du facteur  $\gamma$  de Lorentz. Pour une même énergie, les électrons,  $\sim 250$  fois moins massifs que les pions, engendrent un rayonnement de transition beaucoup plus important que les pions. Des seuils sur l'amplitude des signaux collectés ont été définis pour permettre une discrimination entre électrons et pions. La précision attendue pour

la position est de l'ordre de  $130 \mu\text{m}$  dans le plan  $R - \phi$ .

En combinant les trois sous-détecteurs (une trace reconstruite dans la partie tonneau a typiquement trois impacts dans les pixels, huit dans les SCT et trente-six dans les TRT), la résolution attendue sur la mesure de l'impulsion transverse d'une particule chargée est donnée par :

$$\frac{\sigma_{P_T}}{p_T} = 0.05\% p_T(\text{GeV}) \oplus 1\% \quad (2.7)$$

La quantité de matière dans le détecteur interne moyennée sur  $\phi$  est représentée en unité de longueur de radiation en fonction de  $|\eta|$  figure 2.8. La probabilité pour un photon de se convertir en paire  $e^+e^-$  dépend directement de cette quantité de matière, sa connaissance est donc primordiale afin de simuler correctement les photons convertis et de calculer la fraction de conversions.

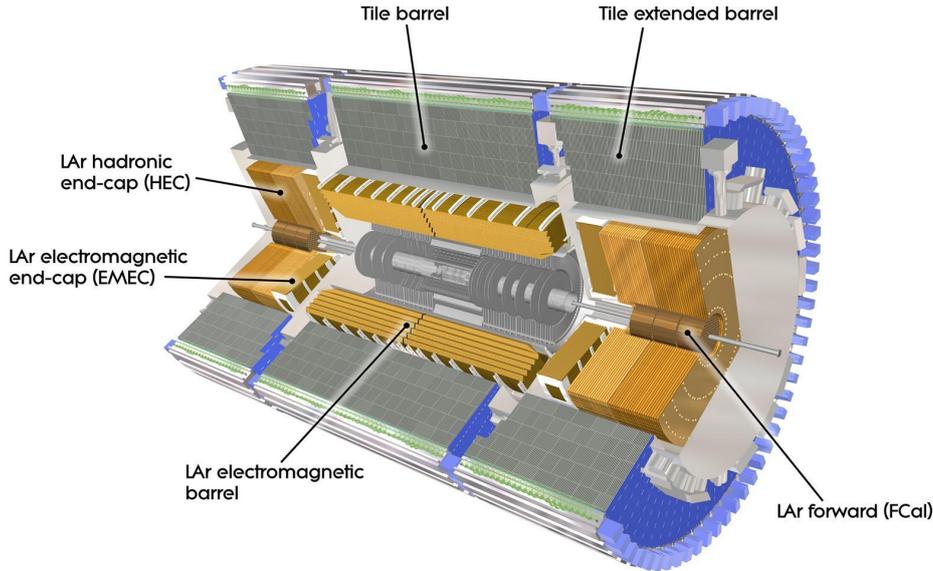


**Figure 2.8** – Quantité de matière (en longueur de radiations) dans le détecteur interne, en fonction de  $|\eta|$  [91].

### 2.2.3 Les calorimètres

Le système calorimétrique d'ATLAS est composé de trois parties. Le calorimètre électromagnétique (EM) mesure l'énergie et la direction des photons et des électrons dans la région  $|\eta| < 3.2$ , et il dispose d'une granularité plus fine jusqu'à  $|\eta| = 2.37$ . Le calorimètre hadronique, dans la région  $|\eta| < 3.2$  permet de contenir totalement l'énergie des jets non déposée dans le calorimètre EM. Enfin, dans la région  $3.2 < |\eta| < 4.9$ , les calorimètres avant, qui comportent une partie électromagnétique et une partie hadronique, permettent de mesurer les jets mais aussi les électrons à très petits angles et améliore ainsi la mesure de l'énergie transverse manquante.

La figure 2.9 montre l'ensemble du système calorimétrique d'ATLAS. Les calorimètres d'ATLAS sont des calorimètres à échantillonnage : une gerbe de particules est initiée dans un matériau dense (l'absorbeur), et une partie de l'énergie déposée dans le milieu actif (détecteur) est mesurée. Deux technologies sont utilisées, la calorimétrie à argon liquide et la calorimétrie à scintillateurs. Le calorimètre électromagnétique (tonneau et bouchons), les bouchons hadroniques et les calorimètres avant utilisent la technologie de l'argon liquide ; le reste du calorimètre hadronique utilise le scintillateur.



**Figure 2.9** – Le système calorimétrique d'ATLAS, composé d'un calorimètre électromagnétique (argon liquide) et d'un calorimètre hadronique (tuiles pour la partie central, argon liquide pour la partie avant [90]).

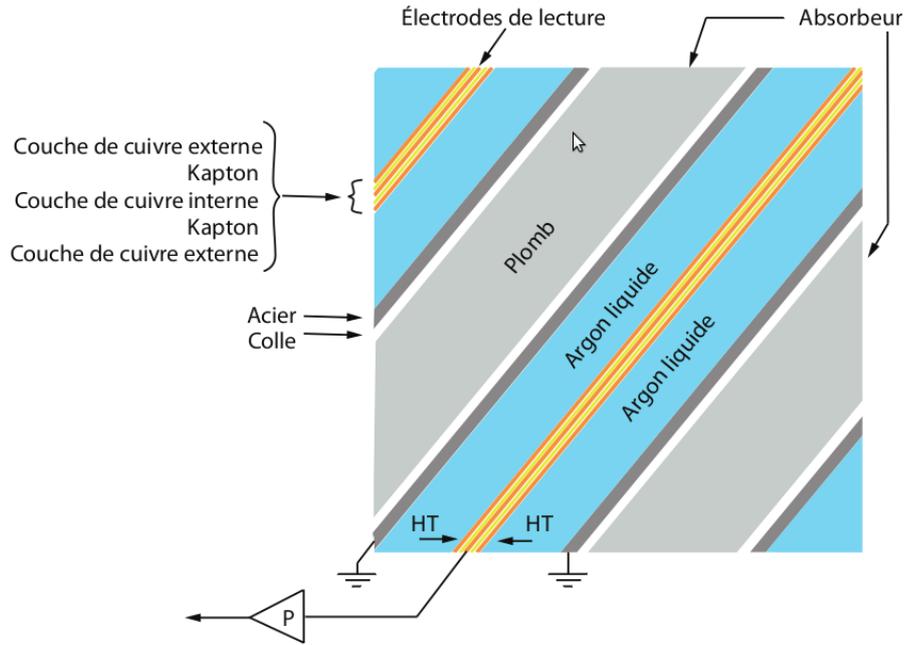
### 2.2.3.1 Le calorimètre électromagnétique (EM) à argon liquide (LAr)

La zone de couverture du calorimètre électromagnétique LAr s'étend jusqu'à  $|\eta| < 3.2$ . Il comprend une partie tonneau (EMB)  $|\eta| < 1.475$  et deux bouchons (EMEC)  $1.375 < |\eta| < 3.2$ , dont l'épaisseur totale est respectivement de 22 (à  $|\eta|=0$ ) et 24 longueurs de radiation  $X_0$ . Le tonneau est contenu dans le même cryostat que le solénoïde central et les bouchons sont chacun contenus dans un cryostat séparé. La zone de transition entre tonneau et bouchon  $1.37 < |\eta| < 1.52$  où passe les câbles et les services du détecteur interne est faiblement instrumentée et dénommée *crack*. Cette région est généralement exclue des analyses de physique car la mesure de l'énergie  $y$  est dégradée. La partie tonneau du calorimètre électromagnétique est divisée en deux demi-tonneaux identiques et distants de 4 mm en  $z = 0$ . Chaque bouchon est quant à lui constitué d'une roue externe,  $1.375 < |\eta| < 2.5$  et d'une roue interne,  $2.5 < |\eta| < 3.2$ . Chaque roue contient huit modules et chaque demi-tonneau en contient seize.

Un module du calorimètre EM est composé d'absorbeurs en plomb et d'argon liquide qui constitue le milieu actif. L'électrode, qui collecte le signal, est faite de trois couches de cuivre séparées par des couches de kapton comme présenté figure 2.10. Une haute tension, nominale 2000 V, est appliquée entre l'absorbeur et les couches externes de cuivre tandis que la couche interne récolte, par couplage capacitif, le courant de dérive induit par les électrons d'ionisation dans l'argon liquide. Afin d'assurer l'herméticité en  $\phi$  et une extraction rapide du signal à l'avant ou à l'arrière de l'électrode, absorbeurs et électrodes sont pliées en forme d'accordéon [92].

Le calorimètre EM est segmenté longitudinalement en trois couches comme présenté figure 2.11 :

- La première couche ou *front*, de profondeur  $\sim 4.4 X_0$ , jusqu'à  $|\eta| < 2.37$  est formée de cellules également appelées *strips* de très fine granularité en  $\eta$ .  $\Delta\eta \times \Delta\phi = 0.025/8 \times 0.1$  dans la région  $|\eta| < 1.4$ . Dans les bouchons la granularité est comparable mais varie en fonction de  $|\eta|$ . Cette première couche avec une excellente segmentation en  $\eta$  joue un rôle particulièrement important pour la séparation des gerbes spatialement proches, comme dans le cas de deux photons provenant de la désintégration d'un  $\pi^0$ .



**Figure 2.10** – Représentation schématique d'une partie d'un module du calorimètre EM [93].

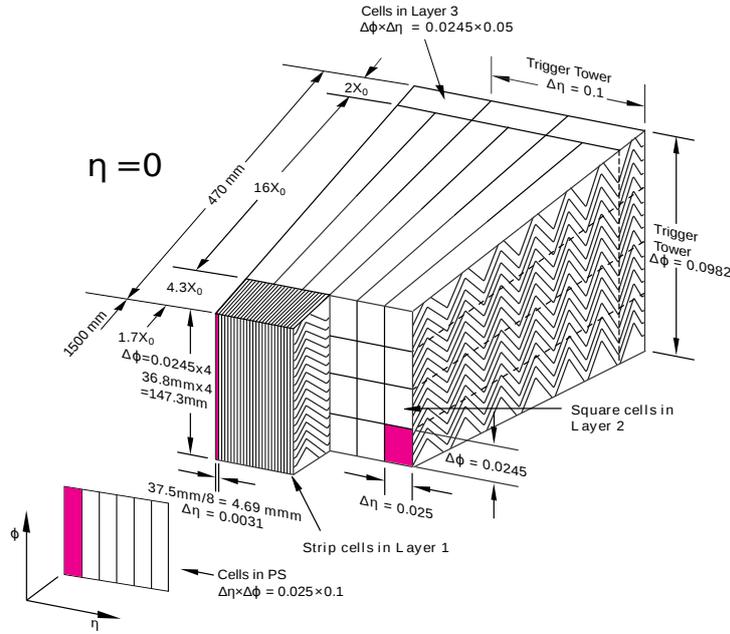
- La seconde couche ou *middle*, dont la profondeur varie de 16 à 18  $X_0$ , reçoit la plus grande partie de l'énergie. Elle a une granularité de  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  et assure avec la première couche, la mesure de la pseudo-rapacité  $\eta$  de la particule incidente et la direction des photons dans le plan  $(r, z)$ .
- La troisième couche ou *back*, de profondeur 2  $X_0$  (à  $|\eta|=0$ ), sert à estimer les pertes vers l'arrière de gerbes non contenues dans le calorimètre EM. Sa granularité est de  $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ .

Dans la région  $|\eta| < 1.8$ , le calorimètre EM est précédé par un pré-échantillonneur (*pre-sampler*) de granularité  $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$ . Son rôle principal est l'estimation des pertes d'énergies des photons et électrons dans la matière située en amont du calorimètre.

Les performances imposées au calorimètre EM ont été optimisées pour la découverte du boson de Higgs dans les canaux  $H \rightarrow \gamma\gamma$  et  $H \rightarrow ZZ^{(*)} \rightarrow 4e$ . Elles concernent la résolution angulaire  $50 \text{ mrad}/\sqrt{E(\text{GeV})}$  et la résolution en énergie du calorimètre EM [95] :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.8)$$

- Le terme  $a/\sqrt{E}$  est le terme d'échantillonnage. Il décrit les fluctuations statistiques liées au développement de la gerbe électromagnétique dans l'argon liquide. Sa valeur est de l'ordre de 10% à  $\eta = 0$ .
- Le terme  $b/E$  est le terme de bruit. Il décrit les fluctuations venant principalement de la chaîne électronique et des événements d'empilement. Sa valeur est de l'ordre de 300 MeV, mais varie avec la luminosité et est particulièrement importante à basse énergie.
- Le terme  $c$  est le terme constant, le plus difficile à estimer. Il reflète la non-uniformité de la réponse du calorimètre. Il est principalement dû aux non-uniformités des matériaux, à la qualité de l'étalonnage et à la linéarité de la réponse. Il domine la résolution à haute



**Figure 2.11** – Schéma des couches successives d'absorbeurs, argon liquide, électrodes composant le calorimètre EM [94].

énergie notamment dans le canal  $H \rightarrow \gamma\gamma$ . La valeur attendue est de l'ordre de 0.7% dans le tonneau. Les dernières mesures effectuées en 2011 estiment le terme constant à  $\sim 1\%$  dans le tonneau et entre 1.2 et 2.1% dans les bouchons.

### 2.2.3.2 Le calorimètre hadronique

Le calorimètre hadronique est divisé en deux parties, le calorimètre à tuile ou *Tile calorimeter* dans la région  $|\eta| < 1.7$  et les bouchons hadroniques (HEC) dans la région  $1.5 < |\eta| < 3.2$ .

Le calorimètre à tuiles est constitué d'une seule partie tonneau dans la région  $|\eta| < 1.0$  mais de deux parties tonneaux, ou tonneaux étendus, dans la région  $0.8 < |\eta| < 1.7$ . Il comporte près de 10 000 voies de lecture. Il s'agit d'un calorimètre à scintillateurs, il est formé d'une alternance de tuiles en polystyrène dopé constituant le milieu actif et de couches d'acier constituant le milieu passif. La lumière produite par les molécules du milieu actif excitées lors du passage des gerbes hadroniques constitue le signal, qui est acheminé vers des photomultiplicateurs grâce à des fibres optiques.

Les bouchons hadroniques sont logés dans les mêmes cryostats que les bouchons électromagnétiques associés et sont composés de deux roues indépendantes. L'absorbeur est ici le cuivre.

Les performances requises pour les calorimètres hadroniques sont moins contraignantes que pour le calorimètre EM, une grande partie de l'énergie des jets est déposée dans le calorimètre EM (en moyenne 2/3 de l'énergie pour des jets  $> 100 \text{ GeV}$ ) et les gerbes hadroniques fluctuent beaucoup d'un événement à l'autre :

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \quad (2.9)$$

### 2.2.3.3 Le calorimètre avant

Les calorimètres avant (FCal) étendent la couverture en pseudo-rapacité jusqu'à  $|\eta| < 4.9$ . Ils utilisent également la technologie à argon liquide et sont situés dans les mêmes cryostats que les calorimètres EM bouchons. Ils sont constitués d'un module électromagnétique utilisant le cuivre comme absorbeur et de deux modules hadroniques utilisant du tungstène comme absorbeur.

Les performances nominales de ce détecteur sont :

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \quad (2.10)$$

### 2.2.4 Le système de déclenchement

En 2012, le taux d'interactions par seconde atteint est de 1 milliard, et le détecteur ATLAS est équipé d'une centaine de millions de voies de lecture. Ce flux de données est beaucoup trop élevé pour pouvoir stocker toutes les cellules de toutes les interactions. Le système de déclenchement à trois niveaux permet de réduire le flux de données de 20 MHz à 400 Hz. Les trois niveaux de déclenchement dans ATLAS utilisent des informations de complexité croissante, comme illustré sur le schéma figure 2.12 :

- Le niveau 1 ou L1 : Le déclenchement se fait de manière électronique. Il utilise les informations des calorimètres (avec une granularité  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ) et des spectromètres à muons pour définir des régions d'intérêt RoI (*Regions of Interest*) quand l'énergie déposée dans une zone est supérieure à un certain seuil. Lorsqu'un événement passe les critères de déclenchement, il est susceptible d'être transmis au niveau suivant. Cependant, afin de maintenir le taux d'événements sous la limite des capacités de traitement de 75 kHz, une fraction des événements doit être rejetée de façon aléatoire pour certains triggers. La fraction d'événement ainsi rejetés est appelée *pre-scale*. La durée maximale de prise de décision du niveau L1 est de 2.5  $\mu$ s.
- Le niveau 2 ou L2 : Le déclenchement se fait au niveau logiciel. Il élabore une décision en utilisant des informations provenant de l'ensemble des sous-détecteurs sur la zone de la RoI. Des candidats objets (reconstruction simplifiée, uniquement dans les RoI) sont alors déterminés, des "menus" sont définis en fonction du type et du nombre d'objets reconstruits. Le taux de déclenchement ne doit pas dépasser les 3.5 kHz, la décision est prise en moins de 40 ms.
- Le niveau 3 ou EF (*Event Filter*) : Le déclenchement se fait également au niveau logiciel. A ce stade, les événements sont reconstruits en utilisant les mêmes algorithmes de reconstruction et d'identification des objets que pour le traitement "offline" des données. La fréquence d'enregistrement est typiquement de 400 Hz.

Les différents menus définissent des flots de données ou *stream*. Les streams sont divisées en deux catégories : les streams de calibration et les streams de physique. Les streams de physique sont constituées des événements destinés à l'analyse physique, classées par signature (muons, objets EM...). Les streams de calibration, quant à elles, fournissent des informations détaillées sur les conditions de la prise de donnée (luminosité, empilement...) et l'état du détecteur. Elles sont indispensables pour établir la qualité des données enregistrées par ATLAS.

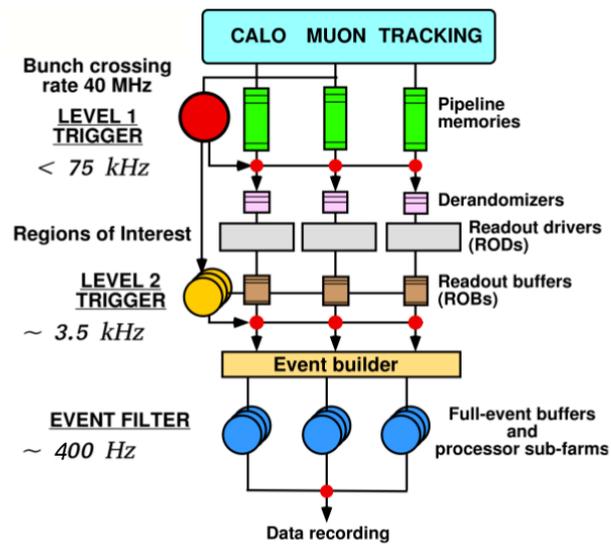
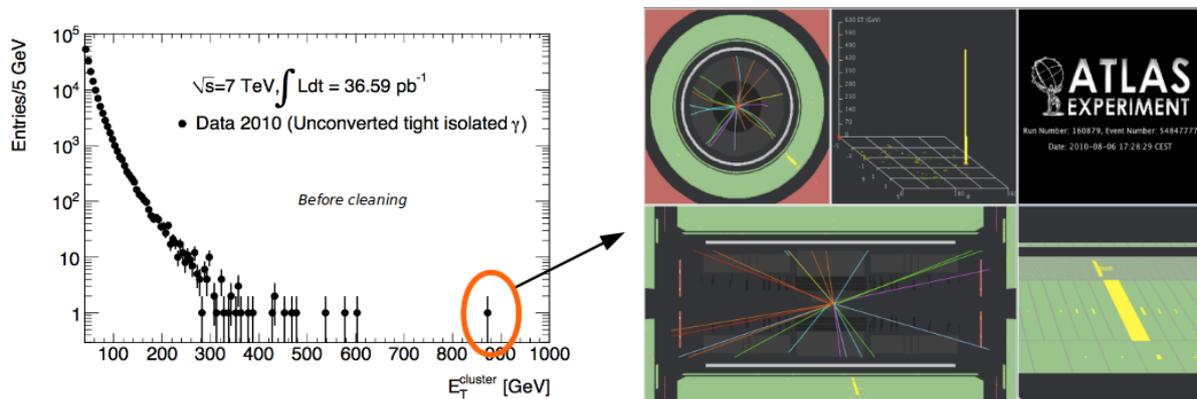


Figure 2.12 – Schéma du système de déclenchement d'ATLAS

## Chapitre 3

# Identification des cellules bruyantes du calorimètre à argon liquide

Une description du détecteur ATLAS a été donnée au chapitre précédent et nous avons vu plus en détails le détecteur interne et les calorimètres, essentiels dans le traitement des photons. Avant d'aborder, au chapitre suivant, la reconstruction et calibration des photons, nous attarderons ici sur la démarche qualité appliquée dans ATLAS et plus particulièrement dans le calorimètre à argon liquide. La garantie de la qualité des données est essentielle pour les analyses de physique. Elle permet par exemple de rejeter un candidat photon de haute énergie (figure 3.1), passant toutes les coupures d'identification et d'isolation usuelles, mais correspondant à du bruit anormal dans des cellules du calorimètre EM.



**Figure 3.1** – Candidat photon dans les données 2010 à  $\sqrt{s} = 7 \text{ TeV}$ . Ce candidat passe les coupures d'identification et d'isolation usuelles mais correspond en réalité à un comportement anormal de cellules du calorimètre EM.

Nous expliquerons dans un premier temps comment est faite l'acquisition des données dans ATLAS. Nous nous concentrerons ensuite sur la chaîne d'acquisition du calorimètre à argon liquide. Après avoir compris la reconstruction de l'énergie d'une cellule du calorimètre à argon liquide, nous verrons comment monitorer les 200 000 cellules et identifier celles qui sont anormalement bruyantes. Enfin, une fois les cellules problématiques identifiées, nous présenterons les moyens utilisés pour leur traitement lors de la reconstruction des données.

### 3.1 Acquisition des données dans ATLAS

L'acquisition des données est dirigée par le TDAQ (*Trigger and Data Acquisition system*) et le DCS (*Detector Control System*). Le TDAQ gère la configuration du détecteur et du système de déclenchement. Le DCS, quant-à-lui permet de configurer les équipements et contrôle l'ensemble des systèmes relatifs à chaque sous-détecteur ; il fournit des informations en temps réel afin d'identifier d'éventuelles défaillances qui compromettraient de façon irrémédiable la qualité des données.

Les données enregistrées par ATLAS sont regroupées en *runs*. Chaque run est identifié par un numéro ou *runnumber* et est constitué de séquences de données appelées *lumiblocks*. La durée typique d'un lumiblock est de 120 secondes en 2010 et de 60 secondes en 2011 et 2012. La qualité des données est estimée par lumiblock afin de limiter la perte des données.

Les données brutes collectées nécessitent un traitement afin d'identifier les défauts potentiels et d'appliquer des corrections *ad-hoc* lors de la reconstruction. Une procédure d'étalonnage, appelée *calibration loop*, basée sur les streams de calibration (introduites section 2.2.4) est mise en oeuvre durant les 48 heures suivant la fin d'un run (figure 3.2). Les streams de calibration sont processées rapidement après la fin du run tandis que les streams de physique sont mises en attente pendant que la calibration est effectuée. Pendant la calibration loop, chacun des défauts identifiés est consigné dans une base de donnée dédiée et les lumiblocks affectés par ce défaut sont spécifiés. Un défaut peut être commun à tout ATLAS, ou n'affecter qu'une seule partie du détecteur. Selon la gravité du défaut et la possibilité de le corriger ou non, il est marqué comme tolérable ou intolérable. A l'issue des 48 heures, lorsque toutes les bases de données sont remplies, les streams de physique commencent à être processées, c'est le *bulk processing*. Après le bulk processing, les lumiblocks toujours affectés par un défaut intolérable sont systématiquement exclus de la liste de lumiblocks autorisés pour l'analyse physique des données, les *Good Run List* ou GRL. Une semaine s'écoule entre la prise de données et la mise à disposition des données jugées bonnes pour les analyses physiques (3.2). En 2012, plus de 95% des données collectées ont pu être analysées.

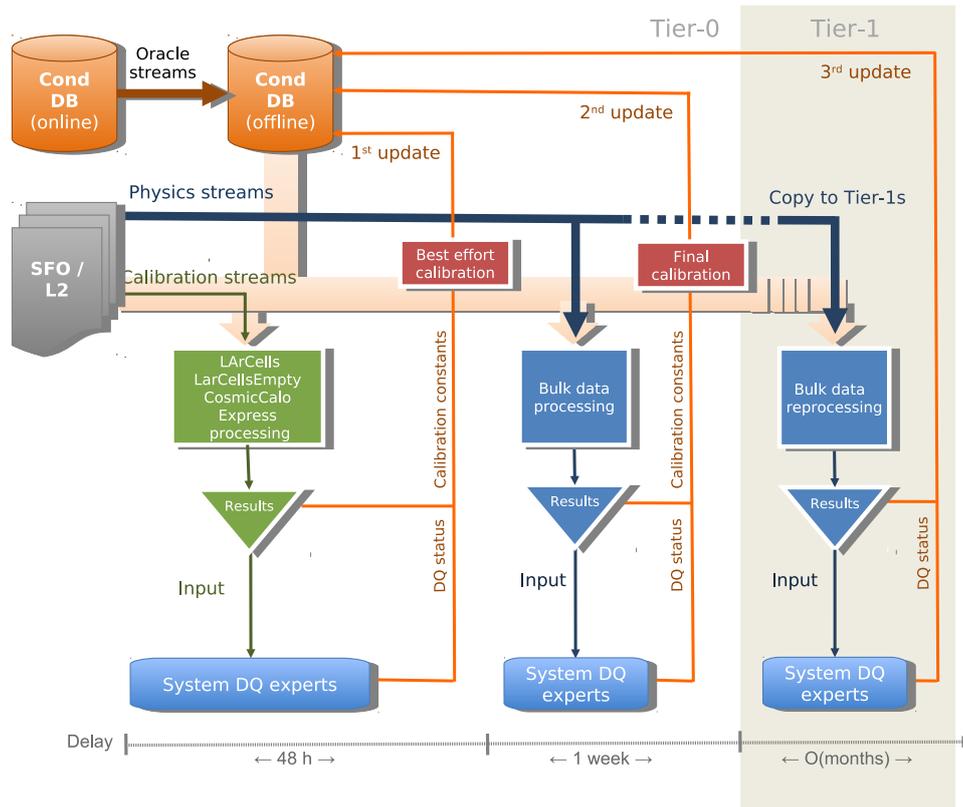
Dans le cas spécifique du calorimètre à argon liquide, quatre streams sont considérées. L'*Express* stream ( $\sim 5\%$  de l'ensemble des données) et la *LArCells* stream contiennent des collisions ; elles sont utilisées pour identifier, entre autres, des problèmes de corruption des données. Les streams qui contiennent des collisions sont celles enregistrant les événements liés aux collisions de paquet contenant des protons. Toutefois les paquets ne sont pas nécessairement remplis de protons, les streams *CosmicCalo* et *LArCellsEmpty* enregistrent les événements liés à ces paquets "vides". Aucune collision n'étant attendue, ces streams permettent de surveiller les 182468 cellules du calorimètre à argon liquide et de juger d'un comportement anormal qui doit être corrigé ou pris en compte.

### 3.2 Chaîne d'acquisition du calorimètre à argon liquide

La chaîne d'acquisition des calorimètres à argon liquide est présentée de manière schématique sur la figure 3.3.

La chaîne d'électronique de lecture est divisée en deux systèmes : le système frontal (*front-end*), situé sur le détecteur, et le système arrière (*back-end*), situé dans la salle de comptage d'ATLAS. Les cartes d'acquisition frontales ou FEB (*Front-End Boards*), installées sur les traversées étanches des cryostats, amplifient, mettent en forme et numérisent le signal des cellules du calorimètre à argon liquide. Chaque FEB, au total 1524 FEBs, est connectée à 128 cellules.

Le signal induit par le dépôt d'énergie d'une particule est un courant provenant de la dérive des électrons d'ionisation dans l'argon, sous l'effet de la haute tension appliquée entre l'absorbéur



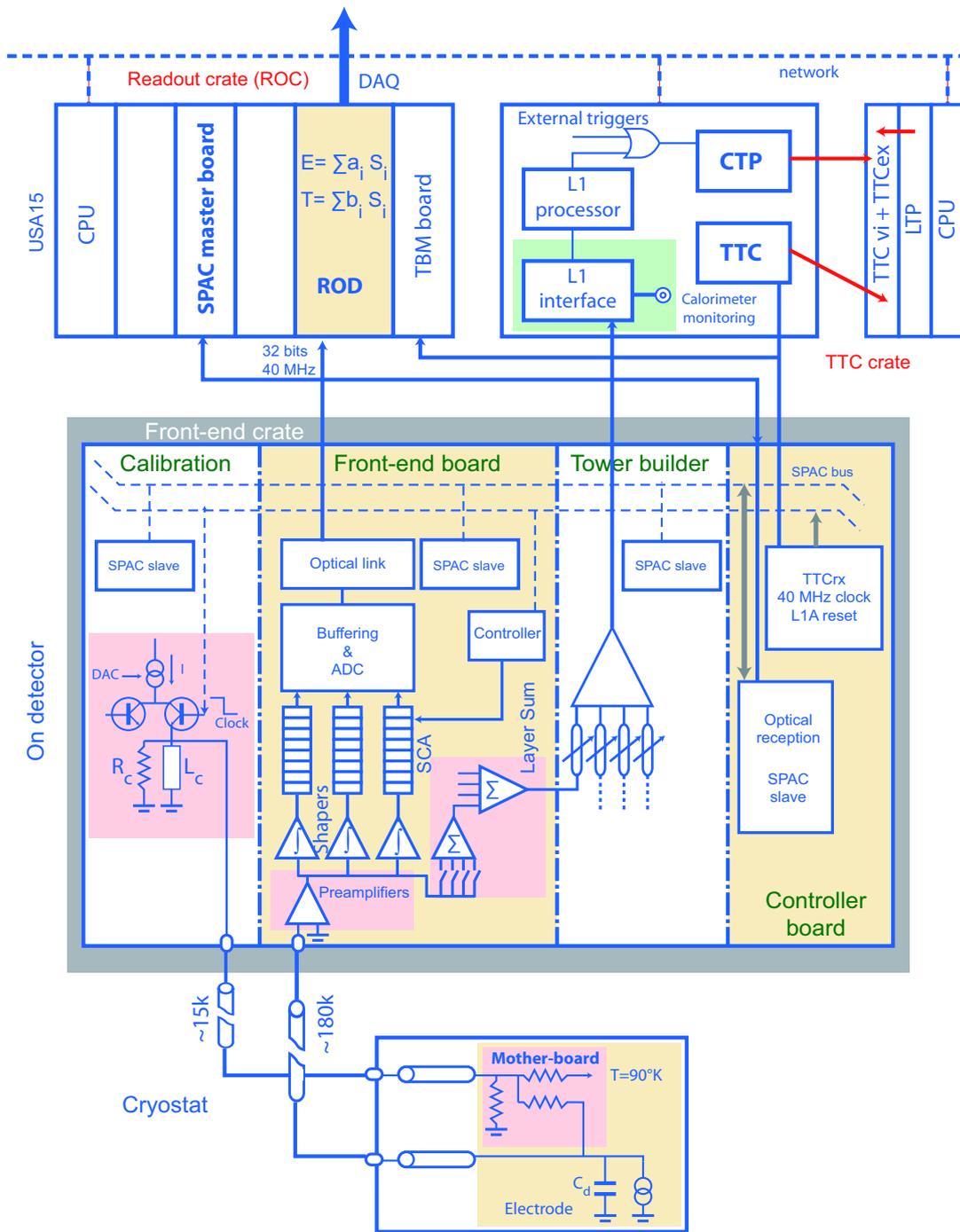
**Figure 3.2** – Organisation de la chaîne de reconstruction et du monitoring des données dans ATLAS [96].

et l'électrode (comme décrit section 2.2.3); il s'agit d'un signal analogique triangulaire. Dans les FEBs, ce signal est d'abord pré-amplifié avant de passer dans un module de mise en forme ou *shaper*. Ce module, constitué d'un filtre  $CR - (RC)^2$ , confère au signal une forme bipolaire. La forme bipolaire du signal (intégrale nulle), présentée figure 3.4, est primordiale pour réduire la contribution moyenne des événements d'empilement. En effet, le temps qui sépare deux paquets de protons au LHC (50 ns en 2011-2012 ou 25 ns en fonctionnement nominal) est plus court que le temps nécessaire pour collecter le signal d'ionisation (400 ns) et il en résulte un empilement des signaux. Les modules de mise en forme produisent en réalité trois signaux mis en forme chacun avec un gain différent : haut, moyen, bas. Les signaux sont ensuite échantillonnés à la fréquence du LHC par des *switched-capacitor array* (SCA). En mode de fonctionnement normal, cinq échantillons par signal sont conservés. Ils sont stockés dans trois pipelines analogiques en attendant la réponse du système de déclenchement de niveau 1. Si la réponse est positive, les échantillons du gain sélectionné (en fonction de l'énergie de la cellule) sont transmis au convertisseur ADC 12 bits, puis transmis vers la salle de comptage où les données sont traitées par les cartes ROD (*Read Out Driver*) en moins de 10  $\mu$ s.

Au niveau des ROD, le processeur DSP (*Digital Signal Processing*) reconstruit l'énergie des cellules suivant la formule :

$$E = f \times G(A_{ADC}) \quad (3.1)$$

- $f$  est un facteur de conversion dépendant de constantes d'étalonnage, du test-beam et de la simulation.
- $G$  est un polynôme, généralement écrit  $G(A) = R_0 + R_1 \times A_{ADC}$ , il permet de convertir un signal numérique en signal analogique.

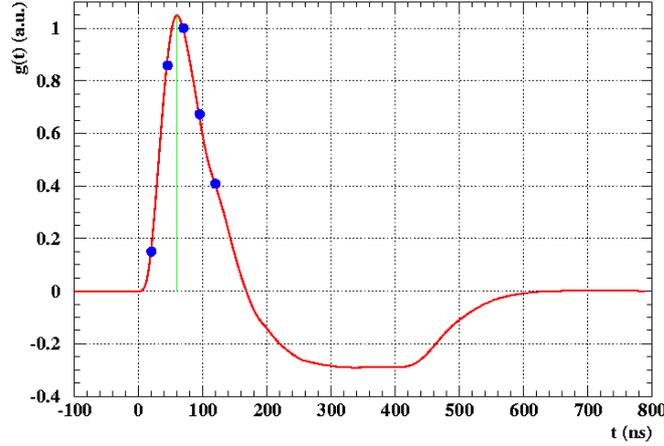


**Figure 3.3** – Diagramme de l'architecture générale de l'électronique de lecture du calorimètre à argon liquide. La partie inférieure représente le circuit électronique froid c'est-à-dire dans le cryostat. La partie centrale montre l'électronique frontale située sur le détecteur en dehors du cryostat. Enfin, la partie supérieure représente l'électronique hors-détecteur, notamment les ROD, montées sur leurs châssis avec d'autres modules [95].

–  $A_{ADC}$  est l'amplitude du signal.

L'amplitude du signal est directement proportionnelle à l'énergie déposée :

$$A_{ADC} = \sum_{i=0}^{N_{\text{echantillon}}} a_i (s_i - p) \quad (3.2)$$



**Figure 3.4** – Amplitude en fonction du temps du courant d’ionisation après mise en forme par le filtre bipolaire. Le maximum de l’amplitude est ici ramené à 1 et le niveau de référence à 0 (i.e. l’intégrale de la forme d’ionisation est nulle).

$p$  est le piédestal, il correspond à l’énergie moyenne de la cellule en l’absence d’événements de physique. Les  $s_i$  sont les valeurs des échantillons et les  $a_i$  sont les coefficients de filtrage optimal. Le filtrage optimal est une technique qui permet de calculer l’amplitude du signal en minimisant les effets du bruit électronique et des événements d’empilement [97].

Ces coefficients sont obtenus lors des procédures d’étalonnage, tout comme le facteur  $f$  défini précédemment. Il est en effet nécessaire d’étalonner la réponse de l’électronique de chaque canal de lecture. Pour se faire, 122 cartes d’étalonnage permettent d’injecter, en sortie des électrodes, un signal de forme proche du signal d’ionisation et d’amplitude connue ; il suit ensuite le même chemin que le signal de physique dans l’électronique de lecture (plus d’informations peuvent être trouvées dans [98]). Les performances requises sur la résolution en énergie imposent un bon contrôle des non-uniformités liées à l’électronique de lecture. Ces cartes assurent la stabilité des constantes de calibration en fonction du temps avec une précision supérieure à 0.1%.

Une fois l’énergie d’une cellule reconstruite, si celle-ci est supérieure à un certain seuil (en général  $5\sigma_{\text{bruit}}$ , où  $\sigma_{\text{bruit}}$  correspond à la RMS du piédestal), le temps et le facteur de qualité  $Q$  sont également calculés :

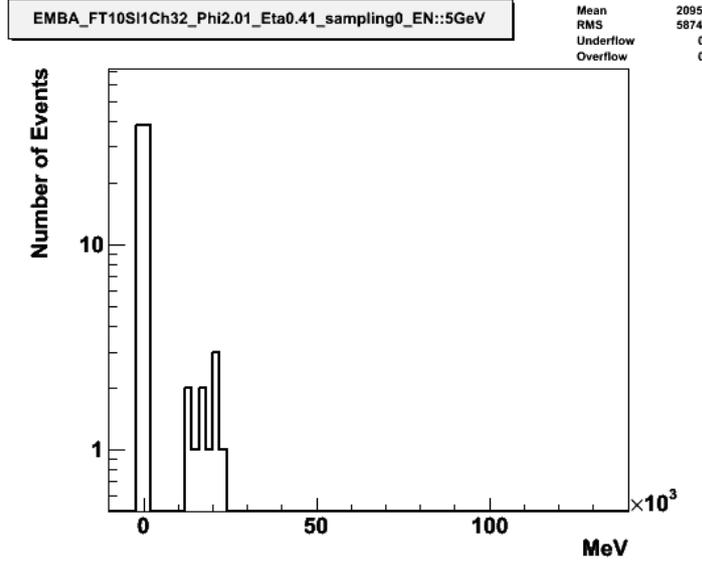
$$Q = \sum_{i=0}^{N_{\text{echantillon}}-1} [(s_i - p) - A_{ADC}(g_i - g'_i\tau)]^2 \quad (3.3)$$

Il permet de quantifier les différences entre la forme d’ionisation prédite  $g_i$  ( $g'_i$  est sa dérivée) et la forme d’ionisation mesurée. Le facteur de qualité est primordial afin de faire la différence entre un signal lié à un dépôt d’énergie physique ou lié à du bruit. Lorsque le facteur de qualité est en dessous d’un certain seuil (4000), la forme du signal est considérée comme correcte ; à l’inverse, le dépassement de ce seuil est l’indication d’un éventuel problème qui nécessite un traitement particulier de la cellule.

### 3.3 Monitoring des cellules bruyantes du calorimètre

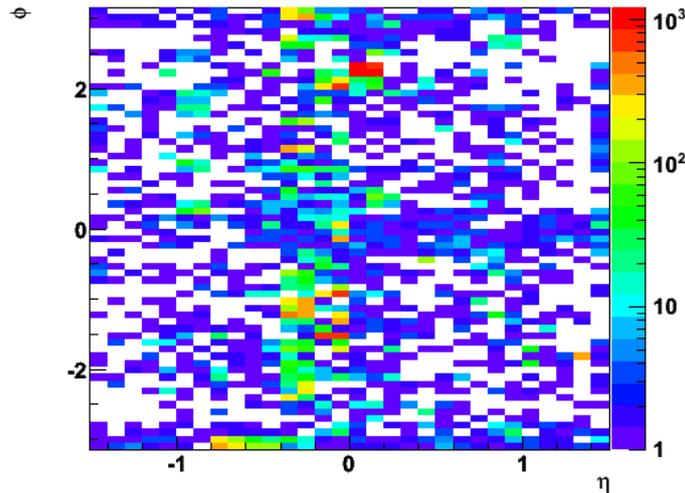
Durant les runs de collisions, chaque cellule est monitorée individuellement dans les streams *CosmicCalo* et *LArCellsEmpty* définies section 3.1. Ces streams étant dépourvues d’événements

de collisions, la distribution d'énergie attendue pour une cellule est une gaussienne centrée en 0. Certaines cellules présentent néanmoins un comportement non gaussien, avec des dépôts d'énergie très supérieurs au bruit moyen attendu comme visible sur la figure 3.5.



**Figure 3.5** – Distribution de l'énergie d'une cellule problématique du pré-échantillonneur A - stream *LArCellEmpty*. On observe une accumulation d'événements autour de 20 GeV en l'absence de collisions.

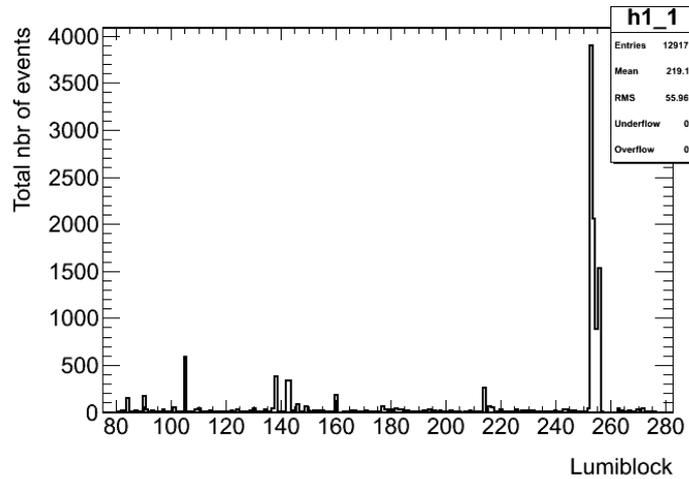
Ces cellules non-gaussiennes sont principalement situées à bas  $|\eta|$  dans le pré-échantillonneur et sont suffisamment énergétiques pour créer des zones problématiques dans la distribution en  $\eta$  et  $\phi$  des amas de cellules du calorimètre EM (section 4.1.1) avec une énergie supérieure à 10 GeV. La figure 3.6 représente la distribution spatiale de ces cellules bruyantes dans la stream *CosmicCalo* pour le run 178021. Les canaux affectés varient d'un run à l'autre; pour



**Figure 3.6** – Amas de cellules du calorimètre EM avec une énergie supérieure à 10 GeV, run 178021 - stream *CosmicCalo*. La zone située entre  $-0.5 < \eta < 0$  montre une accumulation de clusters de haute énergie en l'absence de collisions.

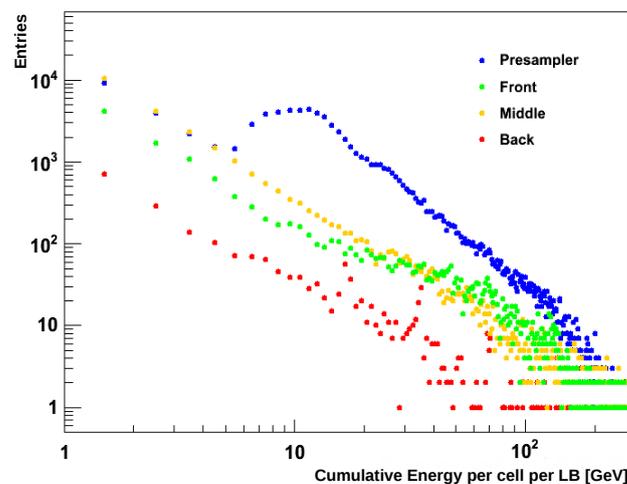
mieux comprendre la topologie de ces phénomènes, diverses études ont été menées. La figure 3.7 représente le nombre total d'événements enregistrés par lumiblock lors d'un run de collisions.

Le nombre très élevé d'événements pendant les lumiblocks 252-256 correspond à une cellule problématique. Ces phénomènes ne sont pas constants, ils ont une durée très limitée, de l'ordre de quelques lumiblocks.



**Figure 3.7** – Nombre total d'événements enregistrés par lumiblock lors d'un run de collisions - stream *LArCellsEmpty*. On observe une importante accumulation dans les LBs [252-256].

La distribution de l'énergie cumulée par cellule et par lumiblock est présentée figure 3.8. Si le pic visible à 8 GeV correspond à un effet du système de déclenchement, il est néanmoins clair que les cellules du pré-échantillonneur présentent une énergie cumulée plus haute que dans les autres couches du calorimètre. Le pré-échantillonneur est particulièrement sensible à ces phénomènes non gaussien dont l'origine est à l'intérieur du cryostat. Il n'a pas été possible d'établir de corrélations entre les régions les plus affectées ( $\eta \in [-0.3; 0]$ ) et des constituants du calorimètre. En août 2010 lors de l'augmentation de la luminosité, ce phénomène a augmenté drastiquement pour les cellules du pré-échantillonneur, ce qui a conduit à la décision de baisser la haute tension sur l'ensemble du pré-échantillonneur de 2 à 1.6 kV. La baisse de la haute tension s'est accompagnée d'une baisse de l'amplitude de ces phénomènes (voir section 3.4).



**Figure 3.8** – Energie cumulée par cellule et par lumiblock - stream *LArCellsEmpty*.

### 3.4 Traitement des cellules bruyantes du calorimètre

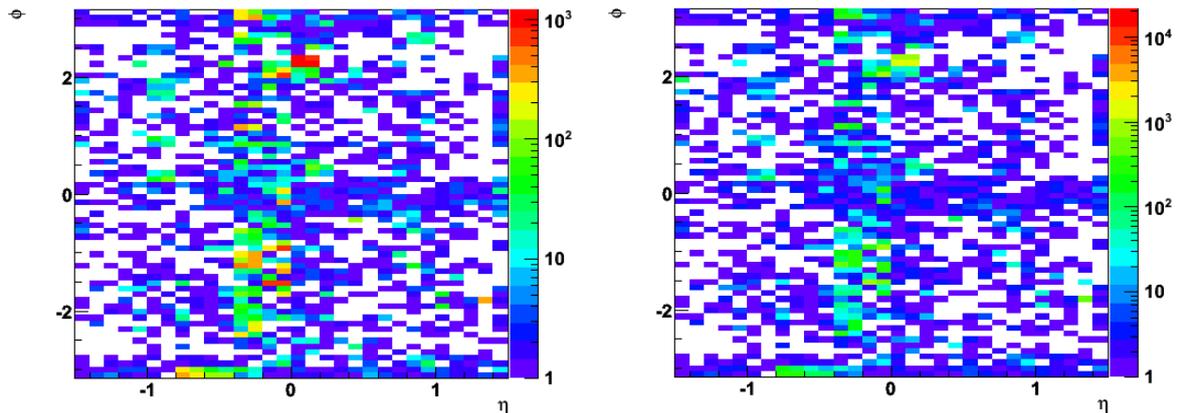
Les phénomènes décrits dans la section précédente ne sont pas constants dans le temps et les canaux affectés varient d'un run à l'autre, on parle de cellules sporadiquement bruyantes. Il est donc nécessaire d'établir pour chaque run, avant la fin de la *calibration loop* une liste complète de ces cellules problématiques. Cette liste est implémentée dans une base de données dédiée et l'ensemble des canaux y figurant seront masqués lors de la reconstruction des streams de physique.

Lorsque la cellule est masquée, son énergie est remplacée par la moyenne de l'énergie de ses huit plus proches voisins dans la même couche du calorimètre. Il existe deux façons de masquer une cellule :

- masquage permanent : la cellule est masquée durant tout le run.
- masquage sélectif : la cellule est masquée uniquement lorsque la distribution de son énergie n'est pas compatible avec un signal physique, c'est-à-dire quand le facteur de qualité  $Q$  est au-dessus d'un certain seuil fixé à 4000.

Le choix entre le masquage permanent ou sélectif se fait au regard du nombre d'événements énergétiques ( $\sim 1$  GeV) par cellule qui persisteraient après un masquage sélectif de la cellule. Un seuil conservatif a été fixé à 80 événements par cellule. Si plus de 80 événements problématiques subsistent au masquage sélectif, la cellule concernée est masquée de façon permanente.

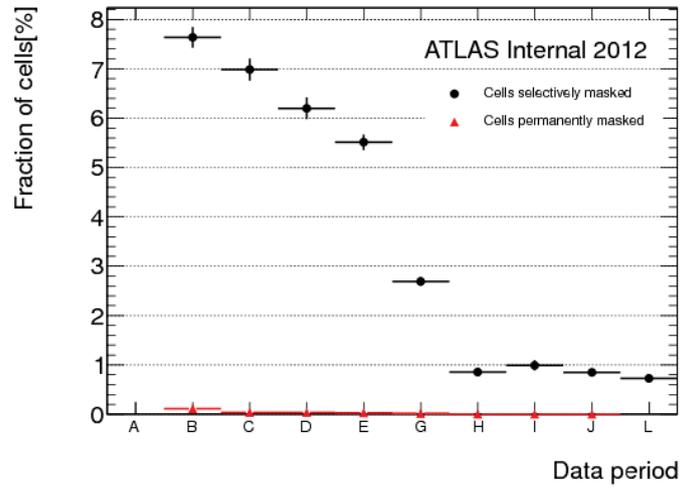
Afin de s'assurer de l'efficacité du masquage opéré pendant la reconstruction, la distribution en  $\eta$  et  $\phi$  des amas de cellules du calorimètre EM supérieurs à 10 GeV est vérifiée pour la même stream sur les données reconstruites après masquage. La figure 3.9 illustre l'effet du masquage des cellules ; à gauche la distribution déjà présentée figure 3.6, avant le masquage des cellules problématiques, à droite la même distribution après masquage. Le nombre limité d'amas valide l'efficacité du masquage des cellules problématiques.



**Figure 3.9** – Distributions en  $\eta$  et  $\phi$  des amas de cellules du calorimètre EM supérieurs à 10 GeV avant masquage des cellules (gauche) et après masquage des cellules (droite), run 205071 - stream CosmicCalo

Durant les premiers mois de l'année 2012, correspondant aux périodes de prises de données A à E, la fraction des cellules du pré-échantillonneur sélectivement masquées avoisinait 7%. Durant l'été 2012, plusieurs lignes haute-tension problématiques ont été identifiées [96]. La valeur de la tension appliquée pour ces lignes a été baissée et ceci s'est accompagné d'une

diminution du nombre de cellules problématiques. Lors de la dernière prise de données, le pourcentage de cellules du pré-échantillonneur affectées était de 1%. La figure 3.10 montre cette évolution. Le pourcentage de cellule masquées de façon permanente, quant à lui, reste inférieur à 0.2% durant toute la prise de données 2012.



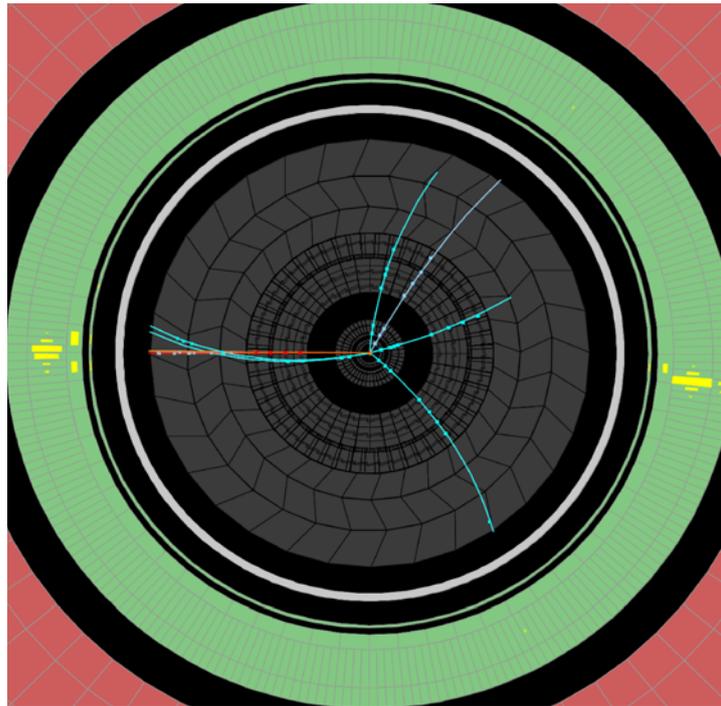
**Figure 3.10** – Pourcentage de cellules du pré-échantillonneur sélectivement masquées en fonction du temps, durant les différentes périodes de prise de données. Seuls les runs de plus  $100 \text{ pb}^{-1}$  sont ici considérés [96].



## Chapitre 4

# Les photons dans ATLAS

Une fois la qualité des données garantie, il est possible de procéder à la reconstruction et calibration des objets électromagnétiques. Les électrons et photons de haute énergie déposent la plupart de leur énergie dans le calorimètre électromagnétique. Le dépôt d'énergie a la forme caractéristique des gerbes électromagnétique, dense, étroit et contenu dans la vingtaine de  $X_0$  du calorimètre électromagnétique ; on l'appelle amas ou *cluster* électromagnétique. Dans le cas des électrons, une trace est associé à ce cluster, pour les photons la situation est plus complexe. Une large fraction des photons ( $\sim 40\%$ ) se convertit en une paire  $e^+e^-$  en traversant la matière du détecteur ; on distingue alors les photons convertis des photons non-convertis (figure 4.1).



**Figure 4.1** – *Event display d'un événement diphoton ( $m_{\gamma\gamma}=188.8$  GeV) avec un candidat converti (gauche) est un candidat non-converti (droite) [99].*

La reconstruction des candidats photons sera détaillée dans un premier temps, nous verrons ensuite comment leur énergie est calibrée à partir du Monte-Carlo mais également à partir des événements  $Z \rightarrow ee$  dans les données. Enfin, les coupures d'isolation et d'identification appliquées aux candidats photons seront décrites, elles sont essentielles pour distinguer les photons des jets identifiés comme photon.

## 4.1 Reconstruction des photons

### 4.1.1 Reconstruction des clusters électromagnétiques

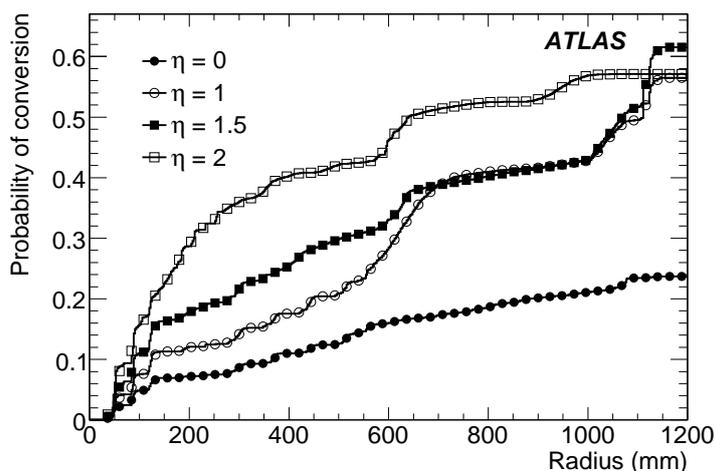
Des clusters électromagnétiques rectangulaires et de taille fixe sont reconstruits à l'aide de l'algorithme *sliding window* [100]. Dans un premier temps, la taille du rectangle considéré est fixée à  $3 \times 5$  cellules en  $\Delta\eta \times \Delta\phi$  de la seconde couche du calorimètre soit  $0.075 \times 0.125$ . Le centre du rectangle est ajusté afin de maximiser l'énergie contenue dans le rectangle ; si celle-ci est supérieure à 2.5 GeV, un pré-cluster est formé.

Afin de distinguer les photons des électrons une procédure de *track matching* est appliquée après la construction du pré-cluster. Pour être associée au pré-cluster, la trace doit être contenue dans une fenêtre en  $\Delta\eta \times \Delta\phi$  de  $0.05 \times 0.10$  autour du barycentre du pré-cluster et son impulsion doit correspondre au moins à 10% de l'énergie du pré-cluster. Si une telle trace au moins est identifiée, l'objet EM est classé comme candidat électron, si aucune trace n'est associée l'objet est classé comme candidat photon. Dans le cas de particule chargées, le champs magnétique courbe la trajectoire dans la direction  $\phi$  ; c'est pourquoi l'énergie du cluster EM est recalculée en utilisant cette fois-ci un rectangle de  $3 \times 7$  cellules en  $\Delta\eta \times \Delta\phi$  de la seconde couche du calorimètre. Dans les bouchons, l'énergie du cluster, quel que soit l'objet EM associé est recalculée en utilisant  $5 \times 5$  cellules dans la seconde couche du calorimètre.

A ce stade quasiment tous les photons non-convertis ( $\sim 85\%$ ) sont identifiés et reconstruits comme tels, et quasiment tous les photons convertis ( $\sim 94\%$ ) sont identifiés comme électrons ; ils doivent être récupérés par la suite.

### 4.1.2 Photons convertis

Pour les photons convertis, l'efficacité de reconstruction dépend de l'endroit où a lieu la conversion. Pour les conversions tardives qui ne laissent de hit que dans le TRT, il est moins probable que la trace soit associée au cluster. A l'inverse les conversions qui ont lieu avant le détecteur à pixel, sont plus souvent reconstruites comme des électrons. La probabilité de conversion des photons dépend de la distribution en  $\eta$  de la quantité de matière dans le détecteur interne présentée section 2.2.2 sur la figure 2.8. La figure 4.2 montre qu'à haut  $|\eta|$ , à la sortie du SCT, un photon a autant de chance de se convertir que de ne pas se convertir.



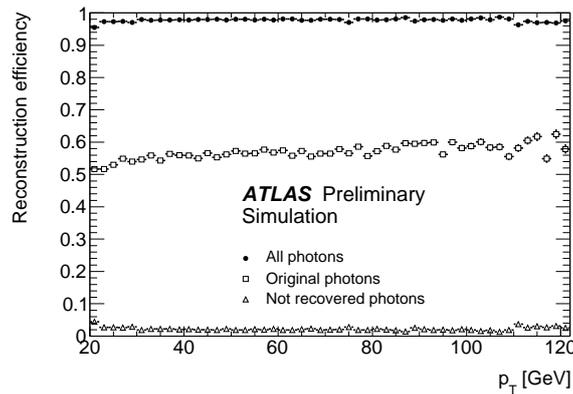
**Figure 4.2** – Probabilité pour un photon de se convertir, en fonction du rayon, pour différentes valeurs de  $\eta$  (calculée pour des photons de plus de 1 GeV) [91].

Les photons convertis sont récupérés depuis la "collection" électron grâce à une procédure

particulière. Cette procédure commence avec le matching de la trace au cluster lors de la reconstruction de l'électron. Quand une trace est associée à un électron, un algorithme vérifie également si cette trace peut être associée à un vertex de conversion. La reconstruction des vertex de conversion est faite par le détecteur interne. Les vertex de conversion peuvent avoir deux traces (de signes opposés), mais ils peuvent également n'avoir qu'une seule trace dans le cas où la conversion est très asymétrique ou si les deux traces sont trop rapprochées. Les candidats électrons pour lesquels la trace coïncide avec une trace provenant d'un vertex de conversion (simple ou double trace) sont alors traités comme des photons convertis. Au niveau de l'analyse, l'objet ainsi défini peut être considéré comme électron ou comme photon converti.

Après cette procédure de récupération, l'efficacité de reconstruction des photons convertis est estimée à 94.33% à partir du Monte-Carlo [101].

Par ailleurs, une partie des photons non-convertis ( $\sim 9\%$ ) ont été reconstruits comme des électrons. Si la trace d'un candidat électron n'est pas associée à un vertex de conversion et si elle est telle que  $p_T < 2$  GeV ou  $E/p > 10$ , le candidat est traité comme un photon non-converti. L'efficacité de reconstruction des photons non-convertis estimée à partir du Monte-Carlo après cette procédure est 99.83%. La figure 4.3 montre l'efficacité de reconstruction avant et après la procédure de récupération en fonction de  $p_T$ . La fraction de photons non reconstruits est également visible.



**Figure 4.3** – Efficacité de reconstruction des photons, avant et après la procédure de récupération, en fonction de  $p_T$  [101].

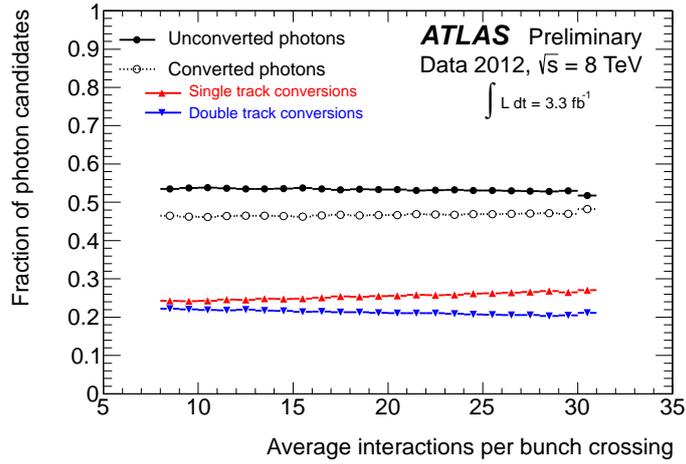
La stabilité de la fraction de photons reconstruits, convertis ( $\sim 46\%$ ) et non-convertis ( $\sim 54\%$ ), en fonction de l'empilement est présentée figure 4.4. La fraction de photon convertis et non convertis varie de 1% entre les différentes conditions d'empilement.

## 4.2 Calibration des clusters électromagnétiques

Une fois l'objet photon reconstruit, son énergie et sa position peuvent être calibrées en conséquence. Les photons convertis et les photons non-convertis sont calibrés de façon indépendante. Il y a deux étapes dans la calibration des photons. La première est basée sur la simulation Monte-Carlo, la deuxième est une calibration *in situ* basée sur les événements  $Z \rightarrow e^+e^-$ .

### 4.2.1 Calibration Monte-Carlo

L'énergie des clusters EM est corrigée à partir de la simulation pour prendre en compte tous les dépôts d'énergie en dehors du cluster. Mais pour se faire, il est d'abord nécessaire de corriger la position du cluster qui est initialement mesurée comme le barycentre (pondéré par



**Figure 4.4** – Fraction de photon convertis et non-convertis reconstruits en fonction du nombre moyen d’interactions par croisement de paquets.

l’énergie) des cellules considérées dans chaque couche du calorimètre. Le fait que la gerbe ne soit pas complètement contenue dans la fenêtre en  $\eta$  et la taille finie des cellules amènent une dépendance de la position mesurée en fonction du point d’impact de la particule dans la cellule. Il est donc nécessaire de corriger la position en  $\eta$  pour ce biais. Par ailleurs, en raison de la géométrie en accordéon du calorimètre EM, la quantité d’absorbeur rencontrée par la particule incidente varie en fonction de  $\phi$ . Cela introduit une modulation en  $\phi$  de l’énergie déposée par le photon qu’il est également nécessaire de corriger.

Les positions mesurées (après correction) à partir des deux premières couches du calorimètre sont alors combinées pour définir le point d’impact de la gerbe dans le calorimètre.

L’énergie du cluster est ensuite calculée en sommant quatre contributions :

- l’énergie déposée dans le pré-échantillonneur et en amont du calorimètre.
- l’énergie déposée dans l’accordéon (c’est-à-dire front, middle et back).
- l’énergie perdue dans le calorimètre EM (fuites latérales).
- l’énergie perdue hors du calorimètre EM (fuite longitudinales en arrière).

L’énergie déposée en amont du calorimètre et celle déposée dans l’accordéon sont corrigées par des facteurs de calibration dérivés de la simulation, afin de prendre en compte la différence entre énergie reconstruite et énergie réellement déposée. Les pertes d’énergie du cluster dans le calorimètre sont typiquement de l’ordre de 5% et les pertes en arrière du calorimètre dépendent très fortement de la profondeur de la gerbe EM, elles vont de 0.1 à 2%. Tous ces coefficients sont estimés à partir de la simulation et optimisés indépendamment pour les photons convertis et les photons non-convertis.

#### 4.2.2 Calibration in-situ

Afin d’améliorer la précision obtenue avec la calibration Monte-Carlo, les événements  $Z \rightarrow e^+e^-$  et la masse bien connue du boson  $Z$  sont utilisés pour extraire une correction et inter-calibrer les différentes régions du détecteur en  $\eta$ . La masse reconstruite des candidats  $Z \rightarrow e^+e^-$  est calculée telle que :

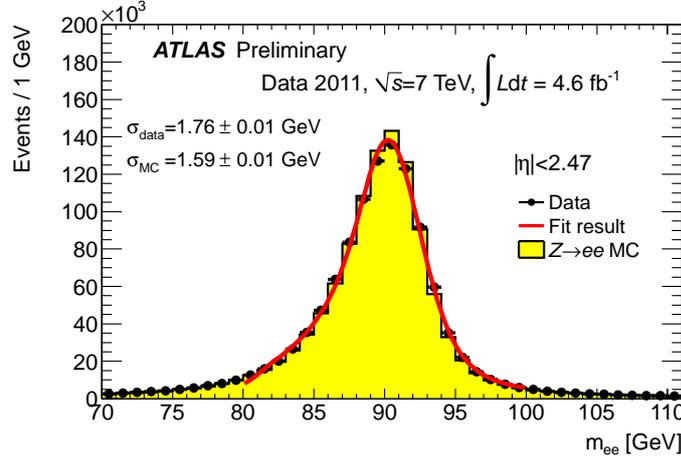
$$m_{12}^{reco} = \sqrt{2E_1^{reco}E_2^{reco}(1 - \cos\theta_{12})} \quad (4.1)$$

$E_1^{reco}$ ,  $E_2^{reco}$  sont les énergies reconstruites des électrons et  $\theta_{12}$  est l’angle les séparant, mesuré par le trajectographe.

L'angle  $\theta_{12}$  est supposé parfaitement connu mais un facteur de correction de l'échelle d'énergie des électrons,  $\alpha_i$ , est introduit pour chaque région  $i$  du calorimètre :

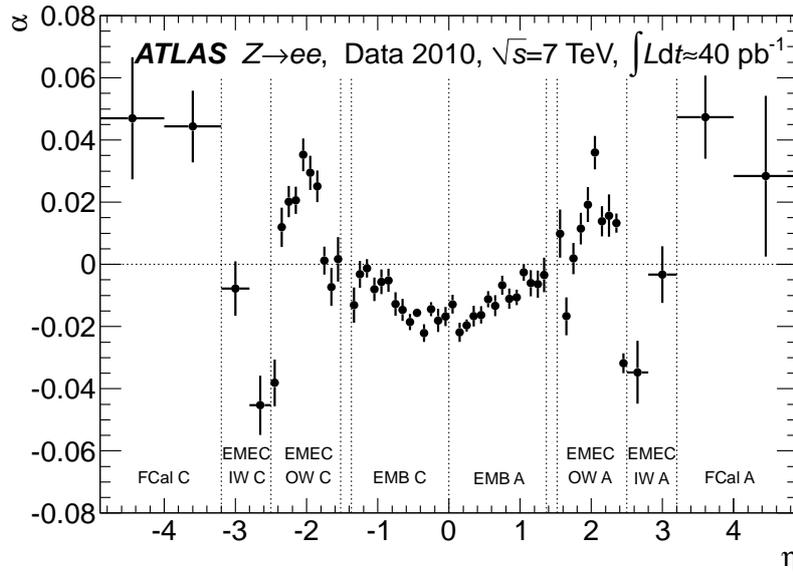
$$E_i^{reco} = E_i^{true}(1 + \alpha_i) \quad (4.2)$$

Ce coefficient  $\alpha_i$  est extrait d'un fit sur la masse reconstruite du boson  $Z$  comme illustré figure 4.5. Les erreurs systématiques principales considérées dans la mesure des  $\alpha_i$  sont l'incertitude sur le fit, la dépendance en énergie, résultante des effets du matériel (estimée à partir du Monte-Carlo) et l'incertitude sur l'échelle d'énergie du pré-échantillonneur. La figure 4.6



**Figure 4.5** – Masse reconstruite des candidats  $Z \rightarrow e^+e^-$  avec les deux électrons dans la région tonneau. Les données sont comparées à la simulation Monte-Carlo. Données et simulation sont fittés par une Breit-Wigner convoluée à une fonction Crystal-Ball [102].

montre les valeurs  $\alpha_i$  obtenues à partir des données 2010 pour une luminosité de  $40 \text{ pb}^{-1}$ .  $\alpha$  est compris entre  $\pm 2\%$  dans la région tonneau et entre  $\pm 5\%$  dans les régions bouchons [103].



**Figure 4.6** – Facteur de correction de l'échelle d'énergie des électrons en fonction de  $\eta$  [103].

Ces corrections sont calculées pour les électrons, et électrons et photons n'intéragissent pas de la même façon avec la matière. En raison de cet effet de matière, appliquer la même

correction que celle appliquée à l'échelle d'énergie des électrons "sur-corrige" l'échelle d'énergie des photons. La différence sur les  $\alpha_i$  entre photons et électrons est estimée à partir de la simulation et est traitée comme une erreur systématique supplémentaire. Pour les photons convertis la différence est très proche de 0, pour les photons non-convertis elle n'excède pas 2%.

### 4.3 Isolation

On a évoqué section 1.3.5 la nécessité d'une coupure sur l'énergie d'isolation des photons pour les différencier des jets qui sont eux entourés d'une intense activité hadronique, comme illustré figure 4.7. L'énergie d'isolation peut être mesurée par le calorimètre, l'isolation calorimétrique ou par le détecteur interne, l'isolation des traces.

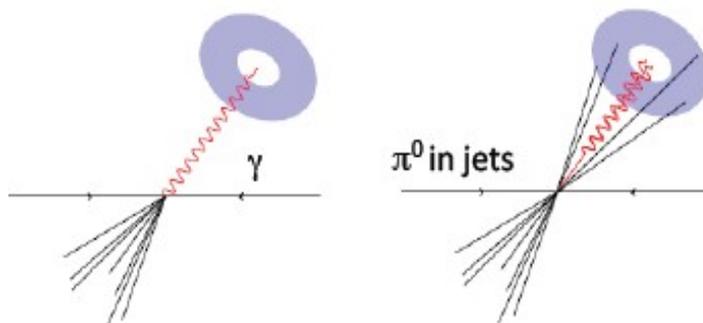


Figure 4.7 – Illustration du principe d'isolation des photons.

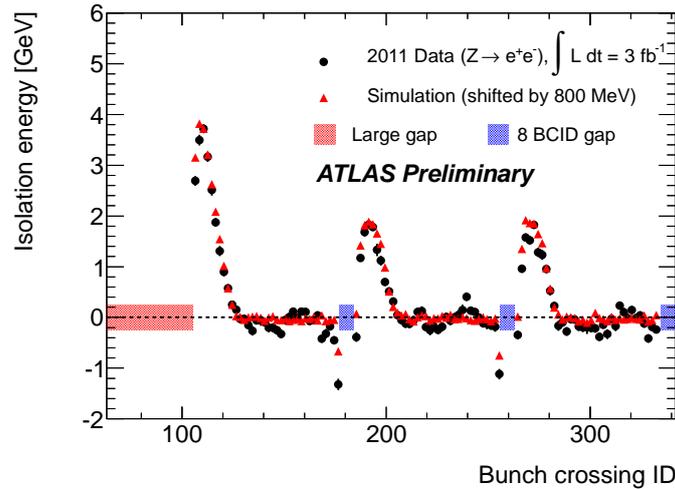
#### 4.3.1 Isolation calorimétrique

L'énergie d'isolation calorimétrique **EtCone** correspond à la somme des énergies déposées dans les cellules du calorimètre contenues dans un cône de rayon  $R=0.4$  autour du photon. Afin d'exclure l'énergie du photon, on soustrait de cette somme l'énergie contenue dans une fenêtre  $5 \times 7$  en  $\Delta\eta \times \Delta\phi$  centrée sur le photon. Le rayon de Molière caractérise le développement latéral de la gerbe électromagnétique, il est défini comme le rayon d'un cylindre contenant en moyenne 90% de l'énergie déposée par la gerbe. Pour le calorimètre électromagnétique d'ATLAS, ce rayon est d'approximativement 4.8 cm soit 1.3 cellule dans la partie tonneau du calorimètre LAr. 90% de l'énergie du photon devrait donc être contenue dans une fenêtre de  $3 \times 3$  cellules en  $\Delta\eta \times \Delta\phi$ , et environ 95% dans une fenêtre de  $5 \times 7$  cellules en  $\Delta\eta \times \Delta\phi$ . Il y a ainsi une fuite estimée à quelques % de l'énergie du photon lui-même dans la variable **EtCone** qu'il est nécessaire de corriger afin d'être indépendant de l'énergie initiale du photon.

L'énergie liée aux événements d'empilement ainsi que l'énergie liée aux événements sous-jacent viennent également s'ajouter à la variable d'isolation calorimétrique. La correction pour ces contributions est basée sur une estimation de la densité d'énergie ambiante événement par événement. La reconstruction des jets est faite via un algorithme  $k_T$  avec  $R=0.5$  [104] ; il tourne sur des clusters topologiques en trois dimensions avec suppression du bruit, les *topocluster*. Les topoclusters, contrairement aux clusters obtenus avec l'algorithme sliding window (section 4.1.1) ne sont pas de taille fixe, l'algorithme commence avec une cellule dont l'énergie est

supérieure à  $4\sigma_{\text{bruit}}$  et rajoute itérativement au cluster des cellules dont l'énergie est supérieure à  $2\sigma_{\text{bruit}}$  (plus de détails peuvent être trouvés dans [105] et [100]). Durant la reconstruction une aire est associée à chaque jet en se basant sur la décomposition de Voronoi<sup>1</sup>. La densité d'énergie pour chaque jet est alors égale à l'énergie du jet considéré divisée par son aire. L'estimation de l'énergie ambiante est donnée par la médiane de la densité d'énergie des jets.

Cette correction, basée sur les topoclusters avec suppression du bruit ne peut pas s'appliquer à l'ensemble des cellules du calorimètre qui n'incluent, quant à elle aucune suppression du bruit. Cette correction ne peut donc pas s'appliquer correctement à la variable  $\text{EtCone}$ , une dépendance en fonction de l'empilement subsiste après les corrections comme illustré figure 4.8. Cette figure représente l'évolution de l'énergie d'isolation, en fonction du numéro de croisement de paquets ou BCID (pour *Bunch Crossing ID*).



**Figure 4.8** – Isolation calorimétrique  $\text{EtCone}$  en fonction de la position du croisement de faisceau dans le train de paquets [106].

Pour remédier à cette dépendance, une nouvelle variable a été imaginée  $\text{topoEtCone}$ . Dans cette variable, ne sont considérées que les cellules appartenant à un topocluster (représentées en orange figure 4.9). L'énergie d'isolation correspond alors à la somme des énergies transverses des clusters (avec une énergie positive) dont le barycentre est situé dans le cône d'isolation. Cette variable est également corrigée pour les fuites latérales en supposant une correction linéaire en  $p_T$ , et la correction pour l'empilement et les événements sous-jacent est obtenue de la même façon que pour la variable  $\text{EtCone}$ . La figure 4.10 montre la stabilité de cette variable en fonction de l'empilement.

La figure 4.11 [107] nous renseigne sur l'équivalence des coupures entre les deux variables d'isolation  $\text{EtCone}$  et  $\text{topoEtCone}$  pour une même efficacité d'isolation.

### 4.3.2 Isolation des traces

On ne considère que les traces dont l'impulsion transverse  $p_T$  est supérieure à 1 GeV afin de minimiser l'effet lié à l'empilement et aux événements sous-jacent. Dans ce cas, le photon est considéré comme isolé si la somme des  $p_T$  des traces comprises dans un cône de rayon  $R = 0.2$  entourant le photon,  $E_T^{\text{iso,trace}}$  est inférieure à un certain seuil fixé à quelques GeV. Seules les traces associées au vertex primaire sont considérées (la sélection du vertex primaire

1. Soit  $m$  un point d'un ensemble de points  $V$  : la cellule de Voronoi associée à  $m$  est l'ensemble des points de l'espace tels que leur distance par rapport à  $m$  est inférieure à leur distance par rapport à n'importe quel autre point de  $V$ .

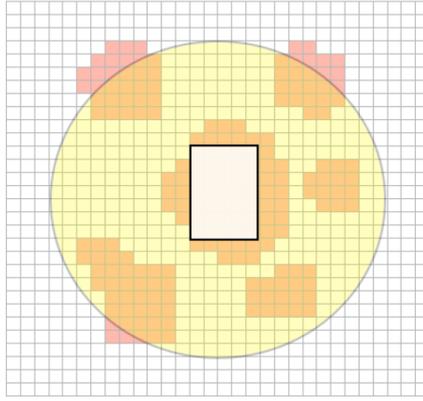


Figure 4.9 – Représentation schématique du calcul de la variable d’isolation  $topoEtCone$ .

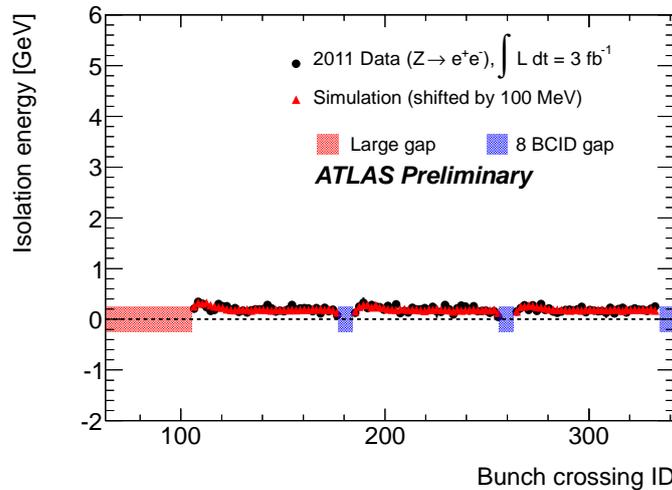


Figure 4.10 – Isolation calorimétrique avec  $topoclusters\ topoEtCone$  en fonction de la position du croisement de faisceau dans le train de paquets [106].

est présentée 6.1.1) et toutes les traces associées à des photons convertis sont exclues de cette somme. La figure 4.12 montre la distribution de l’énergie d’isolation des traces pour les photons et pour le bruit de fond après avoir appliqué les coupures d’identification.

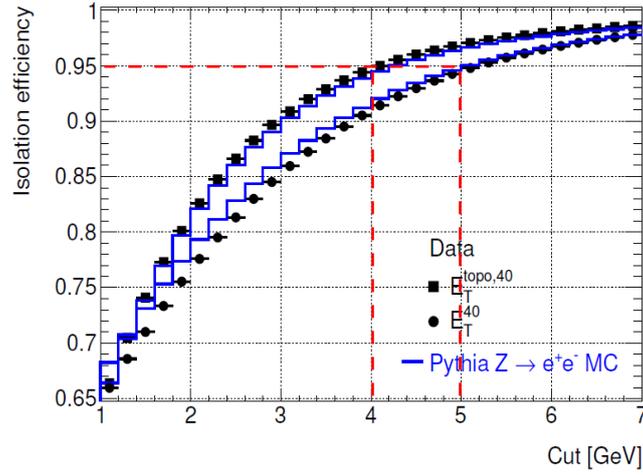
## 4.4 Identification

L’identification des photons directs consiste en un jeu de coupures sur les variables décrivant la forme de la gerbe électromagnétique (EM). Elle permet de rejeter les hadrons, principalement des  $\pi^0$  qui se désintègrent en deux photons 4.13. Deux jeux de coupures sont ici détaillés, la sélection LOOSE et la sélection TIGHT.

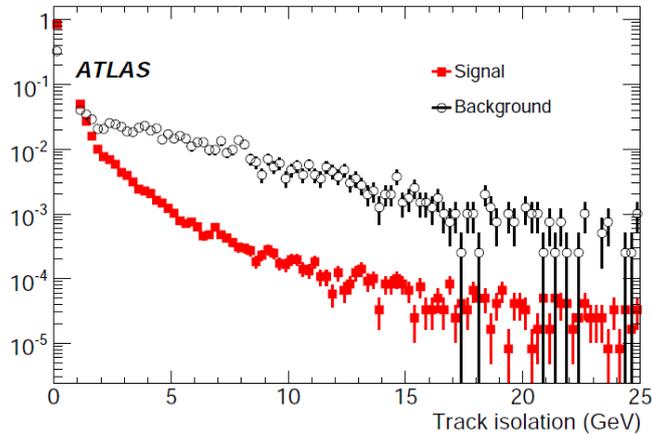
### 4.4.1 Les variables discriminantes

Les variables discriminantes sont divisées en trois catégories. Les variables utilisant la première couche du calorimètre, celles utilisant le seconde couche du calorimètre, et enfin celles utilisant le calorimètre hadronique.

La fine granularité en  $\eta$  de la première couche du calorimètre EM permet de distinguer un unique photon d’une paire de photons collinéaires issue de la désintégration d’un  $\pi^0$ .



**Figure 4.11** – Efficacité d’isolation en fonction de la coupure appliquée sur la variable d’isolation pour les deux variables  $EtCone$  (rond) et  $topoEtCone$  (carré) [107].



**Figure 4.12** – Distribution de l’énergie d’isolation des traces des photons et du bruit de fond après les coupures d’identification [91].

- $f_{side}$  mesure la fraction d’énergie déposée en dehors du coeur de la gerbe EM :

$$f_{side} = \frac{E_{7 \times 1} - E_{3 \times 1}}{E_{3 \times 1}} \quad (4.3)$$

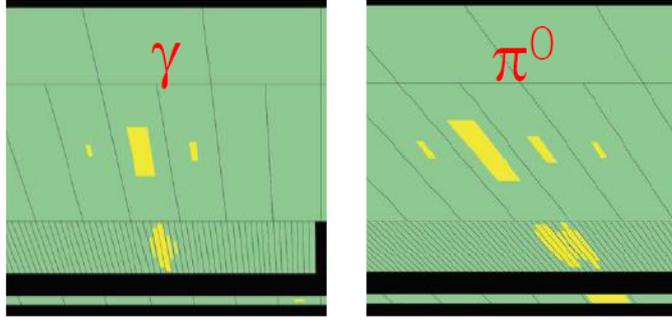
où  $E_{x \times y}$  est l’énergie mesurée dans  $x \times y$  cellules en  $\Delta\eta \times \Delta\phi$  de la première couche du calorimètre EM autour de la cellule la plus énergétique.

- $w_{s,3}$  mesure la largeur de la gerbe EM en  $\eta$  en utilisant les trois cellules entourant la cellule la plus énergétique :

$$w_{s,3} = \sqrt{\frac{\sum E_i (i - i_{max})^2}{\sum E_i}} \quad (4.4)$$

où  $i$  est l’indice de la cellule est  $i_{max}$  désigne la cellule la plus énergétique.

- $w_{tot}$  mesure la largeur de la gerbe EM dans une fenêtre de  $24 \times 2$  cellules en  $\Delta\eta \times \Delta\phi$ , elle est calculée de la même façon que  $w_{s,3}$ .



**Figure 4.13** – Event display d’un photon direct et d’un  $\pi^0$  se désintégrant en deux photons. On observe deux dépôts distincts dans le front pour la désintégration du  $\pi^0$ .

- $\Delta E$  quantifie la présence d’un second dépôt d’énergie :

$$\Delta E = E_{max2} - E_{min} \quad (4.5)$$

où  $E_{max2}$  est la seconde cellule la plus énergétique et  $E_{min}$  est la cellule la moins énergétique entre les deux maxima  $E_{max1}$  et  $E_{max2}$ .

- $E_{ratio}$  mesure la différence relative entre les deux cellules les plus énergétiques :

$$E_{ratio} = \frac{E_{max1} - E_{max2}}{E_{max1} + E_{max2}} \quad (4.6)$$

Cette variable est particulièrement discriminante entre les photons et les  $\pi^0$  qui se désintègrent en deux photons.

La gerbe EM dépose l’essentiel de son énergie dans la seconde couche du calorimètre EM. Les gerbes EM sont moins dispersées que les gerbes hadroniques, l’étalement latéral de la gerbe nous apporte donc une information discriminante.

- $R_\eta$  mesure l’étalement latéral en  $\eta$  de la gerbe EM à l’extérieur du cluster :

$$R_\eta = \frac{E_{3 \times 7}}{E_{7 \times 7}} \quad (4.7)$$

où  $E_{x \times y}$  est l’énergie contenue dans  $x \times y$  cellules en  $\Delta\eta \times \Delta\phi$  centrées autour du cluster.

- $R_\phi$  mesure l’étalement latéral en  $\phi$  de la gerbe EM à l’intérieur et à l’extérieur du cluster :

$$R_\phi = \frac{E_{3 \times 3}}{E_{3 \times 7}} \quad (4.8)$$

où  $E_{x \times y}$  est l’énergie contenue dans  $x \times y$  cellules en  $\Delta\eta \times \Delta\phi$  centrées autour du cluster. Cette variable est moins discriminante pour les photons convertis, leur trajectoire est courbée en  $\phi$  par le champ magnétique.

- $w_{\eta 2}$  mesure la largeur latérale en  $\eta$  dans une fenêtre de  $3 \times 5$  cellules en  $\Delta\eta \times \Delta\phi$  :

$$w_{\eta 2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left( \frac{\sum E_i \eta_i}{\sum E_i} \right)^2} \quad (4.9)$$

où  $i$  désigne l’indice de la cellule, il va de 0 à 14.

Contrairement aux gerbes EM qui sont largement contenues dans le calorimètre EM, les gerbes hadroniques laissent un dépôt d'énergie significatif dans le calorimètre hadronique.

- $R_{had1}$  correspond au rapport de l'énergie transverse déposée par la gerbe dans la première couche du calorimètre hadronique  $E_T^{had1}$  sur l'énergie transverse du cluster  $E_T$  :

$$R_{had1} = \frac{E_T^{had1}}{E_T} \quad (4.10)$$

$R_{had1}$  est utilisée pour  $|\eta| < 0.8$  et  $|\eta| > 1.37$ . Ailleurs, c'est la variable  $R_{had}$  qui est utilisée.

- $R_{had}$  correspond au rapport de l'énergie transverse totale déposée par la gerbe dans le calorimètre hadronique  $E_T^{had}$  sur l'énergie transverse du cluster  $E_T$

$$R_{had} = \frac{E_T^{had}}{E_T} \quad (4.11)$$

Les distributions de ces variables sont présentées, et comparées au bruit de fond, pour les photon non-convertis figure 4.14 et pour les photons convertis figure 4.15.

#### 4.4.2 Les sélections loose et tight

La sélection LOOSE n'utilise que les variables basées sur la seconde couche du calorimètre EM ainsi que la variable  $R_{had}$  (où  $R_{had1}$  selon la région en  $\eta$ ). Les coupures appliquées sur les variables ont été optimisées afin d'obtenir un facteur de réjection proche de 1000 tout en gardant une efficacité de signal d'au moins 97% pour les photons dont l'énergie transverse est supérieure à 20 GeV dans la simulation [101]. Les coupures sont identiques pour les photons convertis et non-convertis. Une sélection LOOSE-trigger est aussi utilisée pour les trigger photons, avec toutefois des coupures plus lâches.

La sélection TIGHT utilise les neuf variables décrites précédemment. Les coupures ont été choisies pour maintenir une efficacité de 85% pour les photons dont l'énergie transverse est supérieure à 40 GeV, tout en ayant un facteur de réjection des jets de l'ordre de 5000 [108]. Dans ce cas les coupures sont optimisées indépendamment pour les photons convertis et non-convertis.

#### 4.4.3 Efficacité d'identification

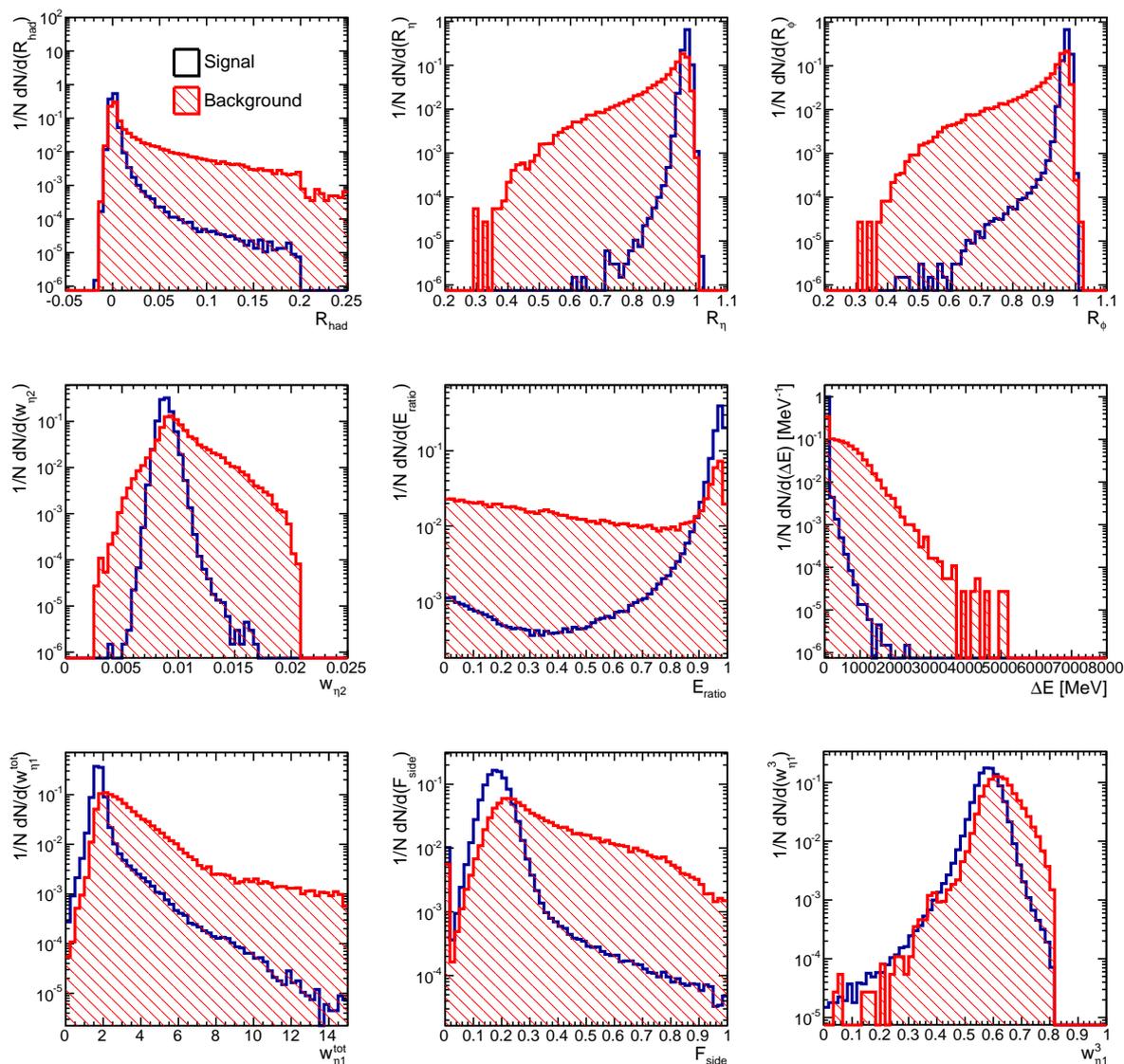
L'efficacité d'identification est définie comme la fraction de photons isolés reconstruits qui passent la coupure d'identification TIGHT. Pour un interval  $k$  en  $\eta$ , on la calcule avec :

$$\epsilon^{reco,k}(E_{T,true}^\gamma) = \frac{dN^\gamma(\eta_{k,1} \leq |\eta_{reco}^\gamma| \leq \eta_{k,2}, E_{T,reco}^{iso} < E_{T,reco|cut}^{iso}, tight - ID)/dE_{T,true}^\gamma}{dN^\gamma(\eta_{k,1} \leq |\eta_{true}^\gamma| \leq \eta_{k,2}, E_{T,true}^{iso} < E_{T,true|cut}^{iso})/dE_{T,true}^\gamma} \quad (4.12)$$

Elle est calculée à partir de la simulation. Les différences entre données et Monte-Carlo dans les distributions des variables discriminantes sont prises en compte par des facteurs correctifs. L'efficacité est calculée à partir des distributions corrigées.

Outre l'estimation à partir de la simulation Monte-Carlo, l'efficacité d'identification peut également être dérivée des données. Trois méthodes différentes sont utilisées [108] :

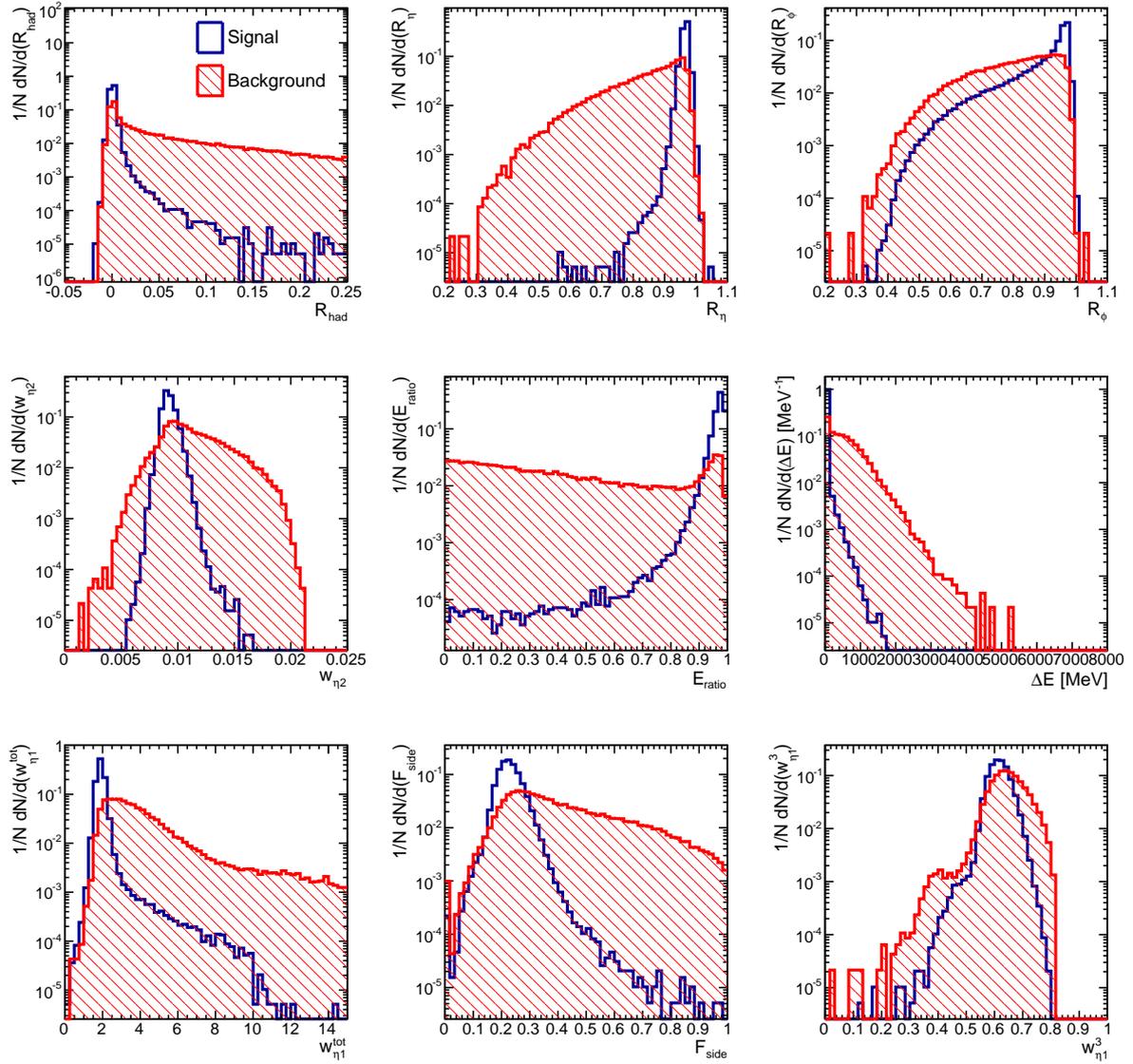
- La première méthode est basée sur la selection de photons directs isolés issus de la désintégration radiative du boson  $Z$  ( $Z \rightarrow l\gamma$ ).



**Figure 4.14** – Distributions normalisées des variables calorimétriques discriminantes dans la région  $0 < |\eta| < 0.6$  pour  $E_T > 20$  GeV pour des photons et des jets reconstruits comme photons non-convertis avant toute sélection (simulation). [101].

- La seconde méthode est basée sur l’extrapolation des électrons obtenus des événements  $Z \rightarrow e^+e^-$ . Si cette extrapolation est quasiment directe dans le cas des photons convertis, elle est plus complexe dans le cas des photons non convertis.
- La dernière méthode, aussi appelée *la méthode de la matrice* utilise l’isolation des traces pour différencier les photons du bruit de fond.

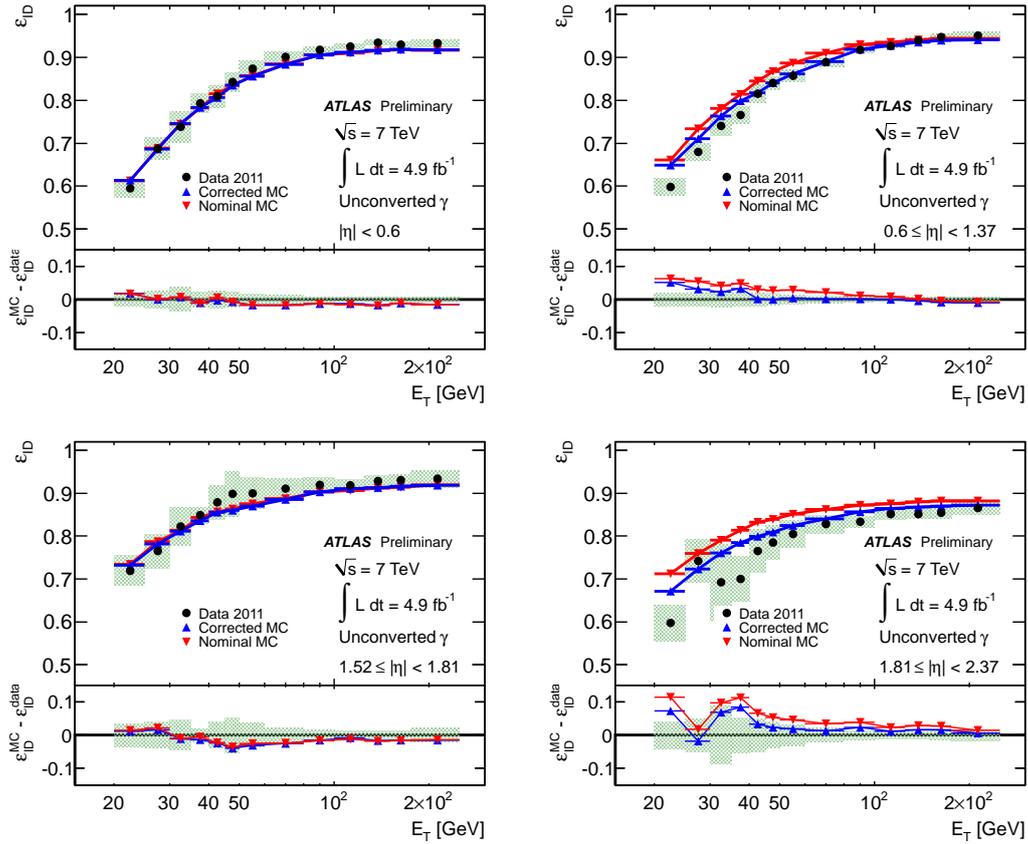
Les trois mesures donnent des résultats compatibles entre eux pour les données 2011 et compatibles avec l’estimation issue de la simulation à  $\pm 5\%$  (figures 4.14 et 4.15). L’efficacité d’identification est typiquement de 85% pour  $E_T > 40$  GeV et le facteur de réjection des jets pour les critères d’identification TIGHT et d’isolation combinés est de l’ordre de 10000. Le critère d’isolation considéré pour la mesure de ces efficacités est  $\text{EtCone} < 5$  GeV. Trois critères d’isolation différents sont utilisés dans les analyses présentées,  $\text{EtCone} < 5$  GeV qui est équivalent à  $\text{topoEtCone} < 4$  GeV (voir figure 4.11) et  $E_T^{\text{iso,trace}} < 2.6$  GeV +  $\text{topoEtCone} < 6$  GeV. La différence sur l’efficacité d’identification entre ces trois critères n’excède pas 2% [108].



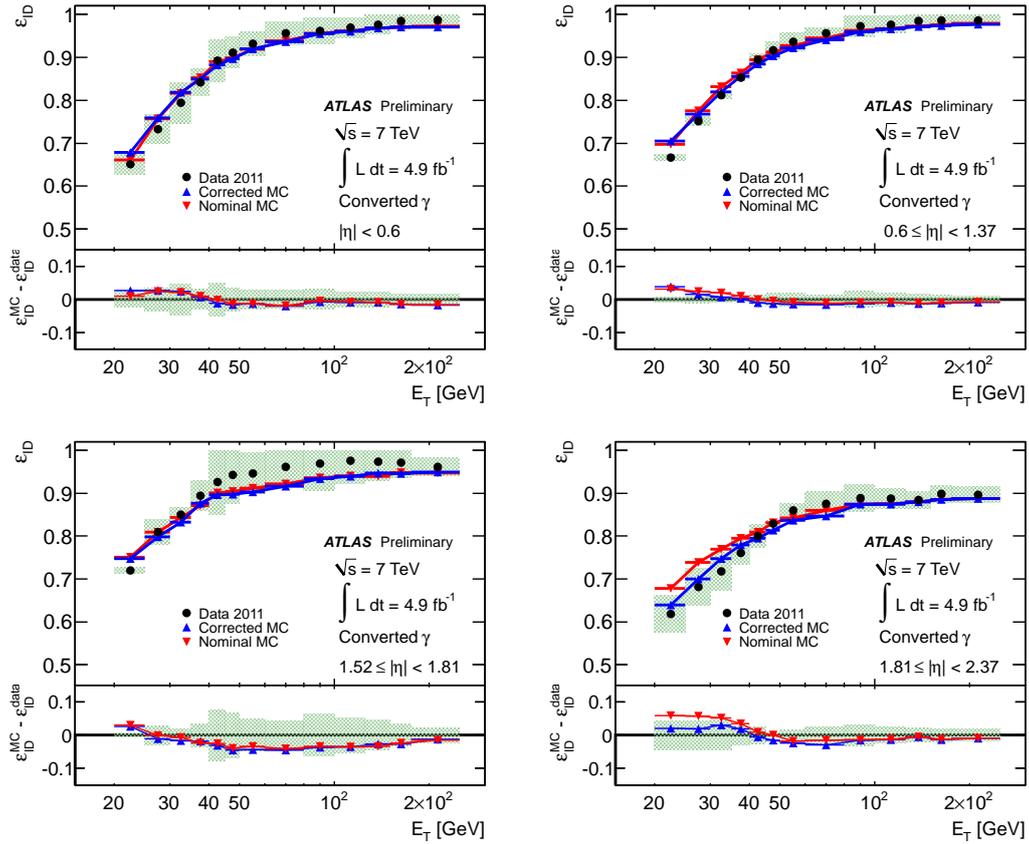
**Figure 4.15** – Distributions normalisées des variables calorimétriques discriminantes dans la région  $0 < |\eta| < 0.6$  pour  $E_T > 20$  GeV pour des photons et des jets reconstruits comme photons convertis avant toute sélection (simulation). [101].

La stabilité des efficacités d'identification pour les photons non-convertis et convertis en fonction de l'empilement est illustrée figures 4.18 et 4.19. Les coupures sur les variables d'identification ont été optimisées pour les conditions d'empilement.

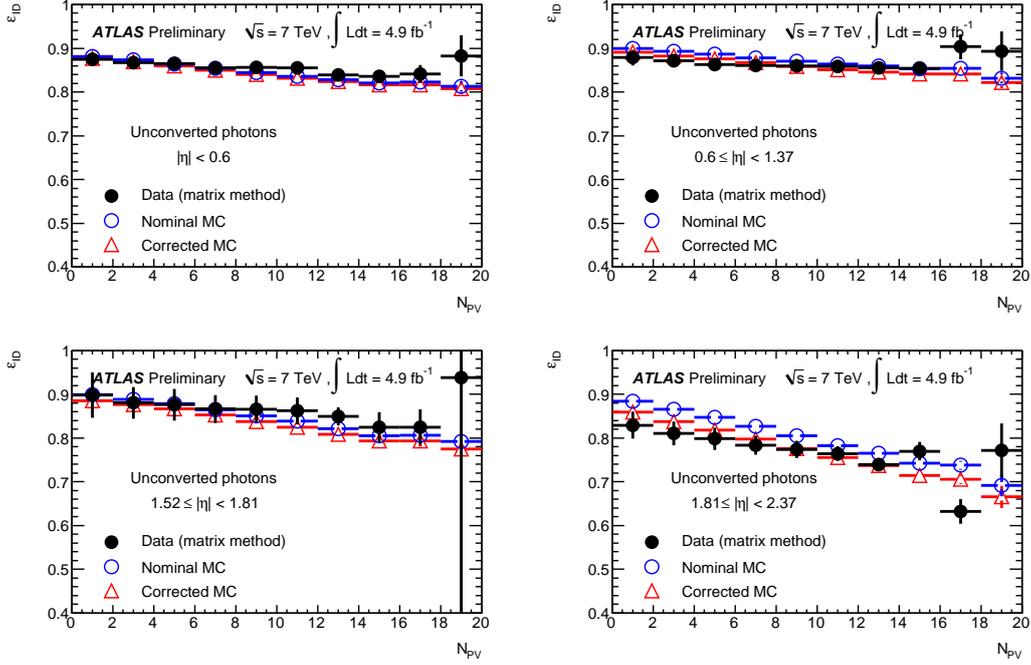
Pour la recherche du boson de Higgs dans le canal diphoton avec les données 2011, l'identification des photons est basée sur un Neural Network [109] qui prend comme variables d'entrée les mêmes que celles utilisées pour le critère TIGHT. L'efficacité d'identification varie entre 85 et 95% pour  $E_T > 40$  GeV [110].



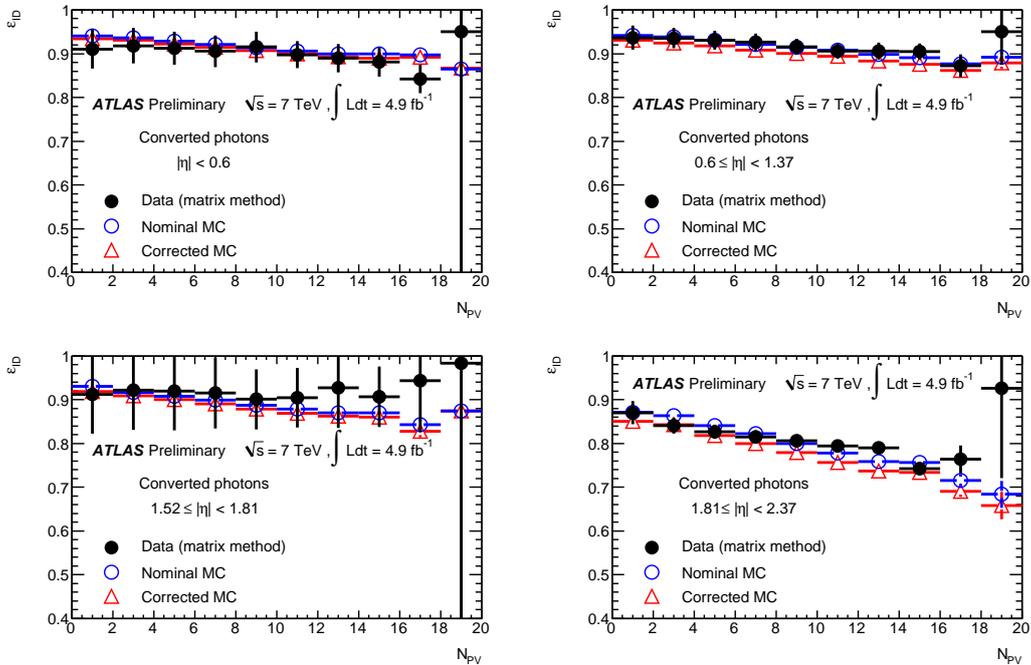
**Figure 4.16** – Comparaison entre le Monte-Carlo et la combinaison des trois méthodes basées sur les données pour extraire l'efficacité d'identification TIGHT pour des photons non-convertis dans la région  $15 < E_T < 300$  GeV. Les barres d'erreurs correspondent à la somme quadratique des erreurs systématiques et statistiques de chaque méthode [108].



**Figure 4.17** – Comparaison entre le Monte-Carlo et la combinaison des trois méthodes basées sur les données pour extraire l'efficacité d'identification TIGHT pour des photons convertis dans la région  $15 < E_T < 300$  GeV. Les barres d'erreurs correspondent à la somme quadratique des erreurs systématiques et statistiques de chaque méthode [108].



**Figure 4.18** – Efficacité d’identification TIGHT en fonction du nombre de vertex primaires pour les photons non-convertis après le critère d’isolation. Pour la méthode basée sur les données les barres d’erreurs correspondent à la somme quadratique des erreurs systematiques et statistiques [108].



**Figure 4.19** – Efficacité d’identification TIGHT en fonction du nombre de vertex primaires pour les photons convertis après le critère d’isolation. Pour la méthode basée sur les données les barres d’erreurs correspondent à la somme quadratique des erreurs systematiques et statistiques [108].

## Chapitre 5

# Mesure de sections efficaces différentielles de production de paires de photons

L'étude de la production de paires de photons (figure 5.1) apporte une compréhension approfondie du bruit de fond dans le canal  $H \rightarrow \gamma\gamma$  et elle permet de tester les prédictions de QCD perturbative afin de raffiner les générateurs Monte-Carlo dans les zones où les données sont jusqu'à présent mal comprises.

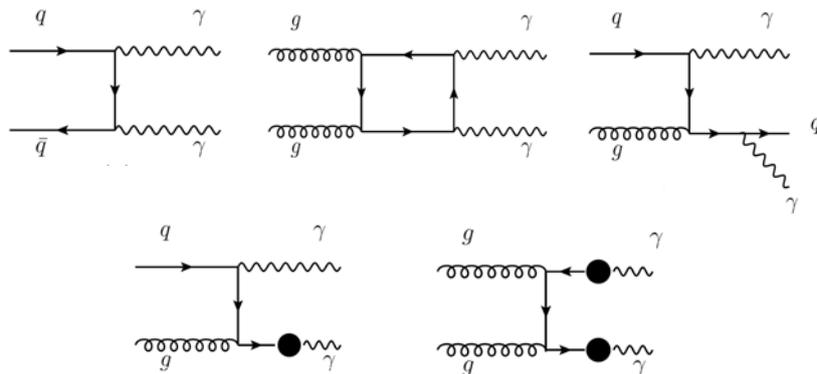


Figure 5.1 – Production de paires de photons au LHC

La section efficace différentielle de production de paires de photons pour l'ensemble des données 2011 à  $\sqrt{s} = 7$  TeV est ici mesurée en fonction de quatre variables cinématiques :

- La masse invariante de la paire de photons,  $m_{\gamma\gamma}$ . Elle est de tout premier intérêt pour la recherche du boson de Higgs dans le canal diphoton mais aussi pour la recherche de nouvelle physique.
- L'impulsion transverse de la paire de photons,  $p_{T,\gamma\gamma}$  et l'angle azimuthal entre les deux photons,  $\Delta\phi_{\gamma\gamma}$ . Toutes deux sont particulièrement sensibles à l'émission de gluons mous et apportent donc des renseignements sur la validité des prédictions théoriques.
- Le cosinus de l'angle polaire entre les deux photons dans le référentiel de Collins-Soper [111],  $\cos\theta^*$ . Cette variable est également très sensible à la contribution de la fragmentation, et elle peut être utilisée plus particulièrement pour l'étude de spin des résonances diphoton.

La sélection TIGHT des photons ainsi que les coupures d'isolation décrites dans le chapitre précédent réduisent significativement le bruit de fond issu de la mauvaise identification des jets en photons. Elles ne sont cependant pas suffisantes pour extraire directement le signal diphoton. Il est nécessaire d'estimer le bruit de fond résiduel pour procéder à la mesure de sections efficaces différentielles de paires de photons. La signification attendue pour un boson de Higgs de masse 125 GeV étant de  $1.6\sigma$  dans le canal  $H \rightarrow \gamma\gamma$  pour l'ensemble des données à  $\sqrt{s} = 7$  TeV avec des coupures optimisées et une catégorisation des événements, la contribution du boson de Higgs a été jugée négligeable dans cette étude.

Nous verrons, en premier lieu, la sélection appliquée aux événements et les critères utilisés pour les classer en candidats signal ou candidats bruit de fond. La méthode utilisée pour extraire le bruit de fond lié à la mauvaise identification des jets en photons sera détaillée dans un second temps. Après l'application de cette méthode, les candidats signal sont divisés en quatre composantes  $\gamma\gamma$ ,  $\gamma$ -jet, jet- $\gamma$  et jet-jet, il est alors nécessaire d'estimer la contamination des électrons dans la composante  $\gamma\gamma$ . Une fois la composante  $\gamma\gamma$  isolée et corrigée la méthode de déconvolution permet d'accéder au nombre réel d'événements  $\gamma\gamma$ . Les sections efficaces différentielles de production mesurées peuvent alors être comparées aux prédictions théoriques. J'ai principalement contribué à l'extraction des nombres d'événements  $\gamma\gamma$ ,  $\gamma$ -jet, jet- $\gamma$ , jet-jet et à l'évaluation de leurs erreurs systématiques, cette partie sera donc particulièrement détaillée.

## 5.1 Sélection des événements

### 5.1.1 Les données et le Monte-Carlo

Les données analysées correspondent aux données collectées par le trigger `2g20_loose` durant l'année 2011 pour des collisions à  $\sqrt{s} = 7$  TeV. Le trigger `2g20_loose` est le plus bas trigger diphoton sans pre-scale (voir 2.2.4) durant l'année 2011 ; il sélectionne les événements avec au moins deux photons passant la sélection LOOSE-trigger et avec une énergie transverse supérieure à 20 GeV. En excluant les données pour lesquelles la qualité n'est pas garantie, les données disponibles pour l'analyse représentent  $4.9 \text{ fb}^{-1}$  et le nombre moyen d'événements d'empilement s'élève à 8.8 par croisement de paquets. L'incertitude sur la luminosité a été estimée en 2011 à 3.9% pour l'ensemble des données à  $\sqrt{s} = 7$  TeV ([85] et [86]).

Comme nous le verrons par la suite la méthode utilisée pour l'extraction du signal s'appuie en partie sur la simulation Monte-Carlo. Les événements diphoton (processus dur et fragmentation) sont produits par deux générateurs Monte-Carlo différents :

- SHERPA [112] avec les fonctions de distribution des partons CTEQ6L1 [32]
- PYTHIA [113] avec les fonctions de distribution des partons MRST2007 [114]

Les événements photon+jet (jusqu'à cinq partons additionnels) sont issus du générateur ALPGEN [115] avec les PDFs CTEQ6L1. Les événements dijets sont issus d'un échantillon PYTHIA dans lequel un filtre est appliqué au niveau particule afin de sélectionner les événements les plus énergétiques.

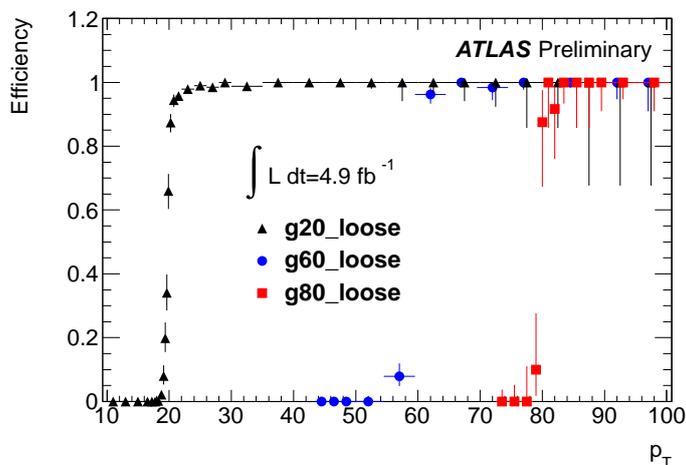
### 5.1.2 La présélection des événements

Pour être sélectionné, chaque événement ayant satisfait le trigger `2g20_loose` doit survivre à un certain nombre de coupures, nommé présélection :

- Vertex primaire : au moins un vertex primaire reconstruit avec au moins trois traces. Cette coupure permet de rejeter les événements non issus d'une collision.

- Acceptance : au moins deux photons LOOSE avec  $0 < |\eta_{s2}| < 1.37 \cup 1.52 < |\eta_{s2}| < 2.37$  où  $\eta_{s2}$  est la pseudo-rapidité mesurée dans la seconde couche du calorimètre. La zone  $1.37 < |\eta_{s2}| < 1.53$  correspond à l'interface entre les régions tonneau et bouchons ; elle est exclue car la réponse du calorimètre y est moins bien contrôlée. La coupure  $|\eta| < 2.37$  correspond à la région couverte par les strips de la première couche du calorimètre, essentielle à l'identification des photons.
- Coupure en  $p_T$  :  $p_T^{\gamma_1} > 25$  GeV (pour le photon le plus énergétique  $\gamma_1$ ) et  $p_T^{\gamma_2} > 22$  GeV (pour le second photon  $\gamma_2$ ). L'efficacité du trigger `2g20_loose` sur l'ensemble de la prise de données à  $\sqrt{s} = 7$  TeV est alors de  $97.8^{+0.8}_{-1.5}(\text{stat})\%$  (section 5.4.1).

Avec un trigger à 20 GeV, la plus petite coupure en  $p_T$  autorisée sur les candidats photons afin d'obtenir une efficacité proche de 1 est de 22 GeV (figure 5.2). La coupure asymétrique en



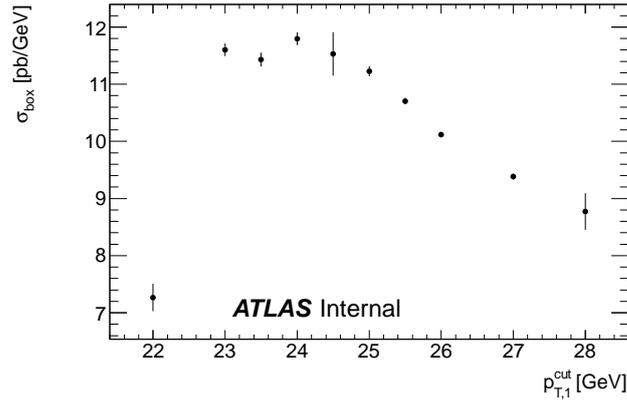
**Figure 5.2** – Efficacités des trigger `g20_loose`, `g60_loose` et `g80_loose` pour les candidats photons en fonction de l'énergie transverse des photons. L'efficacité pour le trigger `g20_loose` correspond aux triangles noirs [116].

$p_T$  est requise uniquement pour pouvoir comparer la section efficace mesurée aux prédictions théoriques. Du fait d'une instabilité dans le calcul du processus de boîte à NLO, les prédictions pour ce processus ne sont pas correctes lorsqu'une coupure symétrique est appliquée sur les photons. Cette instabilité est illustrée figure 5.3, il s'agit de la section efficace calculée à NLO par le générateur `Gamma2MC` pour le processus de boîte en fonction de la coupure en  $p_T$  sur le premier photon quand une coupure de 22 GeV est appliquée sur le  $p_T$  du second photon. L'augmentation de la section efficace avec l'augmentation de la coupure en  $p_T$  sur le premier photon jusqu'à 24 GeV n'est pas physique ; elle est liée à une défaillance du générateur dans la région où les impulsions transverse des deux photons sont proches. Une coupure à 25 GeV est choisie pour éviter cet effet.

### 5.1.3 Choix de la paire de photons

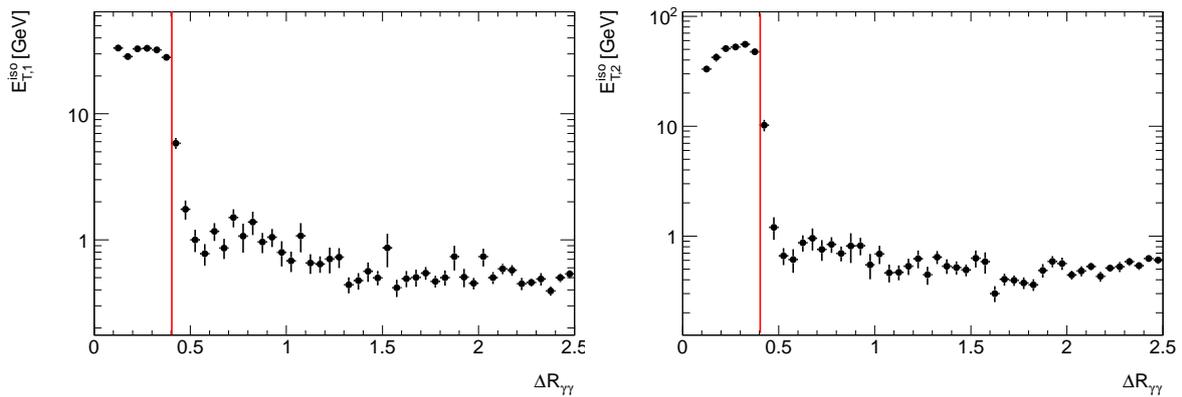
Dans les événements présélectionnés, la paire de photons choisie est la paire la plus énergétique. Les événements avec plus de deux photons sont conservés mais aucun usage n'est fait des photons supplémentaires. La variable d'isolation `topoEtCone` (définie section 4.3.1) est utilisée pour cette analyse et est par la suite notée  $E_T^{iso}$ . Elle correspond à l'énergie transverse mesurée dans un cône de rayon  $R = 0.4$  auquel a été soustraite l'énergie transverse du photon. Afin de limiter la superposition des cônes de chaque photon de la paire, une distance minimale  $\Delta R$  est requise entre les deux photons :

$$\Delta R = \sqrt{(\eta_{\gamma_1} - \eta_{\gamma_2})^2 + (\phi_{\gamma_1} - \phi_{\gamma_2})^2} \quad (5.1)$$



**Figure 5.3** – Section efficace prédite à NLO pour le processus de boîte par le générateur *Gamma2MC* en fonction de  $p_T^{\gamma 1}$  lorsqu’une coupure de 22 GeV est appliquée sur  $p_T^{\gamma 2}$  [117].

Si une séparation à 0.8 serait la plus naturelle pour éviter toute superposition des cônes, elle pose un problème de statistique. La figure 5.4 présente l’énergie d’isolation du premier et du second photon en fonction de la coupure appliquée sur  $\Delta R$  pour les événements diphoton générés avec *SHERPA*. On constate qu’une coupure à 0.4 est suffisante pour réduire le dépôt d’énergie supplémentaire lié à la superposition des cônes.



**Figure 5.4** – Energie d’isolation du premier et du second photon en fonction de la coupure appliquée sur  $\Delta R$  pour les événements diphoton générés avec *SHERPA*.

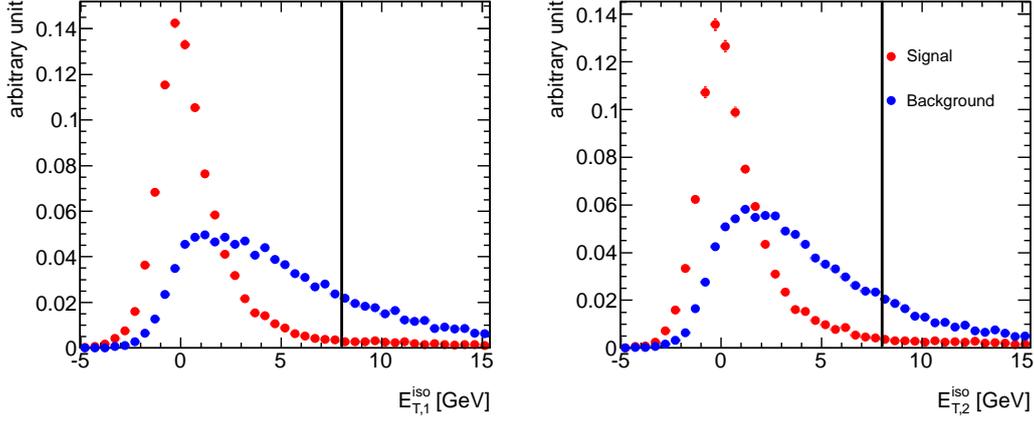
Par ailleurs, une coupure  $E_T^{iso} < 8$  GeV est appliquée pour enrichir les données en photons car la région à haute énergie d’isolation est principalement peuplée par du bruit de fond (figure 5.5).

### 5.1.3.1 Candidats signal

Afin d’être considéré comme candidat signal, les deux candidats photons de la paire ainsi sélectionnée doivent passer le critère d’identification TIGHT (noté aussi **T**) et vérifier  $E_T^{iso} < 4$  GeV (noté **I**). Cette coupure à 4 GeV, avec la nouvelle variable d’isolation *topoEtCone*, est équivalente à celle appliquée pour la mesure de section efficace diphoton précédente en 2010 [40] avec l’ancienne variable d’isolation *EtCone* (voir section 4.3.1 figure 4.11).

### 5.1.3.2 Candidats bruit de fond

Si la paire de photons ne passe pas l’identification TIGHT mais passe la sélection LOOSE’, l’événement est classé comme candidat bruit de fond et noté NON-TIGHT ou  $\tilde{T}$ . La sélection



**Figure 5.5** – Distribution de l'énergie d'isolation pour les photons (signal) et les jets (background) dans l'échantillon ALPGEN. Les distributions de gauche correspondent aux premiers objets les plus énergétiques, les distributions de droite aux seconds.

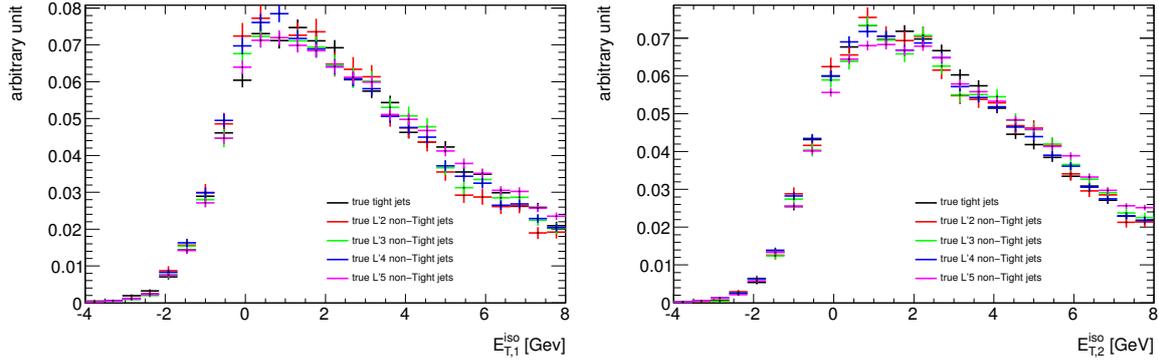
LOOSE' (noté aussi **L'**) diffère de la sélection TIGHT pour les coupures sur les variables de la première couche du calorimètre EM. La méthode d'extraction du bruit de fond (détaillée 5.2.1) étant basée sur la variable d'isolation, il est en effet nécessaire de définir une région de contrôle du bruit de fond telle que le profil de la variable d'isolation pour le bruit de fond dans cette région soit le même que celui passant la coupure d'identification TIGHT. L'énergie d'isolation est calculée en excluant le centre du cluster où la plupart des variables des strips sont définies, ces variables ne sont donc pas corrélées avec la variable d'isolation (voir annexe B). En relâchant une ou plusieurs coupures sur les variables des strips on peut ainsi définir une région de contrôle du bruit de fond. Le choix des coupures qui sont relâchées est cependant arbitraire. On définit quatre sélections :

- LOOSE' ou **L'** : coupures relâchées sur  $E_{ratio}$ ,  $\Delta E$ ,  $F_{side}$ ,  $w_{ts3}$ .
- LOOSE3' ou **L3'** : coupures relâchées sur  $\Delta E$ ,  $F_{side}$ ,  $w_{ts3}$ .
- LOOSE2' ou **L2'** : coupures relâchées sur  $F_{side}$ ,  $w_{ts3}$ .
- LOOSE5' ou **L5'** : coupures relâchées sur  $E_{ratio}$ ,  $\Delta E$ ,  $F_{side}$ ,  $w_{ts3}$ ,  $w_{stot}$ .

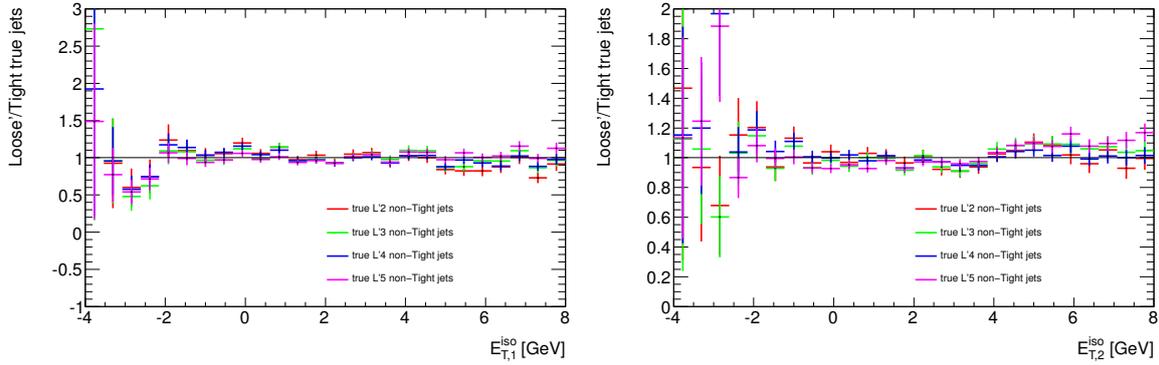
La figure 5.6 montre la distribution pour des événements générés avec ALPGEN de  $E_T^{iso}$  pour des jets TIGHT et pour des jets passant les différentes sélections LOOSE'. Pour rendre la comparaison plus facile, les rapports des distributions TIGHT/LOOSE' sont également présentés figure 5.7. Aucun biais significatif n'est observé pour les sélections LOOSE', LOOSE2', LOOSE3'. Pour la sélection LOOSE5', un très léger biais est observé pour le second jet. La sélection finale choisie est LOOSE' et les autres sélections seront utilisées pour le calcul des erreurs systématiques (voir 5.2.3).

## 5.2 Estimation du bruit de fond $\gamma$ -jet et jet-jet

L'ensemble des candidats de signal est principalement constitué d'événements  $\gamma$ - $\gamma$ ,  $\gamma$ -jet, jet- $\gamma$  et jet-jet, et nous avons besoin d'extraire la composante  $\gamma\gamma$  pour les mesures de sections efficaces. Les quatre composantes peuvent être séparées en utilisant le pouvoir discriminant des variables d'identification et d'isolation. La méthode utilisée, la méthode  $2 \times 2D$  sideband, est une méthode de comptage reposant sur ces deux variables.



**Figure 5.6** – Energie d'isolation du premier et du second jets pour les événements TIGHT et NON-TIGHT simulés avec ALPGEN.



**Figure 5.7** – Ratio de l'énergie d'isolation des jets TIGHT et NON-TIGHT pour les événements simulés avec ALPGEN.

### 5.2.1 Description de la méthode $2 \times 2D$ sideband

L'échantillon de départ **L'L'** est constitué de tous les événements passant la sélection LOOSE' (pour les deux candidats de la paires), soit 653355 événements. Le nombre total de ces événements peut s'écrire comme la somme des différentes composantes  $\gamma\text{-}\gamma$ ,  $\gamma\text{-}jet$ ,  $jet\text{-}\gamma$ ,  $jet\text{-}jet$  :

$$N_{TOTAL} = W_{\gamma\gamma}^{L'L'} + W_{\gamma j}^{L'L'} + W_{j\gamma}^{L'L'} + W_{jj}^{L'L'} \quad (5.2)$$

Parmi ces événements, les candidats de signal sont ceux passant les deux coupures supplémentaires : identification TIGHT et  $E_T^{iso} < 4$  GeV, comme décrit dans la section précédente, soit 165767 événements. On peut là encore diviser ces événements **TITI** (pour  $\gamma_1$  : **T**ight-**I**solé et  $\gamma_2$  : **T**ight-**I**solé) en quatre composantes :

$$W_{tot}^{TITI} = W_{\gamma\gamma}^{TITI} + W_{\gamma j}^{TITI} + W_{j\gamma}^{TITI} + W_{jj}^{TITI} \quad (5.3)$$

On cherche à déterminer les nombres réels d'événements  $W_{xy}^{TITI}$ . On exprime chacun des  $W_{xy}^{TITI}$  en fonction de  $W_{xy}^{L'L'}$  à l'aide des efficacités et coefficients de corrélation définis tels que :

- $\varepsilon_{T1}$  et  $\varepsilon_{T2}$  : efficacités d'identification pour le premier et second photon.
- $\varepsilon_{I1}$  et  $\varepsilon_{I2}$  : efficacités d'isolation pour le premier et second photon.
- $f_{T1}$  et  $f_{T2}$  : probabilités de passer la coupure d'identification pour le premier et second jet quand l'autre candidat de la paire est un photon, c'est ce que l'on nomme par la suite *fake-rates* d'identification.

- $f'_{T1}$  et  $f'_{T2}$  : probabilités de passer la coupure d'identification pour le premier et second jet quand l'autre candidat de la paire est un jet.
- $f_{I1}$  et  $f_{I2}$  : probabilités de passer la coupure d'isolation pour le premier et second jet quand l'autre candidat de la paire est un photon, c'est ce que l'on nomme par la suite *fake-rates* d'isolation.
- $f'_{I1}$  et  $f'_{I2}$  : probabilités de passer la coupure d'isolation pour le premier et second jet quand l'autre candidat de la paire est un jet.
- $\xi_{j1}$  et  $\xi_{j2}$  : coefficients de corrélation entre fake-rates d'isolation et d'identification pour le premier et second jet.
- $\xi_{Ijj}$  : corrélation entre l'isolation des jets pour la composante  $jj$ .

Les corrélations d'isolation pour les événements  $\gamma\gamma$ ,  $\gamma$ -jet et jet- $\gamma$  sont estimées à partir de la simulation Monte-Carlo (voir annexe B). Elles sont respectivement de  $3.0 \pm 0.2\%$ ,  $2.1 \pm 0.7\%$  et  $4.0 \pm 1.0\%$  et sont négligées. La corrélation pour les événements jet-jet (estimée à 8% dans la simulation) est prise en compte ( $\xi_{Ijj}$ ).

Les  $W_{xy}^{TITI}$  s'écrivent alors :

$$W_{\gamma\gamma}^{TITI} = W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} \varepsilon_{I2} \varepsilon_{T2} \quad (5.4)$$

$$W_{\gamma j}^{TITI} = W_{\gamma j}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} f_{I2} f_{T2} \xi_{j2} \quad (5.5)$$

$$W_{j\gamma}^{TITI} = W_{j\gamma}^{L'L'} \varepsilon_{I2} \varepsilon_{T2} f_{I1} f_{T1} \xi_{j1} \quad (5.6)$$

$$W_{jj}^{TITI} = W_{jj}^{L'L'} f'_{I1} f'_{T1} f'_{I2} f'_{T2} \xi_{Ijj} \xi_{j1} \xi_{j2} \quad (5.7)$$

Outre la région **TITI**, les coupures d'identification et d'isolation appliquées aux deux candidats permettent de définir 16 régions comme l'illustre la figure 5.8.

**AA** correspond à la région du signal **TITI** déjà définie, et dans toutes les autres régions, le nombre total d'événements peut également s'écrire en fonction des  $W_{xy}^{L'L'}$ . Les équations complètes sont données en annexe C.

Le système obtenu est un système de 16 équations à 19 inconnues. On fixe six paramètres d'entrée :

- $\varepsilon_{T1}$  et  $\varepsilon_{T2}$  : calculées à partir du Monte-Carlo signal diphoton (voir section 5.2.2).
- $\varepsilon_{I1}$  et  $\varepsilon_{I2}$  : calculées à partir du Monte-Carlo signal diphoton (voir section 5.2.2).
- $\xi_{j1}$  et  $\xi_{j2}$  : fixés à 1, ils sont ensuite variés dans l'intervalle  $(1 \pm 0.01)$  pour évaluer l'erreur systématique correspondante (voir section 5.2.2).

Ces efficacités et corrélations peuvent être calculées à partir de la simulation sans introduire de biais important. Le système se réduit alors à un système sur-contraint de 16 équations à 13 inconnus. Il est résolu en ajustant les 13 inconnues sur les 16 équations (fit). Les paramètres extraits sont alors :

$$- W_{\gamma\gamma}^{L'L'}, W_{\gamma j}^{L'L'}, W_{j\gamma}^{L'L'}, W_{jj}^{L'L'}$$

$$- f_{T1}, f_{T2}, f'_{T1}, f'_{T2}$$

$$- f_{I1}, f_{I2}, f'_{I1}, f'_{I2}$$

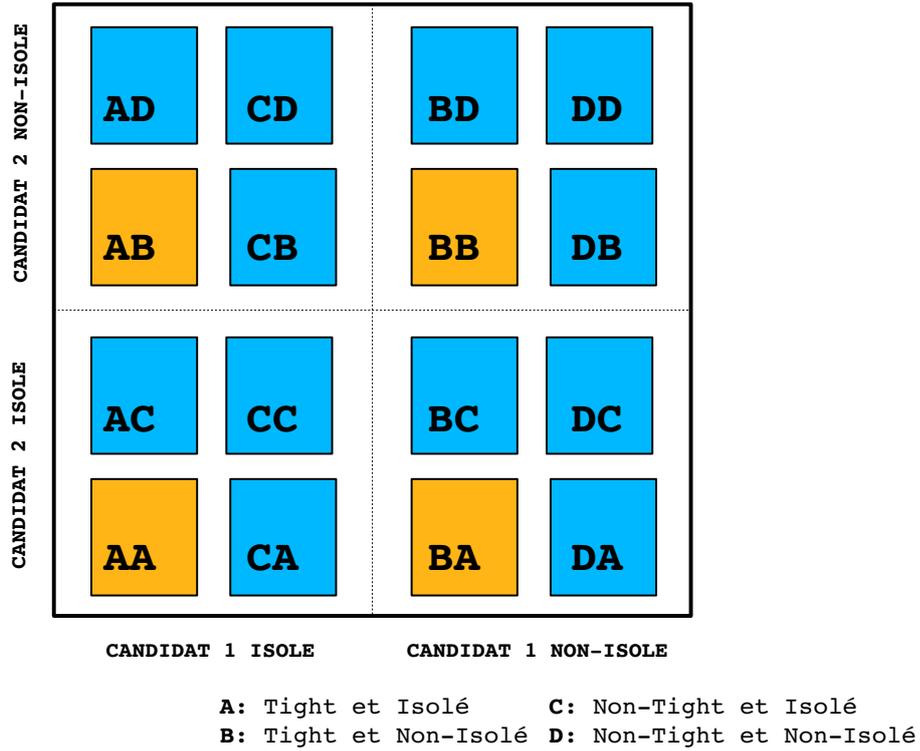


Figure 5.8 – Illustration des 16 régions définies pour la méthode  $2 \times 2D$  sideband.

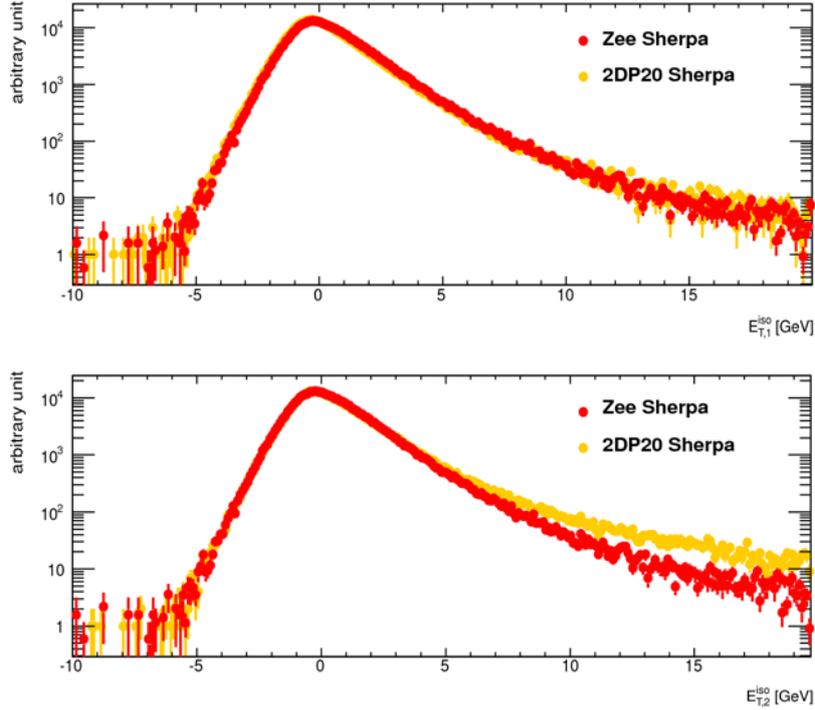
–  $\xi_{Ijj}$

### 5.2.2 Paramètres d’entrée de la méthode $2 \times 2D$ sideband

Il n’a pas été possible de définir dans les données un échantillon de contrôle présentant le même profil d’isolation que les événements diphoton. La figure 5.9 présente la distribution de l’énergie d’isolation pour les événements diphoton et  $Z \rightarrow e^+e^-$  générés avec SHERPA. Les queues présentes à haute énergie d’isolation pour les événements diphoton correspondent aux événements où des partons additionnels issus du processus dur sont produits. Cette contribution spécifique aux événements diphoton rend difficile l’extrapolation du profil d’isolation des photons à partir des électrons. On constate par ailleurs que la queue est plus large pour le second photon ; en effet, dans ces événements, le second photon est en général plus proche des partons issus du processus dur.

Les efficacités d’identification et d’isolation de référence sont déterminées à partir des événements diphoton simulés avec SHERPA après correction d’un décalage entre données et Monte-Carlo pour la distribution de la variable d’isolation, illustré figure 5.10. En raison de la différence observée précédemment entre la distribution de l’énergie d’isolation pour les électrons ( $Z \rightarrow ee$ ) et celle pour les événements diphoton, ce décalage n’est pas estimé à partir de la comparaison entre données et Monte-Carlo pour les événements  $Z \rightarrow ee$ , mais à partir d’une méthode de *template fit* décrite dans [40] et [99]. Le décalage trouvé est de  $160 \pm 10$  MeV pour le premier photon et de  $120 \pm 5$  MeV pour le second photon.

Les efficacités d’isolation et d’identification inclusives calculées à partir de la simulation diphoton SHERPA sont présentées tableau 5.1. Les efficacités sont plus élevées pour le photon de plus haut  $p_T$  ; la pureté augmente avec l’énergie et la probabilité que le photon de plus bas



**Figure 5.9** – Comparaison des énergies d’isolation des photons et des électrons pour les événements diphoton et  $Z \rightarrow e^+e^-$  simulés avec *SHERPA*.

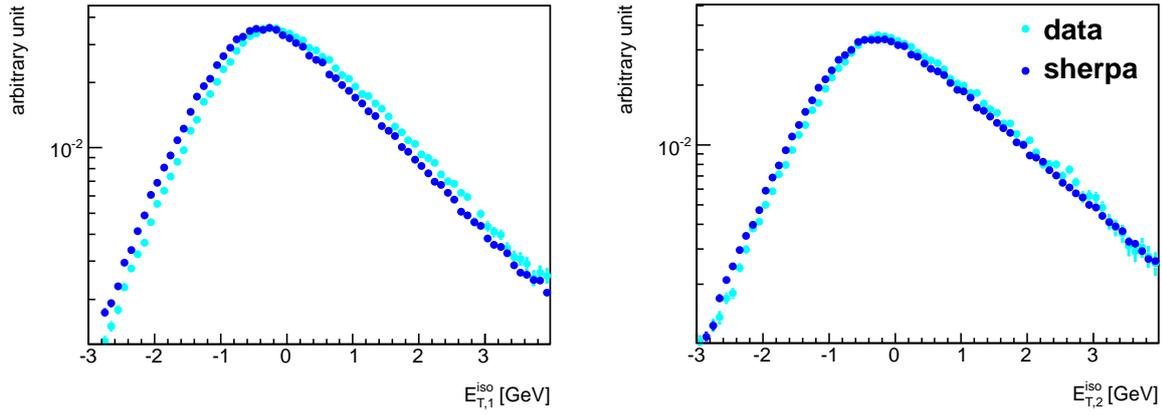
$p_T$  soit issu d’un processus de fragmentation, donc moins isolé, est plus élevée.

Efficacités	
paramètres d’entrée	
$e_{I1}$	$0.9684 \pm 0.0003$
$e_{I2}$	$0.9578 \pm 0.0004$
$e_{T1}$	$0.9334 \pm 0.0005$
$e_{T2}$	$0.9168 \pm 0.0005$

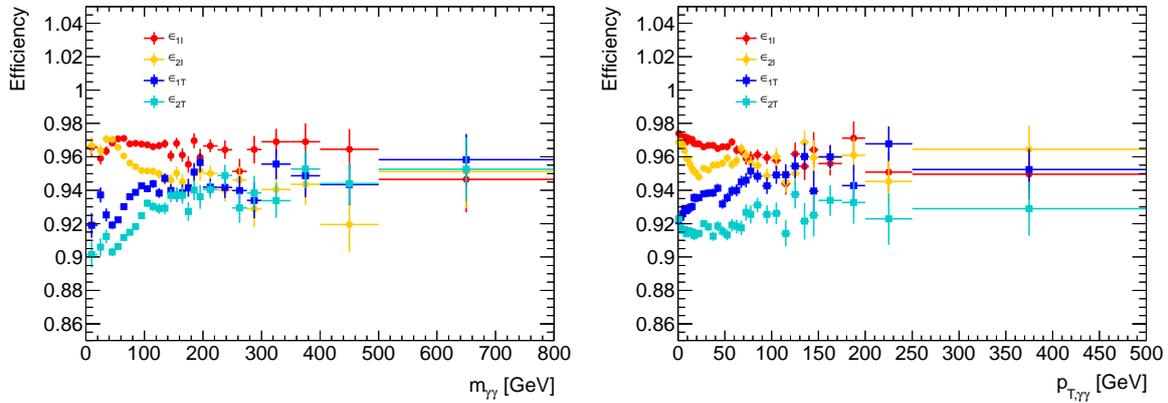
**Tableau 5.1** – Efficacités d’isolation et d’identification utilisées comme paramètres d’entrée à la  $2 \times 2D$  sideband. Ces efficacités sont calculées à partir des événements diphoton générés avec *SHERPA*, les erreurs sont statistiques uniquement.

Pour mesurer des sections efficaces différentielles il faut obtenir la distribution du nombre d’événements  $\gamma$ - $\gamma$  en fonction des variables cinématiques choisies, à savoir  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$  et  $\cos\theta^*$ . La méthode  $2 \times 2D$  sideband est appliquée pour chaque bin de la distribution de chaque variable. Les efficacités d’identification et d’isolation, paramètres d’entrée de la méthode, sont calculées bin par bin, comme illustré figure 5.11 et 5.12. Lorsque la statistique dans un bin est trop faible, celui-ci est groupé avec ses deux bins voisins afin de déterminer les paramètres de sortie.

Les corrélations entre les variables d’isolation et d’identification pour les jets,  $\xi_{j1}$  et  $\xi_{j2}$ , ont été estimées à partir de l’échantillon *ALPGEN* et sont respectivement de  $1.0057 \pm 0.0022$  et de  $0.9919 \pm 0.0019$ . La méthode  $2 \times 2D$  sideband utilise  $\xi_{j1} = \xi_{j2} = 1$  et une erreur systématique



**Figure 5.10** – Energie d’isolation pour le premier (à gauche) et second photon (à droite). La distribution extraite de l’échantillon d’événements diphoton générés avec *SHERPA* est en bleu foncé. La distribution estimée à partir des données, suivant la méthode décrite section 5.2.3.1, est en cyan. L’intégrale de chaque distribution est normalisée à 1.



**Figure 5.11** – Efficacités d’identification et d’isolation pour le premier et second photon en fonction de  $m_{\gamma\gamma}$  (à gauche) et de  $p_{T,\gamma\gamma}$  (à droite) obtenues à partir des événements diphoton générés par *SHERPA*.

est associée à ce choix (voir 5.2.3).

## 5.2.3 Incertitudes systématiques

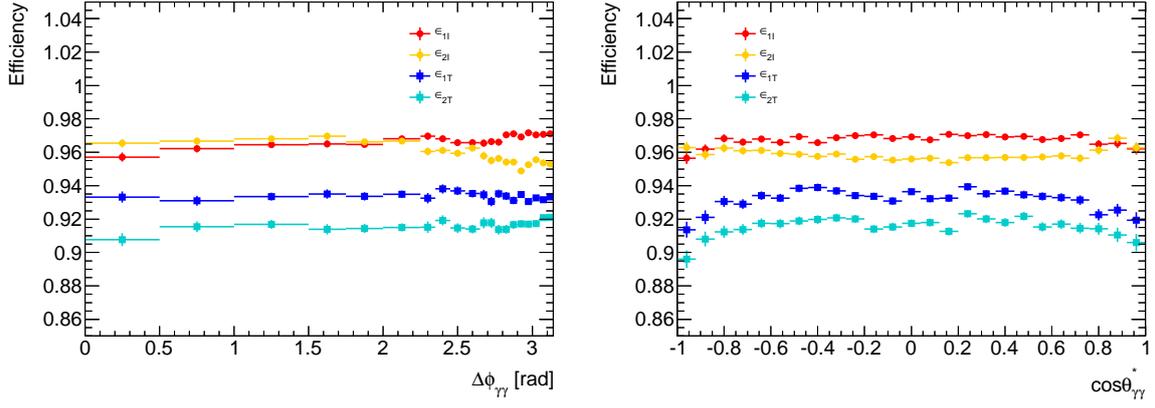
### 5.2.3.1 Définition de la région de contrôle du bruit de fond

La région de contrôle du bruit de fond est basée sur la définition de la sélection LOOSE’. On a vu section 5.1.3 qu’il y a un certain arbitraire dans le choix de cette sélection. Afin d’estimer l’incertitude liée au choix de la sélection LOOSE’, l’extraction du signal est répétée en considérant les trois autres sélections LOOSE’2, LOOSE’3 et LOOSE’5.

L’erreur systématique liée aux coefficients de corrélation entre l’isolation et l’identification des jets est également estimée. La méthode  $2 \times 2D$  sideband utilise  $\xi_{j1} = \xi_{j2} = 1$ . Le choix de l’intervalle parcouru pour l’estimation des erreurs systématiques est fait à partir de la variable  $\theta$  :

$$\theta_i = \frac{\xi_{ji}(1 - f_{Ii} - f_{Ti} + f_{Ii}f_{Ti}\xi_{ji})}{(1 - f_{Ii}\xi_{ji})(1 - f_{Ti}\xi_{ji})} \quad (5.8)$$

avec  $i = 1, 2$  pour le premier et second photon et  $\theta = 1 \pm 0.05$  comme déterminé à partir de la simulation dans [118]. Avec les valeurs moyennes obtenues pour les fake-rates, ceci se traduit



**Figure 5.12** – Efficacités d’identification et d’isolation pour le premier et second photon en fonction de  $\Delta\phi_{\gamma\gamma}$  (à gauche) et de  $\cos\theta^*$  (à droite) obtenues à partir des événements diphoton générés par SHERPA.

par  $\xi_j \pm 1\%$ . Dans la mesure où la corrélation isolation-identification des jets dépend directement de la définition du bruit de fond, cet effet est déjà inclus dans l’erreur systématique liée à la définition de LOOSE’.

Les différences observées sur le nombre d’événements par composante entre les sélections décrites ici et la sélection nominale LOOSE’ avec  $\xi_{j1} = \xi_{j2} = 1$  sont présentées tableau 5.2.

	LOOSE’2	LOOSE’3	LOOSE’5	$\xi_j \pm 0.01$
$W_{\gamma\gamma}^{TITI}$	-7710	-606	+4880	+283
$W_{\gamma j}^{TITI}$	+3880	+610	-2740	-303
$W_{j\gamma}^{TITI}$	+2260	+293	-751	+89
$W_{jj}^{TITI}$	+1580	-288	-1370	-70

**Tableau 5.2** – Différences entre le nombre d’événements inclusif obtenu en modifiant la définitions de la région de contrôle du bruit de fond et le nombre d’événements inclusif obtenu avec la sélection nominale LOOSE’ (pour laquelle  $W_{\gamma\gamma}^{TITI} = 113156$ ,  $W_{\gamma j}^{TITI} = 31495$ ,  $W_{j\gamma}^{TITI} = 12978$ ,  $W_{jj}^{TITI} = 8158$ ). Les erreurs sont statistiques uniquement. La dernière colonne montre l’erreur systématique liée à la variation des coefficients de corrélation  $\xi_j$ .

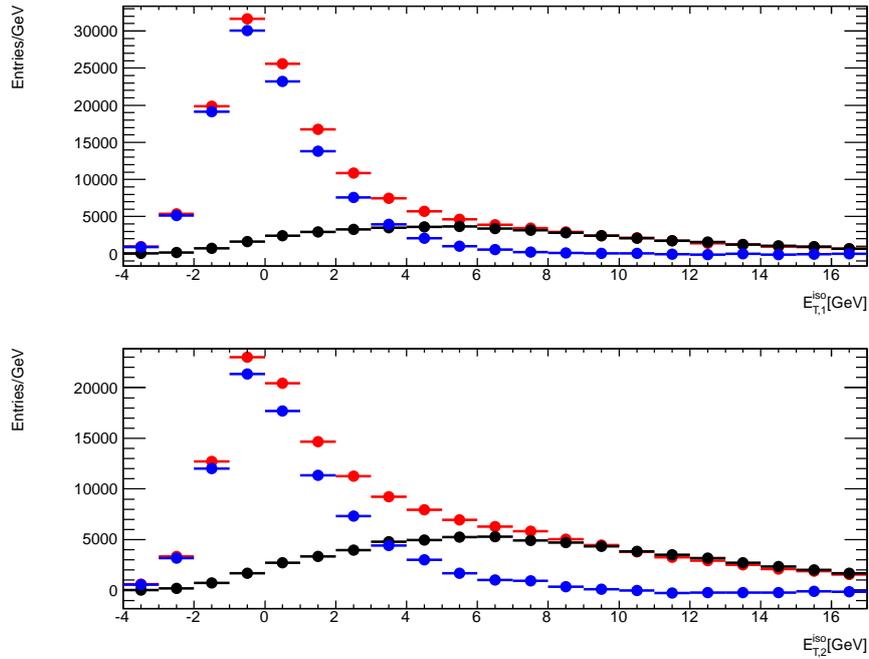
### 5.2.3.2 Efficacité d’isolation

La distribution nominale de l’énergie d’isolation des photons est obtenue à partir des événements diphoton générés par SHERPA. Dans la simulation SHERPA, au niveau des éléments de matrice, les calculs incluent la production d’événements diphoton avec un ou deux partons additionnels. La distance minimale entre un photon et un parton ainsi produits est cependant fixée à 0.3, ce qui réduit le nombre d’événements à haute énergie d’isolation. Pour estimer l’erreur systématique liée au choix du générateur SHERPA, les efficacités d’isolation sont également calculées à partir des données et à partir du générateur PYTHIA.

Afin d’extraire le profil de l’énergie d’isolation du signal dans les données, on considère que pour  $E_T^{iso} > 7$  GeV la distribution de l’énergie d’isolation correspond à du bruit de fond uniquement. La distribution de  $E_T^{iso}$  pour les événements NON-TIGHT est normalisée à celle des événements TIGHT dans l’intervalle 7-17 GeV. La différence de ces deux distributions, ainsi normalisées,

est une estimation de la distribution de l'énergie d'isolation des photons (figure 5.13). L'hypothèse de départ est cependant une approximation, la contribution des photons issus de la fragmentation n'est pas négligeable dans l'intervalle 7-17 GeV ; en faisant la normalisation du profil d'isolation dans cette région, la contribution du bruit de fond est surestimée. Il en résulte une légère surestimation de l'efficacité d'isolation calculée à partir du profil d'isolation obtenu par soustraction ; c'est pourquoi cette valeur n'est pas utilisée comme valeur nominale, mais pour le calcul des erreurs systématiques uniquement.

Dans la simulation PYTHIA, la région de l'espace des phases à haute énergie d'isolation est

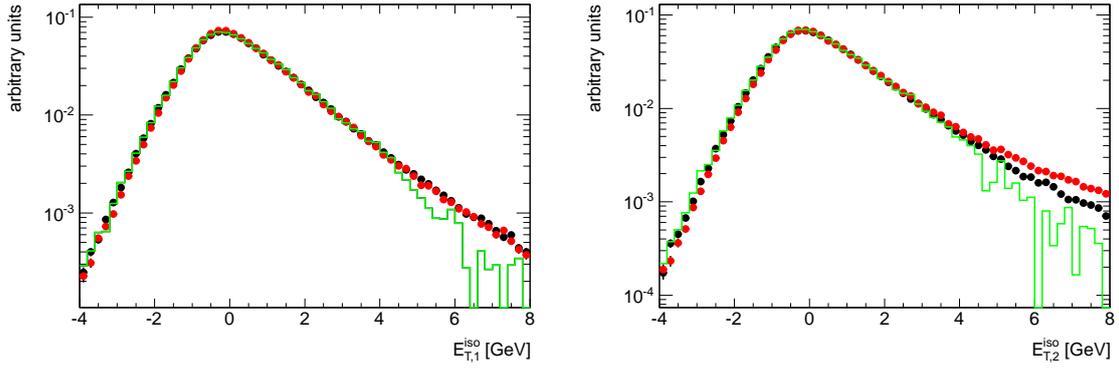


**Figure 5.13** – Profil de l'énergie d'isolation extrait des données pour le premier et second photon. La courbe rouge correspond à l'énergie d'isolation des candidat TIGHT et la courbe noire à celle des candidats NON-TIGHT. La courbe noire est normalisée à la courbe de rouge dans l'intervalle 7-17 GeV, la soustraction des deux courbes ainsi normalisées est la courbe bleue.

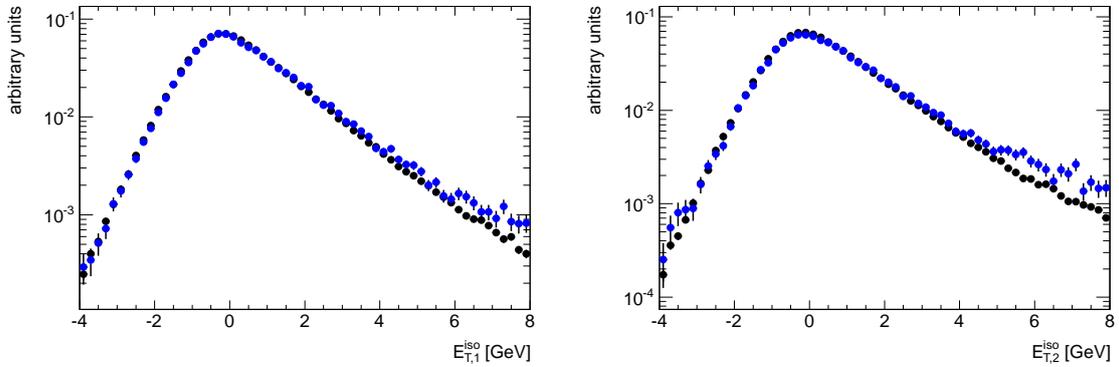
uniquement peuplée par des photons issus de la parton-shower. Figure 5.14 on constate, dans cette région, que la distribution PYTHIA est supérieure à celle de SHERPA. Les efficacités (inclusives) pour SHERPA, PYTHIA et à partir des données sont détaillées tableau 5.3.

Dans la méthode  $2 \times 2D$  sideband, on considère que l'efficacité d'isolation est la même pour le candidat photon si l'autre candidat de la paire est un photon ou un jet. Afin d'estimer l'impact de cette hypothèse, l'extraction du signal est répétée avec les efficacités d'isolation issues de la simulation SHERPA diphoton pour les événements  $\gamma\text{-}\gamma$  et celles issues de la simulation ALPGEN photon+jet pour les événements  $\gamma\text{-jet}$  et  $jet\text{-}\gamma$ . La figure 5.15 montre la comparaison des distributions pour le premier et second photon. Les queues de distribution sont moins importantes pour SHERPA en raison de la coupure à 0.3 appliquée sur la distance entre photon et parton. Les efficacités inclusives correspondantes sont détaillées tableau 5.3.

La distribution de l'énergie d'isolation dépend du développement de la gerbe EM et est sensible à la quantité de matière rencontrée par le photon. Il est ainsi nécessaire de vérifier l'impact de la connaissance limitée de la distribution du matériel en amont du calorimètre. Les efficacités d'isolation sont calculées à partir d'un échantillon d'événements diphoton générés avec PYTHIA pour lequel la géométrie du détecteur est modifiée, 5 à 20% de matériel sont rajoutés en fonction de  $\eta$  en amont du calorimètre. La figure 5.16 montre la comparaison de l'énergie d'isolation



**Figure 5.14** – Energie d’isolation pour le premier (à gauche) et second photon (à droite). La distribution nominale extraite de l’échantillon diphoton SHERPA est en noire, la distribution extraite de l’échantillon diphoton PYTHIA est en rouge et l’histogramme vert correspond à la distribution extraite directement des données. L’intégrale de chaque distribution est normalisée à 1.



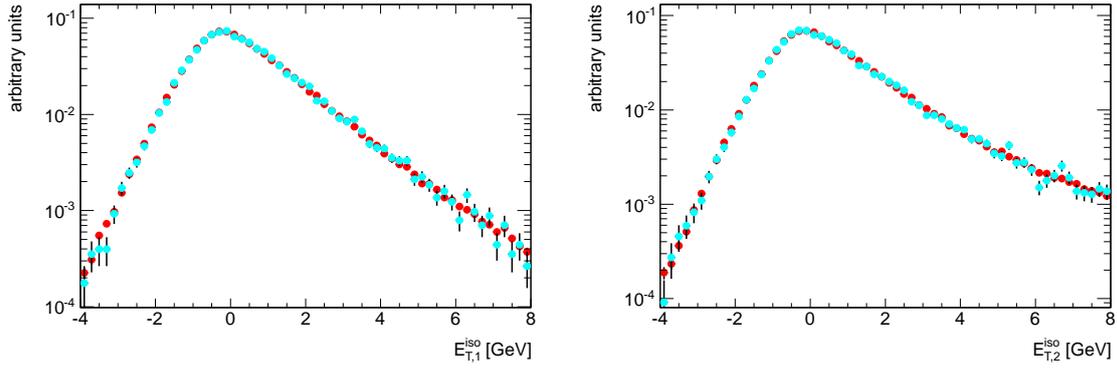
**Figure 5.15** – Energie d’isolation pour le premier (à gauche) et second photon (à droite). La distribution nominale extraite de l’échantillon diphoton SHERPA est en noire, la distribution extraite de l’échantillon photon+jets ALPGEN est en bleue. L’intégrale de chaque distribution est normalisée à 1.

issue des échantillons PYTHIA avec la géométrie nominale et avec une géométrie modifiée. Les efficacités d’isolation correspondantes sont détaillées tableau 5.3. La différence entre les deux échantillons PYTHIA est jugée négligeable, aucune erreur systématique n’est donc associée à cet effet.

	SHERPA	PYTHIA	PYTHIA géométrie modifiée	data-driven	ALPGEN
$\varepsilon_{I1}$	$0.9684 \pm 0.0003$	$0.9696 \pm 0.0003$	$0.9691 \pm 0.0012$	$0.9794 \pm 0.0003$	$0.9663 \pm 0.0013$
$\varepsilon_{I2}$	$0.9578 \pm 0.0004$	$0.9461 \pm 0.0005$	$0.9449 \pm 0.0015$	$0.9652 \pm 0.0009$	$0.9545 \pm 0.0019$

**Tableau 5.3** – Efficacités d’isolation utilisées en paramètres d’entrée de la méthode  $2 \times 2D$  sideband et calculées pour différents échantillons Monte-Carlo ou à partir des données. La région de contrôle du bruit de fond est LOOSE<sup>7</sup>. Les erreurs sont statistiques uniquement.

Comme cela a déjà été mentionné, un décalage est appliqué à la variable d’isolation dans la simulation Monte-Carlo afin de mieux décrire les données. L’erreur systématique liée à ce



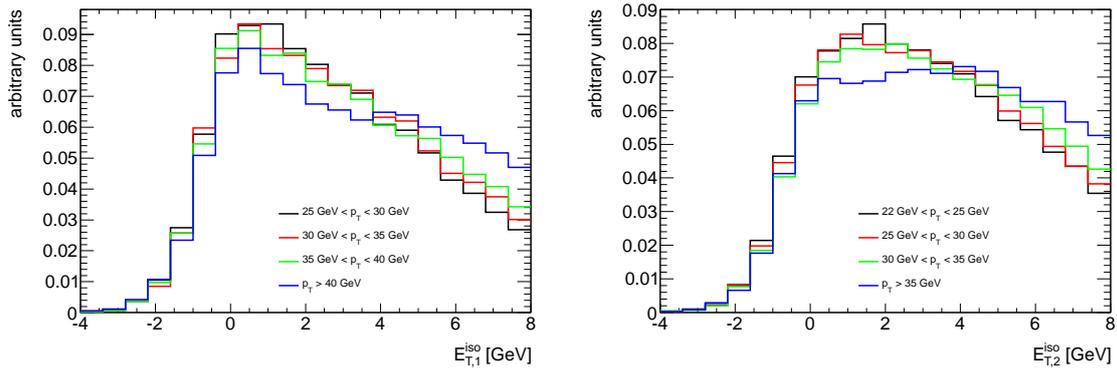
**Figure 5.16** – Energie d’isolation pour le premier (à gauche) et second photon (à droite). La distribution extraite de l’échantillon diphoton PYTHIA avec la géométrie nominale est en rouge, la distribution extraite de l’échantillon diphoton PYTHIA avec une géométrie modifiée est en cyan. L’intégrale de chaque distribution est normalisée à 1.

décalage est estimée en répétant l’extraction du signal en faisant varier le décalage appliqué dans l’intervalle décrit par son incertitude,  $160 \pm 10$  MeV pour le premier photon et de  $120 \pm 5$  MeV pour le second photon.

L’énergie d’isolation est corrigée de l’effet des fuites latérales en énergie du photon dans le cône d’isolation comme décrit section 4.3.1. La dépendance résiduelle en  $p_T$  après correction de l’énergie d’isolation des photons est faible. En revanche, pour le bruit de fond, la figure 5.17 montre une claire dépendance de l’énergie d’isolation en fonction du  $p_T$ . L’effet de cette dépendance a été estimé en extrayant le signal avec la méthode  $2 \times 2D$  sideband sur 13 catégories en  $p_T$  :

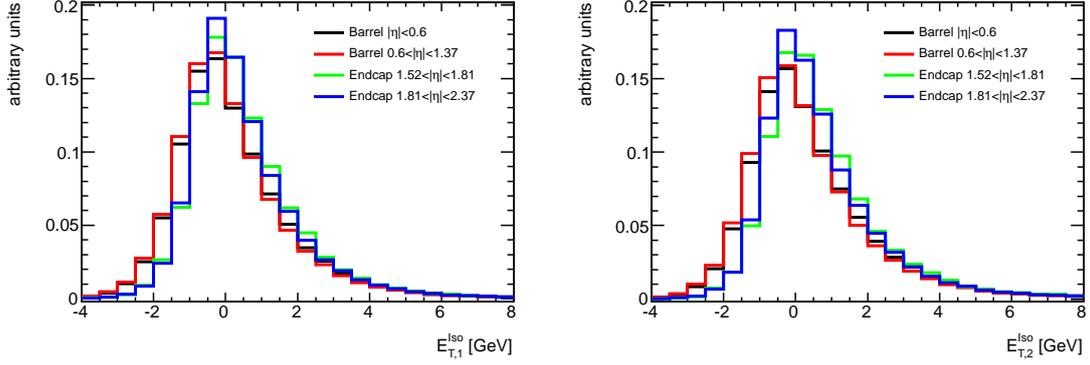
- 1<sup>er</sup> jet :  $25 < p_T < 30$ ,  $30 < p_T < 35$ ,  $35 < p_T < 40$  et  $p_T > 40$  GeV
- 2<sup>nd</sup> jet :  $22 < p_T < 25$ ,  $25 < p_T < 30$ ,  $30 < p_T < 35$  et  $p_T > 35$  GeV

Un biais de  $-0.3\%$  sur le nombre d’événements  $\gamma\text{-}\gamma$  a été mesuré et ajouté aux erreurs systématiques (tableau 5.4).

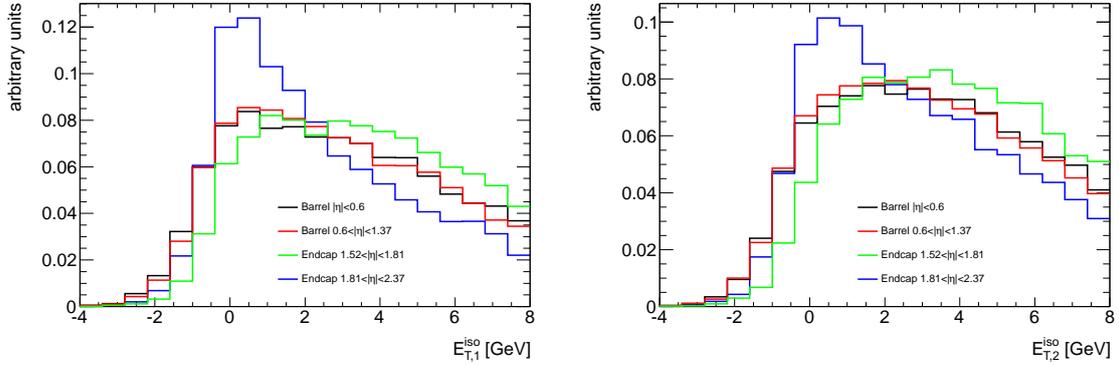


**Figure 5.17** – Energies d’isolation pour le premier (à gauche) et second (à droite) candidat bruit de fond (sélection LOOSE’) dans les données pour différentes régions en  $p_T$ . L’intégrale de chaque distribution est normalisée à 1.

Les figures 5.18 et 5.19 montrent une dépendance de l’énergie d’isolation en fonction de  $\eta$  pour les photons comme pour les jets. De la même façon que pour la dépendance en  $p_T$ , cet effet est estimé en répétant l’extraction du nombre d’événements pour plusieurs catégories en  $\eta$  :  $|\eta| < 0.6$ ,  $0.6 < |\eta| < 1.37$ ,  $1.52 < |\eta| < 1.81$  et  $1.81 < |\eta| < 2.37$  pour chacun des deux candidats objets. Un biais de  $0.02\%$  a été constaté sur la composante  $\gamma\text{-}\gamma$  et jugé négligeable (5.4). Aucune erreur systématique n’est donc associée à cet effet.



**Figure 5.18** – Energies d’isolation pour le premier (à gauche) et second photon (à droite). Les distributions sont extraites de l’échantillon diphoton *SHERPA* pour différentes régions en  $|\eta|$  du calorimètre EM. L’intégrale de chaque distribution est normalisée à 1.



**Figure 5.19** – Energies d’isolation pour le premier (à gauche) et second (à droite) candidat bruit de fond (sélection LOOSE’) dans les données pour différentes régions en  $|\eta|$ . L’intégrale de chaque distribution est normalisée à 1.

	$\eta$ categorisation	$p_t$ categorisation
$W_{\gamma\gamma}^{TITI}$	+23	-340
$W_{\gamma j}^{TITI}$	+104	+567
$W_{j\gamma}^{TITI}$	-42	+247
$W_{jj}^{TITI}$	-86	-522

**Tableau 5.4** – Déviation maximale obtenue sur le nombre d’événements inclusif obtenu avec les catégories en  $\eta$  et  $p_t$  par rapport aux nombres d’événements nominaux ( $W_{\gamma\gamma}^{TITI} = 113156$ ,  $W_{\gamma j}^{TITI} = 31495$ ,  $W_{j\gamma}^{TITI} = 12978$ ,  $W_{jj}^{TITI} = 8158$ )

L’effet sur l’ensemble des contributions est détaillé tableau 5.4. On constate que l’impact des dépendances en  $|\eta|$  et  $p_T$  est plus significatif pour les contributions  $\gamma$ -jet et jet-jet.

### 5.2.3.3 Efficacité d’identification

Comme pour l’isolation, la géométrie du matériel peut affecter le profil des variables d’identification. L’effet associé est estimé à partir des événements diphoton générés par *PYTHIA* avec une géométrie modifiée. Les efficacités correspondantes sont présentées tableau 5.5 et

les différences sur le nombre d'événements inclusifs  $\gamma\gamma$ ,  $\gamma$ -jet, jet- $\gamma$  et jet-jet entre les deux échantillons PYTHIA sont respectivement de 212, 166, 109 et 67. L'effet est jugé négligeable en comparaison de l'erreur liée à la définition de la sélection LOOSE' (tableau 5.6). Aucune erreur systématique n'est donc associée à cet effet.

Comme pour l'isolation, l'efficacité d'identification pour le candidat photon est la même si l'autre candidat de la paire est un photon ou un jet. La simulation ALPGEN photon+jet est utilisée pour estimer l'incertitude résultante de cette approximation. Les efficacités correspondantes sont présentées tableau 5.5.

Comme mentionné section 4.4.3, les distributions des variables d'identification du Monte-Carlo sont ajustées aux données par des coefficients correctifs. Afin d'estimer l'incertitude liée à cette correction, les erreurs sur les efficacités d'identification sont calculées avec l'échantillon SHERPA diphoton sans avoir appliqué ces corrections, et sont ensuite symétrisées pour obtenir l'effet sur le nombre d'événements (tableau 5.5).

	SHERPA	SHERPA sans corrections MC-données	PYTHIA	PYTHIA géométrie modifiée	ALPGEN
$\varepsilon_{T1}$	$0.9334 \pm 0.0005$	$0.9352 \pm 0.0005$	$0.9336 \pm 0.0005$	$0.9237 \pm 0.0018$	$0.9298 \pm 0.0018$
$\varepsilon_{T2}$	$0.9168 \pm 0.0005$	$0.9202 \pm 0.0005$	$0.9178 \pm 0.0005$	$0.9055 \pm 0.0019$	$0.9142 \pm 0.0026$

**Tableau 5.5** – Efficacités d'identification utilisées en paramètres d'entrée de la méthode  $2 \times 2D$  side-band et calculées pour différents échantillons Monte-Carlo. La région de contrôle du bruit de fond est LOOSE'. Les erreurs sont statistiques uniquement.

L'ensemble des erreurs systématiques sur le nombre d'événements inclusif par composante est présenté tableau 5.6. Le choix de la région de contrôle du bruit de fond est la principale source d'erreurs systématiques, elle est de l'ordre de  ${}^{+4}_{-6}\%$  sur le nombre d'événements  $\gamma\gamma$ .

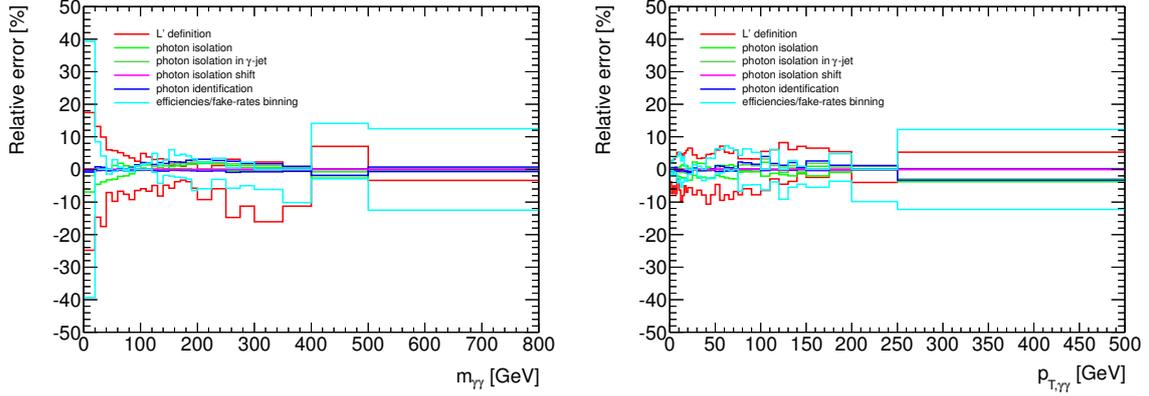
	L'2	L'5	PYTHIA	données	ALPGEN	Corrections MC-données	$\gamma$ MC décalage
$W_{\gamma\gamma}^{TITI}$	-7710	+4880	+1110	-1980	-99	+593 -585	+34 -36
$W_{\gamma j}^{TITI}$	+3880	-2740	-1470	+90.3	+299	+337 -343	+1.9 -2.2
$W_{j\gamma}^{TITI}$	+2260	-751	+571	+933	+144	+152 -154	+20 -18
$W_{jj}^{TITI}$	+1580	-1370	-243	+1020	-323	+96 -96	+20 -19

**Tableau 5.6** – Ensemble des erreurs systématiques considérées. Les différences sur les nombres d'événements sont calculés par rapport à la valeur centrale obtenue avec la sélection LOOSE' pour la définition de la région de contrôle du bruit de fond (pour laquelle  $W_{\gamma\gamma}^{TITI} = 113156$ ,  $W_{\gamma j}^{TITI} = 31495$ ,  $W_{j\gamma}^{TITI} = 12978$ ,  $W_{jj}^{TITI} = 8158$ ).

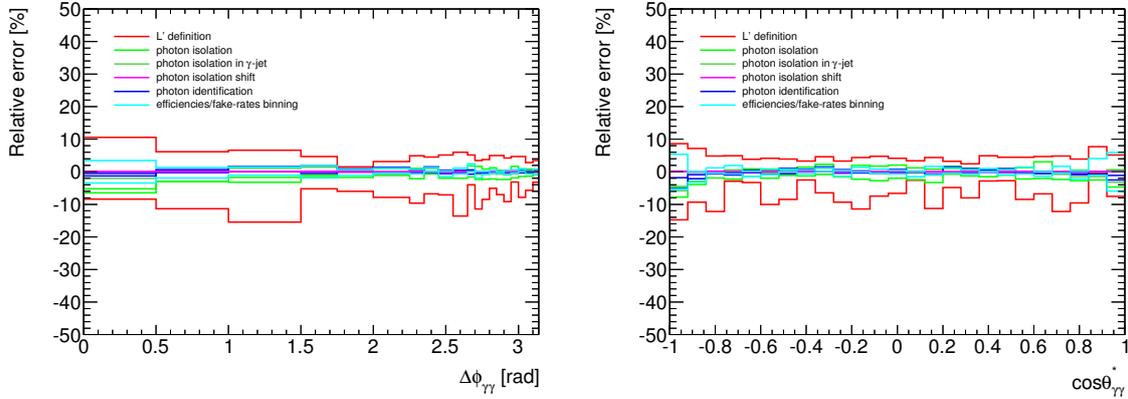
### 5.2.3.4 Erreurs systématiques en fonction des variables cinématiques

Pour les distributions différentielles du nombre d'événements, les erreurs systématiques sont estimées bin par bin. Les sources sont les mêmes que pour l'analyse inclusive, une seule source

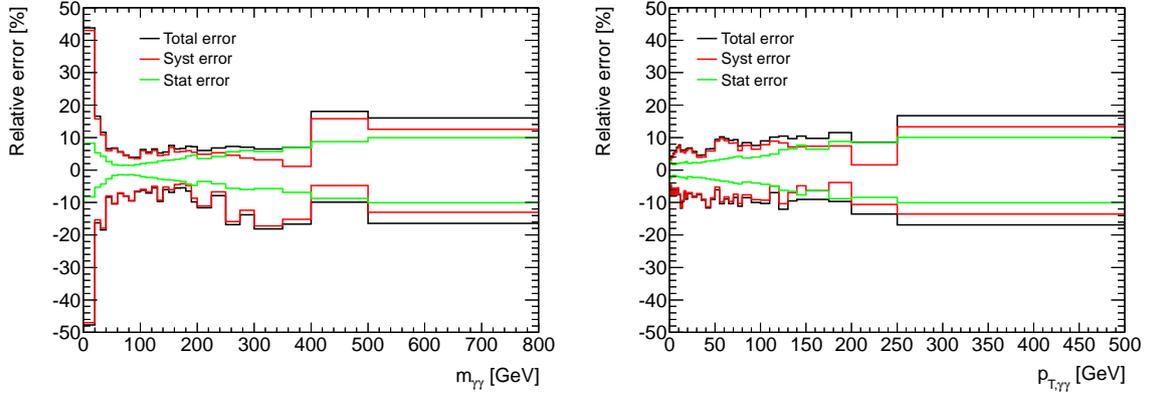
d'erreur s'ajoute ; lorsque la statistique est trop faible dans un bin, les paramètres d'entrée sont extraits en groupant le bin avec ses voisins. L'effet de cette approximation est évalué en répétant l'extraction avec le groupement de bins suivant et précédent. Pour le nombre d'événements  $\gamma\gamma$ , les variations des principales erreurs systématiques en fonction des variables cinématiques sont présentées figure 5.20 et 5.21 ; l'erreur liée à cet effet de binning est visible en cyan. L'erreur systématique totale, qui correspond à la somme en quadrature des erreurs systématiques individuelles, ainsi que l'erreur statistique sont présentées figure 5.22 et figure 5.23. Dans la très grande majorité de l'espace des phases, l'erreur dominante est l'erreur systématique, pour laquelle la principale contribution vient de la définition de la région de contrôle du bruit de fond.



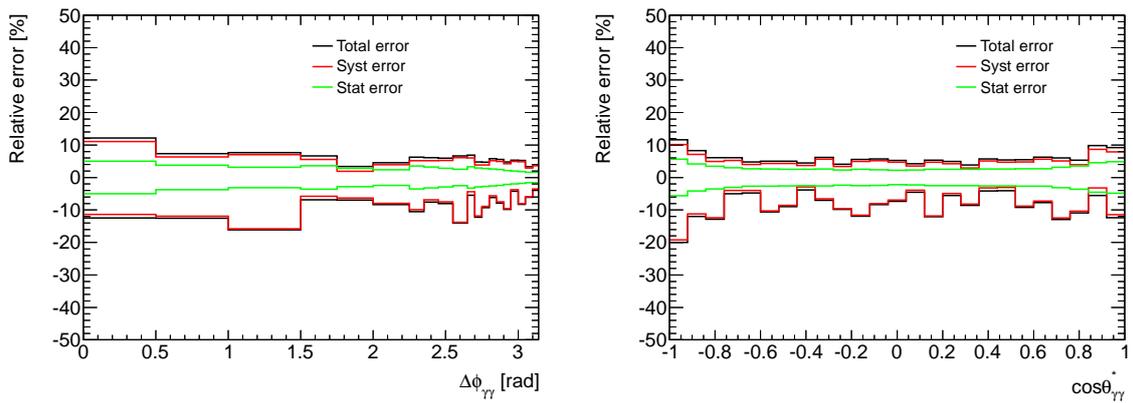
**Figure 5.20** – Incertitudes systématiques relatives aux nombres d'événements  $\gamma$ - $\gamma$  en fonction de  $m_{\gamma\gamma}$  (à gauche) et  $p_{T,\gamma\gamma}$  (à droite).



**Figure 5.21** – Incertitudes systématiques relatives aux nombres d'événements  $\gamma$ - $\gamma$  en fonction de  $\Delta\phi_{\gamma\gamma}$  (à gauche) et  $\cos\theta^*$  (à droite).



**Figure 5.22** – Incertitudes systématiques et statistiques relatives aux nombres d'événements  $\gamma\text{-}\gamma$  en fonction de  $m_{\gamma\gamma}$  (à gauche) et  $p_{T,\gamma\gamma}$  (à droite).



**Figure 5.23** – Incertitudes systématiques et statistiques relatives aux nombres d'événements  $\gamma\text{-}\gamma$  en fonction de  $\Delta\phi_{\gamma\gamma}$  (à gauche) et  $\cos\theta^*$  (à droite).

### 5.2.4 Décomposition en événements $\gamma\gamma$ , $\gamma$ -jet, jet- $\gamma$ et jet-jet

Le tableau 5.7 montre les fake-rates et coefficients de corrélation obtenus à partir de la  $2 \times 2D$  sideband. Les fake-rates d'identification  $f_{Ti}$  et  $f'_{Ti}$  étant quasiment identiques, l'approximation  $f_{Ti} = f'_{Ti}$  est finalement faite. Les fake-rates d'isolation  $f'_{Ii}$  pour la composante  $j$ - $j$  sont inférieurs aux fake-rates d'isolation  $f_{Ii}$  pour les composantes  $\gamma$ -jet et jet- $\gamma$ ; en effet, un jet contamine plus facilement l'énergie d'isolation de l'autre candidat. De même, le coefficient de corrélation entre l'isolation des jets est supérieur à 1 car un jet a plus de chance d'être non-isolé si l'autre jet est également non-isolé.

Fake-rates et coefficient de corrélation	
paramètres de sortie	
$f_{T1}$	$0.444 \pm 0.002$
$f_{T2}$	$0.440 \pm 0.002$
$f_{I1}$	$0.603 \pm 0.005$
$f_{I2}$	$0.586 \pm 0.003$
$f'_{I1}$	$0.592 \pm 0.003$
$f'_{I2}$	$0.553 \pm 0.003$
$\xi_{Ijj}$	$1.040 \pm 0.003$

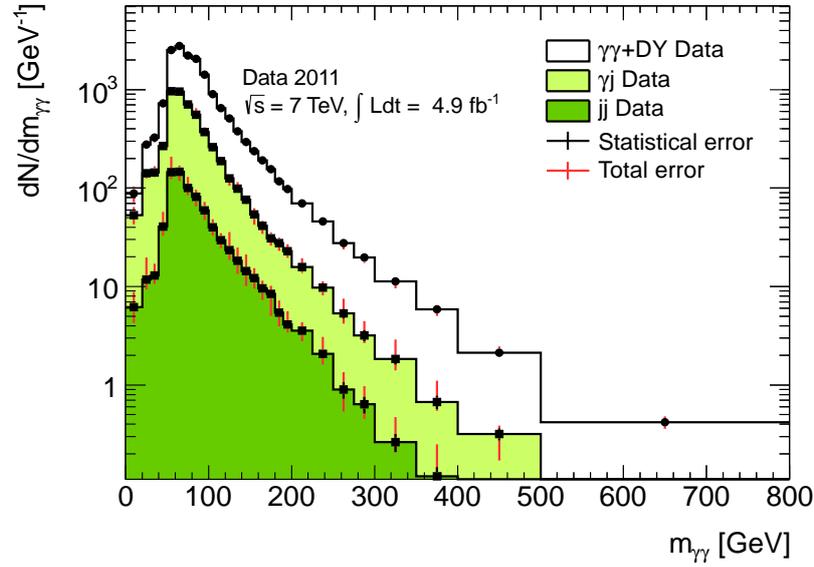
**Tableau 5.7** – Fake-rates et coefficient de corrélation entre l'isolation des jets obtenus par la méthode  $2 \times 2D$  sideband avec la sélection LOOSE' et avec en paramètres d'entrée les efficacités calculées à partir des événements diphoton générés avec SHERPA.

Les nombres d'événements  $\gamma$ - $\gamma$ ,  $\gamma$ -jet, jet- $\gamma$  et jet-jet dans la région du signal sont donnés tableau 5.8. Les erreurs systématiques totales correspondent à la somme quadratique des différents effets détaillés tableau 5.6.

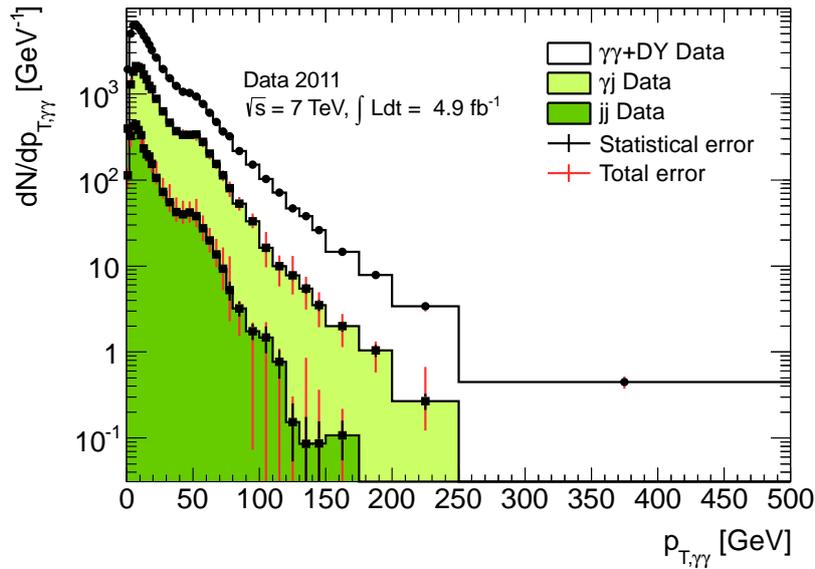
Nombre d'événements	
$W_{\gamma\gamma}^{TITI}$	$113156^{+629}_{-631}$ (stat) $^{+5040}_{-7980}$ (syst)
$W_{\gamma j}^{TITI}$	$31495^{+362}_{-359}$ (stat) $^{+3910}_{-3130}$ (syst)
$W_{j\gamma}^{TITI}$	$12978^{+255}_{-253}$ (stat) $^{+2450}_{-767}$ (syst)
$W_{jj}^{TITI}$	$8158^{+143}_{-141}$ (stat) $^{+1880}_{-1430}$ (syst)

**Tableau 5.8** – Nombres d'événements total par composante dans la région **TITI** obtenus avec la méthode  $2 \times 2D$  sideband. Les erreurs sont statistiques et systématiques.

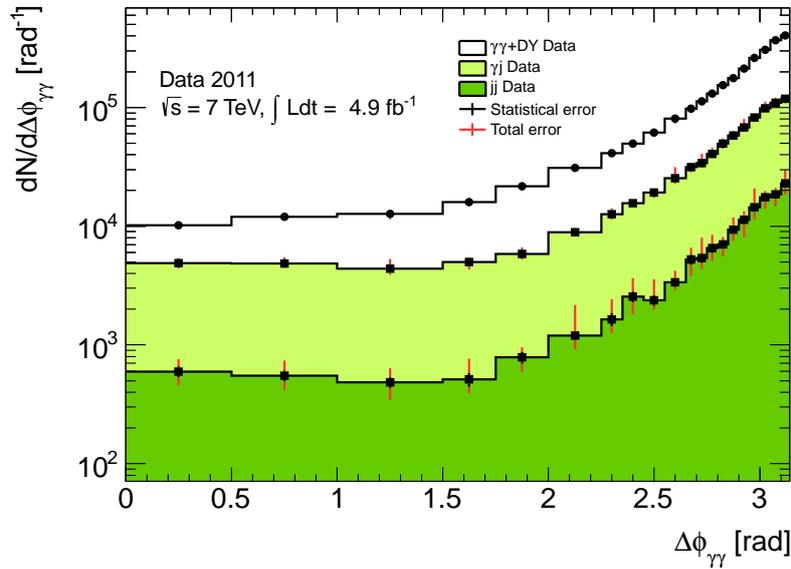
Les spectres différentiels du nombre d'événements sont présentés avec les erreurs statistiques et systématiques figures 5.24 à 5.27. Les distributions des variables cinématiques seront discutées dans la dernière partie de ce chapitre.



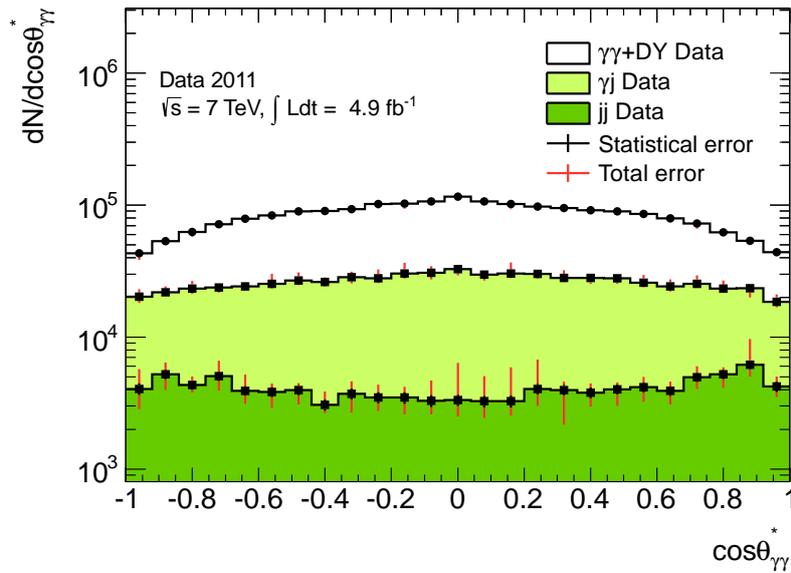
**Figure 5.24** – Spectre différentiel  $m_{\gamma\gamma}$  pour les contributions  $N_{\gamma\gamma}$ ,  $N_{\gamma j+j\gamma}$  et  $N_{jj}$ . Le nombre d'événements et les erreurs (statistiques+systematiques) sont normalisés à la largeur du bin.



**Figure 5.25** – Spectre différentiel  $p_{T,\gamma\gamma}$  pour les contributions  $N_{\gamma\gamma}$ ,  $N_{\gamma j+j\gamma}$  et  $N_{jj}$ . Le nombre d'événements et les erreurs (statistiques+systematiques) sont normalisés à la largeur du bin.



**Figure 5.26** – Spectre différentiel  $\Delta\phi_{\gamma\gamma}$  pour les contributions  $N_{\gamma\gamma}$ ,  $N_{\gamma j+j\gamma}$  et  $N_{jj}$ . Le nombre d'événements et les erreurs (statistiques+systematiques) sont normalisées à la largeur du bin.



**Figure 5.27** – Spectre différentiel  $\cos\theta^*$  pour les contributions  $N_{\gamma\gamma}$ ,  $N_{\gamma j+j\gamma}$  et  $N_{jj}$ . Le nombre d'événements et les erreurs (statistiques+systematiques) sont normalisées à la largeur du bin.

### 5.3 Estimation du bruit de fond électron

La méthode 2×2D sideband permet d'estimer la contamination du bruit de fond  $\gamma$ -jet et jet-jet. Basée sur l'isolation calorimétrique et l'identification des photons, elle permet de distinguer les événements caractérisés par une activité hadronique significative. Elle ne permet pas de distinguer les événements dans lesquels un électron isolé a été identifié comme un photon. Plusieurs processus contribuent à l'état final  $ee$  (Drell-Yan,  $Z \rightarrow ee$ ,  $WW \rightarrow e\bar{\nu}e\nu$ ) et  $e\gamma$  ( $\gamma W \rightarrow \gamma e\nu$ ,  $\gamma Z \rightarrow \gamma ee$ ). Au lieu d'estimer la contamination de chacun de ces bruits de fond dans la composante  $\gamma\gamma$  obtenue avec la méthode 2×2D sideband, une méthode générale basée sur le nombre d'événements  $N_{\gamma e}$  et  $N_{ee}$  est utilisée dans notre analyse.

#### 5.3.1 Description de la méthode

$N_e$  ( $N_\gamma$ ) désigne le nombre d'événements à l'état final avec un électron (photon) reconstruit passant les coupures d'identification TIGHT et d'isolation [103], soit l'ensemble **TI** (pour TIGHT et Isolé). Si on considère  $n_e$  ( $n_\gamma$ ) le nombre d'événements à l'état final avec un électron (photon),  $N_e$  ( $N_\gamma$ ) peut alors s'écrire :

$$\begin{pmatrix} N_\gamma \\ N_e \end{pmatrix} = \begin{pmatrix} \varepsilon_\gamma & \rho_{e \rightarrow \gamma} \\ \rho_{\gamma \rightarrow e} & \varepsilon_e \end{pmatrix} \times \begin{pmatrix} n_\gamma \\ n_e \end{pmatrix} \quad (5.9)$$

$\rho_{\gamma \rightarrow e}$  ( $\rho_{e \rightarrow \gamma}$ ) est la probabilité qu'un photon (électron) soit identifié comme un électron (photon) **TI**.  $\varepsilon_e$  ( $\varepsilon_\gamma$ ) est la probabilité qu'un électron (photon) soit reconstruit comme tel et identifié comme **TI**. Cette efficacité  $\varepsilon_e$  ( $\varepsilon_\gamma$ ) peut également s'écrire en fonction du nombre d'électrons (photons) identifiés comme tels et passant la sélection **TI**,  $N_\gamma^{sig}$  ( $N_e^{sig}$ ) :

$$N_e^{sig} = \varepsilon_e n_e \quad \text{et} \quad N_\gamma^{sig} = \varepsilon_\gamma n_\gamma \quad (5.10)$$

L'équation 5.9 devient :

$$\begin{pmatrix} N_\gamma \\ N_e \end{pmatrix} = \begin{pmatrix} 1 & f_{e \rightarrow \gamma} \\ f_{\gamma \rightarrow e} & 1 \end{pmatrix} \times \begin{pmatrix} N_\gamma^{sig} \\ N_e^{sig} \end{pmatrix} \quad (5.11)$$

$f_{\gamma \rightarrow e}$  ( $f_{e \rightarrow \gamma}$ ) correspondent aux ratios entre les objets faussement identifiés et ceux correctement identifiés.

$$f_{e \rightarrow \gamma} = \frac{\rho_{e \rightarrow \gamma}}{\varepsilon_e} = \frac{N_{e \rightarrow \gamma}}{N_{e \rightarrow e}} \quad (5.12)$$

$$f_{\gamma \rightarrow e} = \frac{\rho_{\gamma \rightarrow e}}{\varepsilon_\gamma} = \frac{N_{\gamma \rightarrow e}}{N_{\gamma \rightarrow \gamma}} \quad (5.13)$$

Le même raisonnement peut être appliqué pour le nombre d'événements à l'état final avec deux objets.  $N_{xy}$  désigne le nombre d'événements reconstruits comme  $xy$  passant les coupures d'identification et d'isolation pour les deux objets, soit **TITI**.  $N_{xy}^{sig}$  est le nombre d'événements pour lesquels cette classification est correcte. Les équations dans ce cas, se traduisent par la matrice :

$$\begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix} = \begin{pmatrix} 1 & f_{e \rightarrow \gamma} & (f_{e \rightarrow \gamma})^2 \\ 2f_{\gamma \rightarrow e} & (1 + f_{e \rightarrow \gamma}f_{\gamma \rightarrow e}) & 2f_{e \rightarrow \gamma} \\ (f_{\gamma \rightarrow e})^2 & f_{\gamma \rightarrow e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma}^{sig} \\ N_{\gamma e}^{sig} \\ N_{ee}^{sig} \end{pmatrix} \quad (5.14)$$

Elle est inversée afin d'accéder aux nombres d'événements de signal :

$$\begin{pmatrix} N_{\gamma\gamma}^{sig} \\ N_{\gamma e}^{sig} \\ N_{ee}^{sig} \end{pmatrix} = \frac{1}{(1 - f_{e\rightarrow\gamma}f_{\gamma\rightarrow e})^2} \begin{pmatrix} 1 & -f_{e\rightarrow\gamma} & (f_{e\rightarrow\gamma})^2 \\ -2f_{\gamma\rightarrow e} & (1 + f_{e\rightarrow\gamma}f_{\gamma\rightarrow e}) & -2f_{e\rightarrow\gamma} \\ (f_{\gamma\rightarrow e})^2 & -f_{\gamma\rightarrow e} & 1 \end{pmatrix} \times \begin{pmatrix} N_{\gamma\gamma} \\ N_{\gamma e} \\ N_{ee} \end{pmatrix} \quad (5.15)$$

Connaissant les  $f_{e\rightarrow\gamma}$  et  $f_{\gamma\rightarrow e}$ ,  $N_{\gamma\gamma}^{sig}$  peut être calculé à partir de  $N_{\gamma\gamma}$  :

$$N_{\gamma\gamma}^{sig} = \frac{N_{\gamma\gamma} - [f_{e\rightarrow\gamma}N_{\gamma e} - (f_{e\rightarrow\gamma})^2N_{ee}]}{(1 - f_{e\rightarrow\gamma}f_{\gamma\rightarrow e})^2} \quad (5.16)$$

$f_{e\rightarrow\gamma}$  et  $f_{\gamma\rightarrow e}$  dépendent de plusieurs effets expérimentaux comme la quantité de matière et les performances du trajectographe dont la description dans la simulation est limitée. Ces paramètres sont extraits directement des données. On considère la région de masse invariante proche du pic du  $Z$ , dans laquelle l'état final  $ee$  est dominant par rapport à  $e\gamma$  et  $\gamma\gamma$ . Le bruit de fond dans cette région est soustrait en utilisant les side-bands.

A partir de la désintégration du  $Z$  en  $ee$  et  $e\gamma$ , on peut estimer  $f_{e\rightarrow\gamma}$  :

$$\begin{aligned} N_{ee} &= N_{ee}^{sig} \\ N_{\gamma e} &= N_{ee}^{sig} \cdot 2f_{e\rightarrow\gamma} \end{aligned}$$

d'où :

$$f_{e\rightarrow\gamma} = \frac{N_{\gamma e}}{2N_{ee}} \quad (5.17)$$

De même, à partir de la désintégration radiative du  $Z$  en  $ee\gamma$ , et en comptant le nombre d'événements  $eee$  et  $ee\gamma$  identifiés dans les données on peut estimer  $f_{\gamma\rightarrow e}$  :

$$\begin{aligned} N_{ee\gamma} &= N_{ee\gamma}^{sig} \\ N_{eee} &= N_{ee\gamma}^{sig} \cdot f_{\gamma\rightarrow e} \end{aligned}$$

d'où :

$$f_{\gamma\rightarrow e} = \frac{N_{eee}}{N_{ee\gamma}} \quad (5.18)$$

Pour mesurer les sections efficaces différentielles, le nombre d'événements issus du bruit de fond des électrons isolés doit être calculé pour chaque bin des distributions des variables cinématiques choisies ( $m_{\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\cos\theta^*$ ). Les histogrammes  $h_{\gamma\gamma}$ ,  $h_{\gamma e}$  sont remplis respectivement pour les événements  $\gamma\text{-}\gamma$ ,  $\gamma\text{-}e$  et  $e\text{-}\gamma$  passant la sélection **TITI**. Pour chaque histogramme ainsi rempli et à partir de l'équation 5.16 l'impureté est calculée par bin selon :

$$I_e[bin] = 1 - \frac{h_{\gamma\gamma}^{sig}[bin]}{h_{\gamma\gamma}[bin]} \quad (5.19)$$

$$= \frac{f_{e\rightarrow\gamma} \frac{h_{\gamma e}[bin]}{h_{\gamma\gamma}[bin]} - (f_{e\rightarrow\gamma})^2 \frac{h_{ee}[bin]}{h_{\gamma\gamma}[bin]}}{(1 - f_{e\rightarrow\gamma}f_{\gamma\rightarrow e})^2} - \frac{2f_{e\rightarrow\gamma}f_{\gamma\rightarrow e} - (f_{e\rightarrow\gamma}f_{\gamma\rightarrow e})^2}{(1 - f_{e\rightarrow\gamma}f_{\gamma\rightarrow e})^2} \quad (5.20)$$

L'histogramme "signal" est ensuite obtenu avec :

$$h_{\gamma\gamma}^{sig}[bin] = h_{\gamma\gamma}[bin](1 - I_e[bin]) \quad (5.21)$$

### 5.3.2 Extraction de la contamination électron

Les électrons doivent passer la sélection TIGHT et le même critère d'isolation que les photons  $E_T^{iso} < 4$  GeV. En outre, une coupure sur la distance entre les coordonnées des clusters électron et photon est appliquée afin de supprimer la reconstruction simultanée d'un objet comme électron et photon.

$$\Delta R = \sqrt{(\eta_e - \eta_\gamma)^2 + (\phi_e - \phi_\gamma)^2} < 0.05$$

$N_{\gamma e}$  et  $N_{ee}$  (nécessaires à l'estimation de  $f_{e \rightarrow \gamma}$ ) ainsi que  $N_{eee}$  et  $N_{ee\gamma}$  (nécessaires à l'estimation de  $f_{\gamma \rightarrow e}$ ) sont extraits des données dans une fenêtre de masse invariante entourant le pic du  $Z$ . Pour chaque distribution, le pic du boson  $Z$  dans la distribution de la masse invariante di-électron est corrigé et fitté avec une fonction gaussienne de largeur  $\sigma$ . Les événements considérés sont compris dans une fenêtre de  $\pm 1\sigma$ . Les régions de sideband utilisées pour l'estimation du bruit de fond sont définies symétriquement par rapport au pic du  $Z$  : 51-61 GeV et 121-131 GeV. Les erreurs systématiques liées aux choix de la fenêtre et des régions de sideband sont calculées en modifiant la fenêtre de masse de  $\pm 1$  et  $\pm 2\sigma$ , et en décalant les régions de sideband de  $\pm 5$  GeV.

La méthode est appliquée sur les événements  $Z \rightarrow ee$  générés par PYTHIA qui incluent également les événements  $Z \rightarrow ee\gamma$ . Les résultats sont présentés tableaux 5.9 et 5.10. L'accord est satisfaisant entre la valeur extraite du Monte-Carlo au niveau détecteur et au niveau particule pour  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$ , ce qui permet de valider la méthode. Toutefois des différences de l'ordre de 1 et  $2\sigma$  sont observées entre les valeurs Monte-Carlo et les données pour respectivement  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$ .

	$N_{\gamma e}$	$N_{ee}$	$f_{e \rightarrow \gamma}$	$f_{e \rightarrow \gamma}$ vérité
$Z \rightarrow ee$ MC	165253	1284798	$0.0643 \pm 0.0001(\text{stat})_{-0.0048}^{+0.0526}(\text{syst})$	$0.0643 \pm 0.0002$
données	79041	640559	$0.0617 \pm 0.0002(\text{stat})_{-0.0100}^{+0.0400}(\text{syst})$	

**Tableau 5.9** – Valeurs de  $f_{e \rightarrow \gamma}$  estimées dans les données et dans le Monte-Carlo, comparées à la valeur  $f_{e \rightarrow \gamma}$  au niveau vérité [117].

	$N_{eee}$	$N_{ee\gamma}$	$f_{\gamma \rightarrow e}$	$f_{\gamma \rightarrow e}$ vérité
$Z \rightarrow ee\gamma$ MC	49.4	1849.9	$0.027 \pm 0.004(\text{stat})_{-0.002}^{+0.021}(\text{syst})$	$0.0267 \pm 0.0039$
données	81.9	2138.3	$0.038 \pm 0.004(\text{stat})_{-0.006}^{+0.024}(\text{syst})$	

**Tableau 5.10** – Valeurs de  $f_{\gamma \rightarrow e}$  estimées dans les données et dans le Monte-Carlo, comparées à la valeur  $f_{\gamma \rightarrow e}$  au niveau vérité [117].

L'impureté bin par bin est estimée à partir de l'équation 5.20 pour les quatre distributions cinématiques considérées,  $m_{\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\cos\theta^*$ . Les valeurs de  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$  sont fixées et utilisées pour l'estimation de l'impureté dans chaque bin. La principale source d'erreur systématique vient de la dépendance en  $E_T$  de ces coefficients  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$ . Les erreurs systématiques associées ont été évaluées en faisant varier les valeurs de  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$  de  $\pm 1\sigma$ . Par ailleurs, l'erreur totale sur les valeurs de  $f_{e \rightarrow \gamma}$  et  $f_{\gamma \rightarrow e}$  est également considérée comme une source d'erreur systématique dans l'estimation de l'impureté.

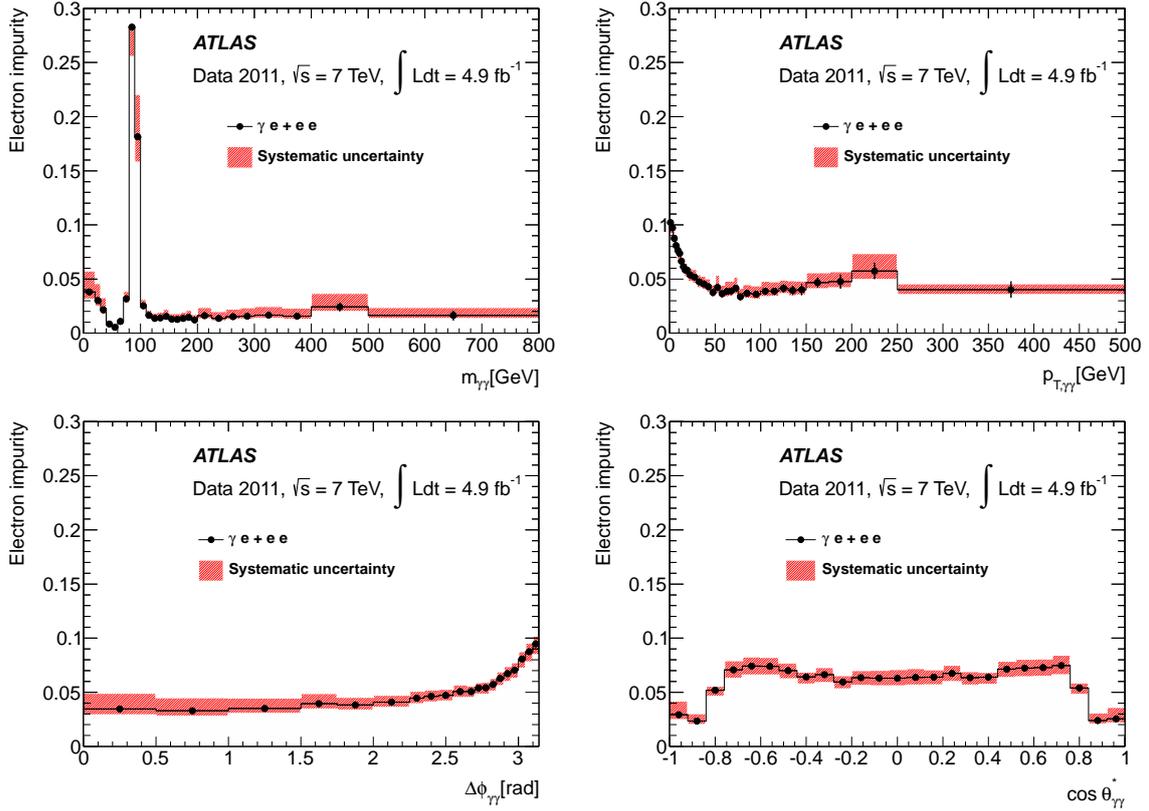


Figure 5.28 – Impureté bin par bin pour  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$ ,  $\cos\theta_{\gamma\gamma}^*$  [99].

## 5.4 Déconvolution

La déconvolution est la dernière étape dans la mesure de section efficace. A partir du nombre d'événements observés, elle permet de remonter au nombre d'événements réellement produits.

### 5.4.1 Description de la méthode

La section efficace différentielle en fonction de la variable  $X$ , dans le bin  $i$  s'écrit :

$$\frac{d\sigma_i}{dX^{\text{true}}} = \frac{N_i^{\gamma\gamma,\text{true}}}{\int \mathcal{L} dt \Delta X_i^{\text{true}}} \quad (5.22)$$

$\int \mathcal{L} dt$  est la luminosité intégrée et  $\Delta X_i$  est la largeur du  $i^{\text{ème}}$  bin de la variable  $X$ . La déconvolution permet de faire le lien entre le vrai nombre d'événements avec une paire de photons  $N_i^{\gamma\gamma,\text{true}}$ , dont la quantité  $X$  dans le bin  $i$  a pour vraie valeur  $X_i^{\text{true}}$ , et le nombre d'événements diphoton passant les coupures de sélection et reconstruits dans le bin  $j$  de la quantité  $X^{\text{reco}}$ . La formule utilisée est :

$$N_i^{\gamma\gamma,\text{select}} = \varepsilon_{\text{trig}} \sum_j \Lambda_{ji} N_j^{\gamma\gamma,\text{true}} \quad (5.23)$$

$\varepsilon$  est l'efficacité du trigger `2g20_loose` et les  $\Lambda_{ji}$  sont les coefficients de déconvolution. Les nombres d'événements véritables peuvent donc être écrits comme une combinaison linéaire des nombres d'événements reconstruits.

L'efficacité de trigger est définie comme le ratio des événements diphoton passant le trigger sur le nombre total d'événements diphoton. Cette efficacité est estimée avec une technique *bootstrap* [119] à partir d'un ensemble de données sélectionnées par le plus bas trigger non biaisé et en prenant en compte les corrélations cinématiques entre les deux candidats photons. Elle est

de  $97.8^{+0.8}_{-1.5}(\text{stat})\%$  et elle est estimée constante en fonction des quatre variables cinématiques considérées.

Les coefficients de déconvolution correspondent aux éléments des matrices d'efficacité de sélection. Pour chacune des quatre variables cinématiques étudiées  $X = m_{\gamma\gamma}, p_{T,\gamma\gamma}, \Delta\phi_{\gamma\gamma}, \cos\theta^*$ , les matrices d'efficacité de sélection sont construites avec en abscisse la quantité reconstruite et en ordonnée la quantité au niveau particule. Les éléments de matrice  $\Lambda_{ij}$  prennent en compte les efficacités de reconstruction et de sélection ainsi que les migrations entre les bins liées aux effets de résolution du détecteur. Les coupures appliquées au niveau particule et au niveau générateur sont celles appliquées pour l'analyse de données. L'énergie d'isolation au niveau particule correspond à l'énergie transverse totale de toutes les particules contenues dans un cône de rayon 0.4. Les corrections liées à l'empilement et aux événements sous-jacent sont calculées de la même façon au niveau vérité que sur les objets reconstruits (densité d'énergie ambiante).

Les matrices d'efficacité obtenues pour les quatre variables cinématiques à partir des événements diphoton générés par SHERPA sont présentées figure 5.29. Ces matrices sont presque diagonales, ceci témoigne de la bonne résolution du détecteur par rapport à la largeur des bins et justifie le choix de la méthode de déconvolution "bin par bin". Dans cette méthode, les coefficients de déconvolution correspondent simplement à l'inverse des efficacités définies par l'équation :

$$\varepsilon(X^{\text{true}}) \equiv \frac{dN^{\gamma\gamma}/dX^{\text{select}}}{dN^{\gamma\gamma}/dX^{\text{true}}}, \quad (5.24)$$

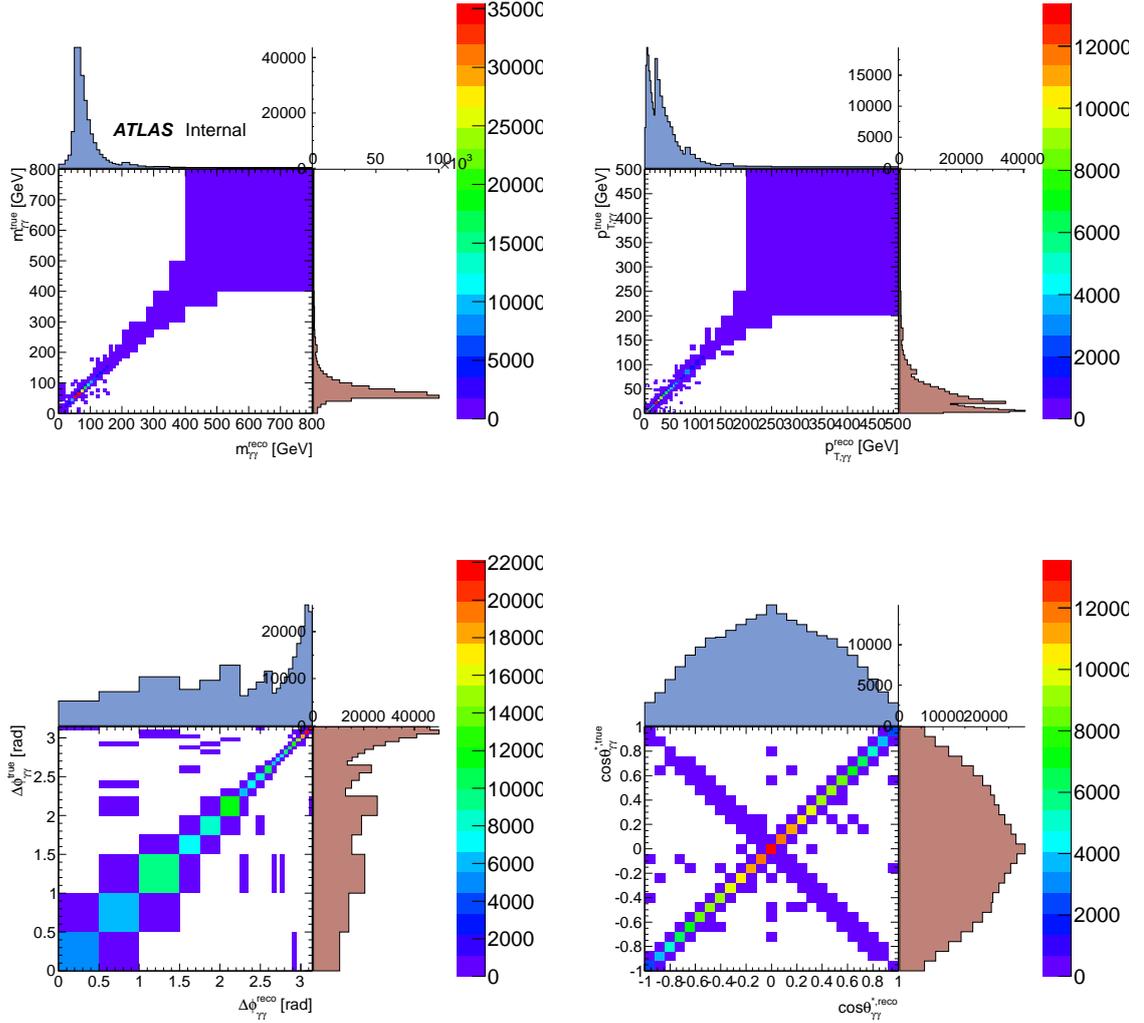
Les migrations bin à bin sont considérées négligeables ; la matrice se réduit alors à sa diagonale.

Un test de fermeture a été réalisé avec les générateurs SHERPA et PYTHIA afin de juger d'un biais inhérent à la méthode. Les coefficients de déconvolution sont calculés à partir de la simulation SHERPA et la déconvolution est appliquée aux variables reconstruites de la simulation PYTHIA. Les variables reconstruites après déconvolution sont comparées aux distributions des variables au niveau particule pour les mêmes événements PYTHIA. Un léger biais est observé et une erreur systématique de 1% sur la section efficace finale est associée.

Deux autres méthodes sont utilisées pour vérification : la méthode itérative, basée sur la probabilité conditionnelle qu'un événement reconstruit dans le bin " $E_j$ " corresponde à un événement du bin " $C_i$ " de la distribution véritable, et la méthode *TUnfold* basée sur l'utilisation d'un fit. Elles ne seront pas détaillées ici, mais plus d'explications peuvent être trouvées respectivement dans [120] et [121]. Les trois méthodes de déconvolution sont appliquées aux données et comparées. Aucun biais significatif n'est observé, aucune incertitude systématique n'est donc associée au choix de la méthode de déconvolution.

#### 5.4.1.1 Incertitudes systématiques

Les coefficients de déconvolution par bin sont calculés à partir de l'échantillon diphoton SHERPA et incluent efficacité de reconstruction et efficacité de sélection. Les erreurs systématiques sur l'efficacité de reconstruction sont évaluées à partir d'événements diphoton obtenus avec la simulation PYTHIA nominale et la simulation PYTHIA avec une géométrie modifiée. Les effets pris en compte sont : la quantité de matière en amont du calorimètre, la proportion de photons issus du processus de fragmentation et l'usage de différentes fonctions de structure. L'impact estimé sur la section efficace différentielle est compris entre  $^{+2}_{-5}\%$  pour  $m_{\gamma\gamma}$ , entre  $\pm 3\%$  pour  $p_{T,\gamma\gamma}$ , entre  $^{+3}_{-4}\%$  pour  $\Delta\phi_{\gamma\gamma}$  et entre  $^{+2}_{-3}\%$  pour  $\cos\theta^*$ . L'incertitude liée aux efficacités d'identification est estimée en variant les efficacités d'identification dans l'intervalle déterminé par leurs erreurs (évaluées dans [108]). Les incertitudes liées à l'échelle d'énergie des photons et à la résolution sont prises en compte de la même façon. L'effet sur la section efficace différentielle est typiquement de l'ordre de  $^{+1}_{-2}\%$ . Les incertitudes liées à l'empilement et au décalage



**Figure 5.29** – Matrices d’efficacité de sélection obtenues à partir de SHERPA pour :  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$  et  $\cos\theta^*$  [117].

entre données et Monte-Carlo pour la distribution de la variable d’isolation sont considérées comme négligeables.

Afin d’estimer l’incertitude sur l’efficacité de trigger, celle-ci est calculée à partir des échantillons Monte-Carlo en appliquant la technique de bootstrap au trigger photon simple, négligeant ainsi les corrélations entre les deux candidats photons. L’erreur systématique est extraite des différentes valeurs obtenues, elle est de  $\pm 0.8\%$ , on a donc  $\varepsilon_{trig} = 97.8^{+0.8}_{-1.5}(\text{stat}) \pm 0.8(\text{syst})\%$ .

### 5.4.2 Efficacité inclusive

Outre les efficacités de sélection par bin pour la mesure de sections efficaces différentielles, l’efficacité de sélection inclusive est nécessaire pour la mesure de section efficace totale. La section efficace totale est calculée à partir du nombre total d’événements  $\gamma\text{-}\gamma$ ,  $N_{\gamma\gamma}$ , obtenu après correction du bruit de fond  $\gamma\text{-jet}$ ,  $\text{jet-jet}$  et électron :

$$\sigma(pp \rightarrow X \rightarrow \gamma\gamma) = \frac{N_{\gamma\gamma}}{\bar{\varepsilon}_{sel}\varepsilon_{trig} \int \mathcal{L} dt} \quad (5.25)$$

$\bar{\varepsilon}_{sel}$  est l'efficacité moyenne de sélection, elle est calculée à partir de l'échantillon diphoton SHERPA qui a été pondéré afin de mieux décrire les données. Pour chaque distribution cinématique, le profil SHERPA est corrigé pour correspondre au spectre  $\gamma\text{-}\gamma$  extrait des données (après soustraction des bruits de fond). L'efficacité de sélection est calculée pour chaque variable. Les efficacités obtenues sont compatibles entre elles et leur valeur moyenne est  $\bar{\varepsilon}_{sel}=49.6^{+1.9}_{-1.7}\%$ . Les erreurs systématiques principales sur cette estimation viennent de l'incertitude sur l'efficacité d'identification  $\pm 1.2\%$ , de l'incertitude sur l'échelle d'énergie  $^{+1.2}_{-0.5}\%$  et du choix de l'échantillon de simulation  $\pm 0.9\%$ .

## 5.5 Sections efficaces

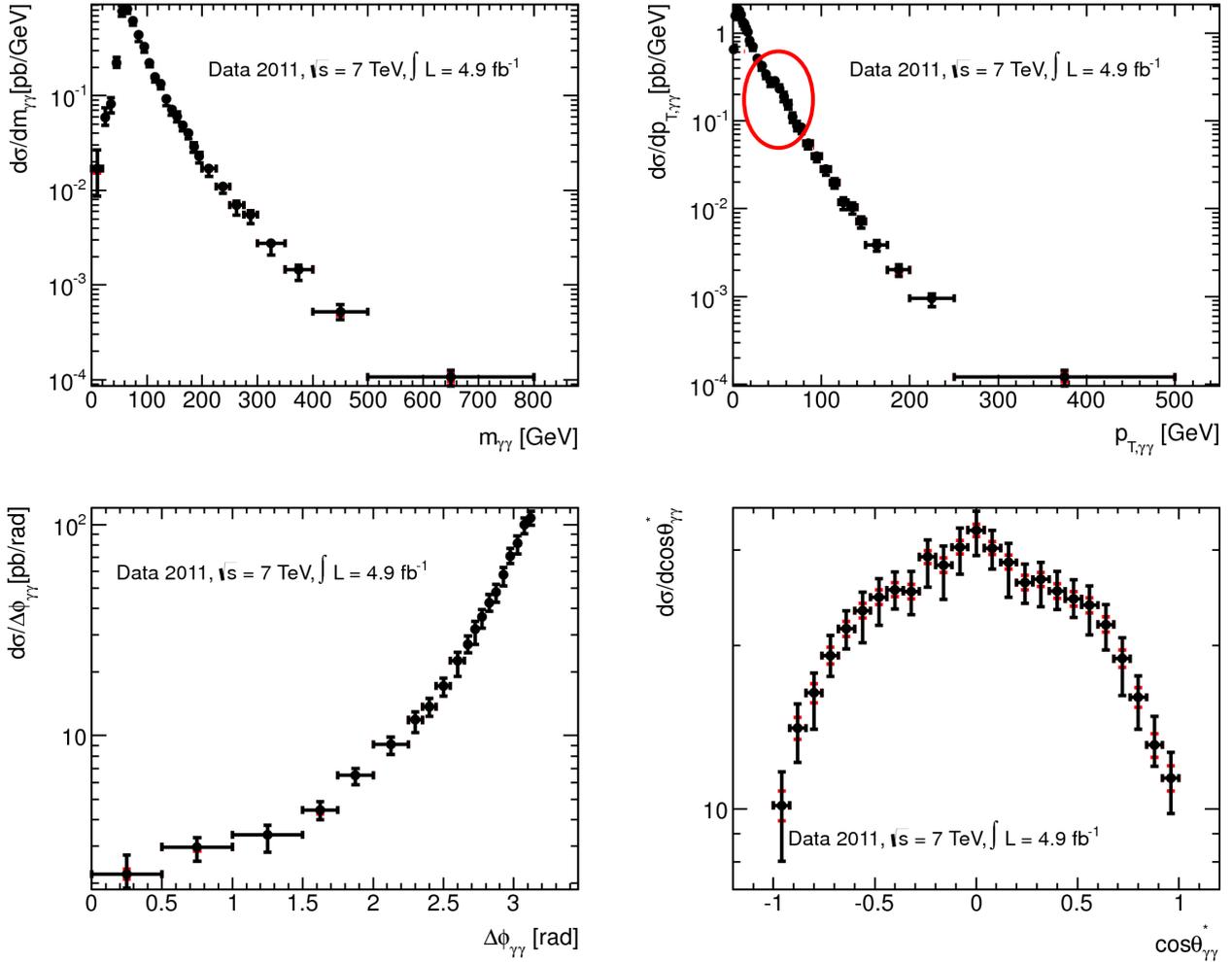
### 5.5.1 Sections efficaces mesurées

Le nombre total d'événements  $N_{\gamma\gamma}$  est divisé par la luminosité intégrée, l'efficacité de trigger et l'efficacité  $\bar{\varepsilon}_{sel}$  définie auparavant (équation 5.25). La luminosité intégrée étant de  $4.9\pm 0.2 \text{ fb}^{-1}$ , la section efficace totale mesurée est :

$$\sigma(pp \rightarrow X \rightarrow \gamma\gamma) = 44.0^{+3.2}_{-4.2} \text{ pb} \quad (5.26)$$

Les sources d'erreurs systématiques principales sur cette mesure sont l'efficacité de sélection et la soustraction du bruit de fond  $\gamma\text{-jet}$ ,  $jet\text{-jet}$ .

Les distributions du nombre d'événements  $\gamma\gamma$  obtenues après déconvolution (incluant la correction pour l'efficacité de trigger) sont divisées par la luminosité intégrée et la largeur du bin. Les sections efficaces différentielles obtenues en fonction de  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$  et  $\cos\theta^*$  sont présentées figures 5.30, l'erreur totale obtenue sur la mesure est dominée par l'erreur systématique dans tout l'espace des phases. La structure observée dans le spectre en  $p_{T,\gamma\gamma}$  correspond à un effet dénommé *Guillet shoulder* et documenté dans [122]. Elle est liée à l'apparition de la contribution NLO qui commence pour  $\Delta\phi_{\gamma\gamma} < \pi/2$  et qui correspond à  $p_{T,\gamma\gamma} > p_{T,\gamma\gamma}^{min}$  avec  $p_{T,\gamma\gamma}^{min} = \sqrt{p_{T,\gamma 1}^2 + p_{T,\gamma 2}^2} \simeq 33 \text{ GeV}$ .



**Figure 5.30** – Sections efficaces différentielles mesurées en fonction de  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$  et  $\cos\theta^*$ . Les événements  $\gamma\gamma$  sont obtenus à partir de la méthode 2×2D sideband et corrigés pour la contamination des électrons avant d’être déconvolués. L’effet encerclé en rouge correspond à la *Guillet shoulder* [122]. L’erreur totale correspond à la somme en quadrature des erreurs statistiques et systématiques. L’erreur statistique est représentée par les barres rouges.

## 5.5.2 Prédictions théoriques

Les deux types de prédictions théoriques (voir chapitre 1), les prédictions à ordre fixe et les générateurs Monte-Carlo avec parton shower sont ici considérées.

La contribution du boson de Higgs du Modèle Standard (pour  $m_H = 125$  GeV), n'est pas inclut dans les prédictions. Elle est estimée à environ 1% dans la fenêtre de masse 120-130 GeV et est négligeable ailleurs.

### 5.5.2.1 Prédictions à ordre fixe

Dans DIPHOX, à l'exception du processus de boîte, les contributions de tous les processus (Born, Bremsstrahlung, fragmentations simple et double) sont calculées à NLO. Afin d'obtenir un calcul complet à NLO, la contribution du processus de boîte (à NLO) a été extraite à partir de Gamma2MC et incluse dans nos prédictions. A proprement parlé, le processus de boîte est à NNLO, mais la forte densité de gluons au LHC rend sa contribution comparable à une contribution à LO [26].

La sélection appliquée au niveau partonique dans DIPHOX est équivalente à celle appliquée dans l'analyse :

- Acceptance : Pour les deux photons  $|\eta| \leq 1.37$  ou  $1.52 \leq |\eta| \leq 2.37$  et  $p_T^{\gamma 1} > 25$  GeV,  $p_T^{\gamma 2} > 22$  GeV.
- Distance entre les deux photons :  $\Delta R = \sqrt{(\eta^{\gamma 1} - \eta^{\gamma 2})^2 + (\phi^{\gamma 1} - \phi^{\gamma 2})^2} > 0.4$ .
- Isolation : L'énergie partonique déposée dans un cône de rayon  $R=0.4$  autour du photon ne doit pas excéder 4 GeV et toute activité associée au processus dur est exclue.

La figure 5.31 montre les contributions des composantes directes, de fragmentation simple et de fragmentation double dans les prédictions DIPHOX. Seule la somme des différentes contributions est un observable physique (la dépendance en  $\mu_f$ , l'échelle de fragmentation, s'annule seulement quand les trois contributions sont sommées [26]), il est cependant intéressant de distinguer les différentes composantes pour pouvoir les comparer. Les contributions des simple et double fragmentations après la coupure en isolation sont respectivement quatre fois et vingt fois plus petites que celles des processus directs.

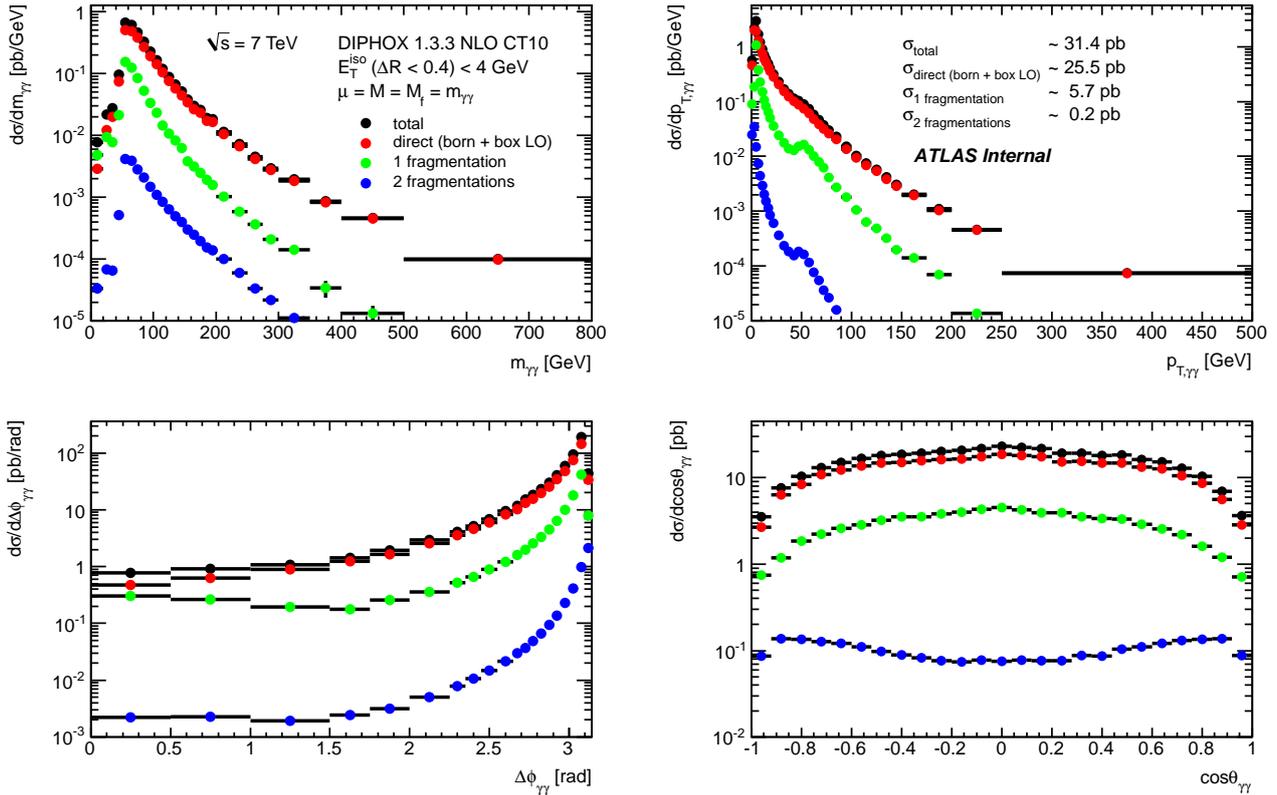
La figure 5.32 montre les sections efficaces différentielles obtenues avec DIPHOX+Gamma2MC et les PDFs CT10 NLO [123], [124]. Les échelles de renormalisation  $\mu$ , de factorisation  $M$  et de fragmentation  $\mu_f$  sont fixées égales à  $m_{\gamma\gamma}$ . Les erreurs d'échelle sont évaluées en variant les trois échelles entre  $0.5 m_{\gamma\gamma}$  et  $2 m_{\gamma\gamma}$  [73]. L'enveloppe de toutes les variations correspond à l'erreur systématique mais l'erreur finale est dominée par la configuration pour laquelle les échelles sont modifiées de façon incohérente. Les prédictions sur la section efficace varient ainsi de 5 à 45% en fonction de  $m_{\gamma\gamma}$ , de 5 à 25% en fonction de  $p_{T,\gamma\gamma}$ , de 15 à 25% en fonction de  $\Delta\phi_{\gamma\gamma}$  et de 5 à 25% en fonction de  $\cos\theta^*$ .

Les erreurs liées aux PDFs ont été estimées en variant de  $\pm 1\sigma$  chacune des 20 valeurs propres, les variations positives et négatives de la section efficace sont ensuite sommées quadratiquement de façon indépendante. Cette méthode est appliquée indépendamment pour chaque bin des distributions de la figure 5.32. Les variations correspondantes sur la section efficace sont de 2% à 5% pour  $m_{\gamma\gamma}$ , de 2% à 5% pour  $p_{T,\gamma\gamma}$ , de 2% à 3% pour  $\Delta\phi_{\gamma\gamma}$ , de 3% à 4% pour  $\cos\theta^*$ .

Les prédictions NNLO ont été fournies par les auteurs de 2 $\gamma$ NNLO [125]. Les coupures d'acceptance appliquées au niveau générateur sont les mêmes que celles appliquées pour l'analyse des données. Pour l'isolation, une coupure équivalente ne peut pas être appliquée. L'isolation implémentée dans 2 $\gamma$ NNLO est l'isolation Frixione [126] définie telle que :

$$E_T^{iso}(R) < E_T^{max} \chi(R), \quad \chi(R) = \left( \frac{1 - \cos R}{1 - \cos R_0} \right)^n \quad (5.27)$$

Avec  $R_0 = 0.4$  et  $E_T^{max} = 4$  GeV. Le critère d'isolation doit être respecté pour tout  $R < R_0$  (plus le parton est proche du photon, plus la coupure sur l'énergie d'isolation est stricte). Il s'agit



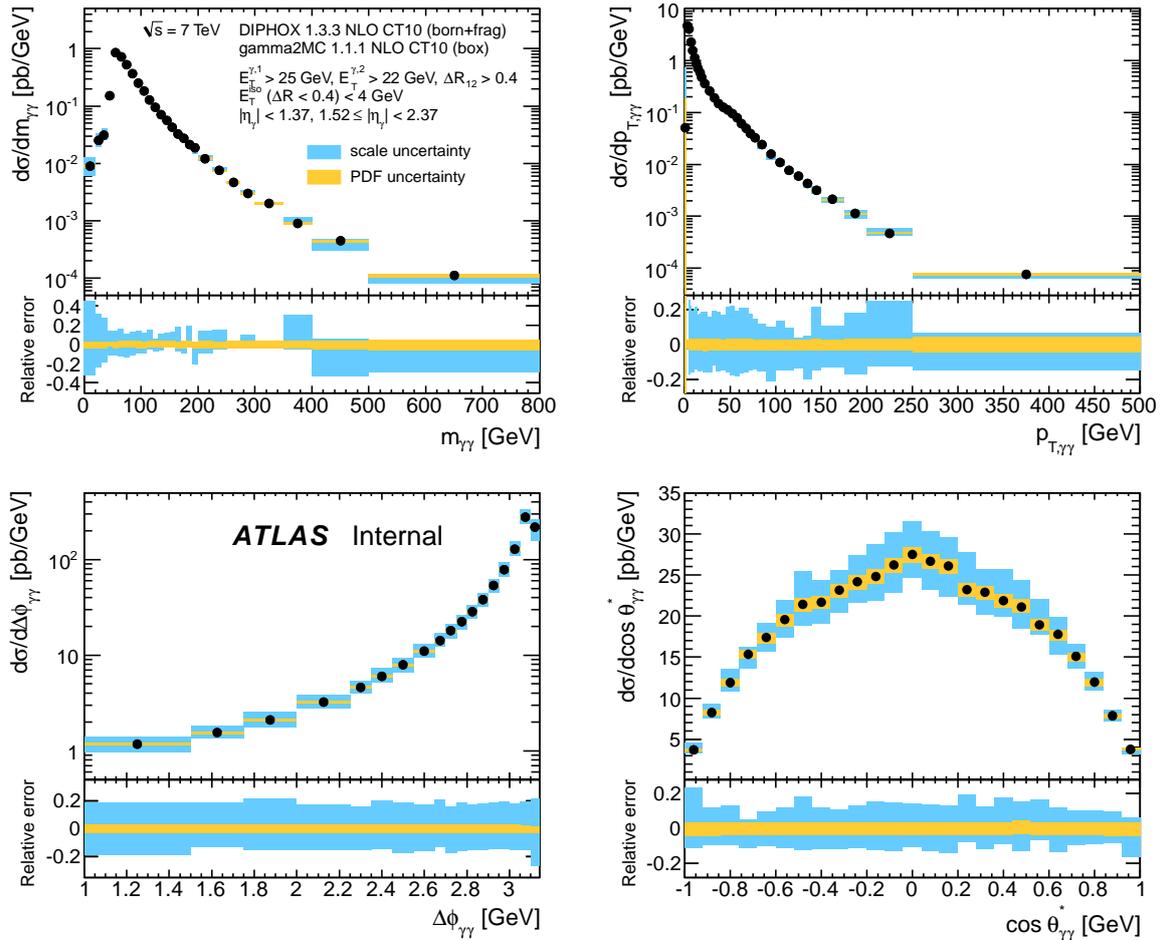
**Figure 5.31** – Contributions directes et de fragmentation obtenues avec DIPHGX et les PDFs CT10. Les échelles sont fixées à  $m_{\gamma\gamma}$ . Les erreurs sont statistiques uniquement [117].

d'un critère continu qui élimine complètement les composantes de fragmentation. Toutefois, les auteurs ont constaté que la section efficace obtenue à NLO avec  $2\gamma$ NNLO pour  $n = 1$  était une bonne approximation de la section efficace obtenue avec DIPHGX pour le même rayon et énergie maximale. Pour  $1 < E_T^{iso} < 6$  GeV, la section efficace calculée par  $2\gamma$ NNLO sous-estime au plus de quelques % la section efficace calculée par DIPHGX (comprenant donc composante directe et fragmentations). Par ailleurs, il est intéressant de noter qu'une version discrétisée du critère d'isolation Frixione, plus proche de l'approche expérimentale, a aussi été développée. La comparaison entre le critère discret et continu de l'isolation Frixione est discutée en annexe D [127]. Les PDFs MSTW2008 sont utilisés pour nos prédictions, mais les erreurs liées aux PDFs n'ont pas été fournies par les auteurs. Suivant leurs recommandations, les erreurs calculées pour DIPHGX à NLO+LO pour les PDF MSTW2008 sont appliquées à cette prédiction NNLO, ce qui est extrêmement conservatif.

Comme pour DIPHGX, les échelles ont été fixées égales à  $m_{\gamma\gamma}$ . L'incertitude liée à ce choix est évaluée en considérant deux cas :  $\mu = m_{\gamma\gamma}/2$ ,  $M = 2m_{\gamma\gamma}$  et  $\mu = 2m_{\gamma\gamma}$ ,  $M = m_{\gamma\gamma}/2$ .

### 5.5.2.2 Générateur Monte-Carlo avec parton shower

Dans le cas des générateurs Monte-Carlo avec parton shower, les sections efficaces sont obtenues à l'aide d'un programme RIVET [128]. Pour chaque événement de la simulation, le programme RIVET analyse l'événement au niveau vérité, les coupures appliquées sont les mêmes que celles appliquées pour l'analyse de données. Pour l'isolation au niveau particule, l'énergie transverse totale de toutes les particules contenues dans un cône de rayon 0.4 autour du photon doit être inférieure à 4 GeV (comme mentionné section 5.4.1). Les corrections sur l'énergie d'isolation liées à l'empilement et aux événements sous-jacent sont calculées de la même façon que sur les objets reconstruits (densité d'énergie ambiante).



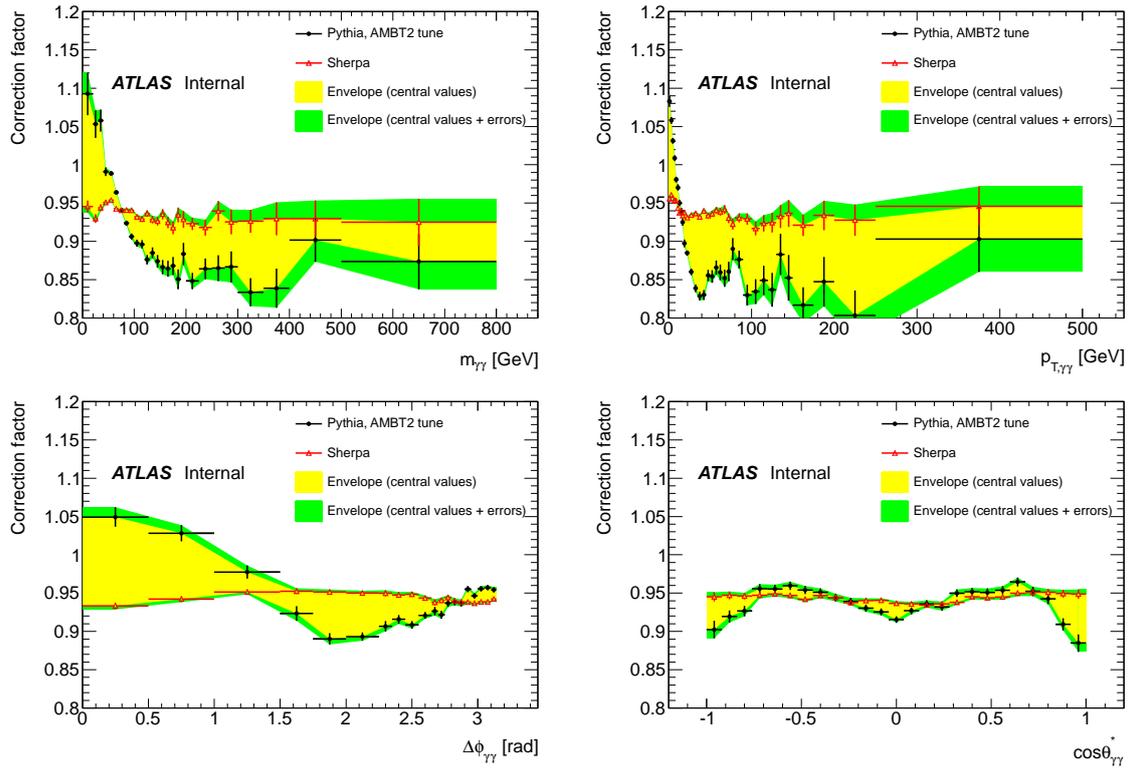
**Figure 5.32** – Sections efficaces différentielles prédites par *DIPHOX* +*Gamma2MC* pour les PDFs *CT10*. Toutes les échelles sont fixées à  $m_{\gamma\gamma}$  [117].

*PYTHIA* est un générateur LO avec parton shower. Les calculs des éléments de matrice sont fait à LO, les termes d'ordre plus élevé sont simulés par la production  $\gamma$ -jet et di-jet en combinaison avec ISR et/ou FSR. Un modèle pour les événements sous-jacent et l'hadronisation sont également inclus. La configuration de *PYTHIA* utilisé par ATLAS comprend les PDF *MRST2007* et comme pour tous les générateurs avec parton shower, les événements générés sont injectés dans le programme *GEANT4* [129] pour la simulation détecteur. La section efficace totale de production de paires de photons estimée par *PYTHIA* avoisine les 36 pb.

*SHERPA* est également un générateur Monte-Carlo (LO) avec parton shower. La fragmentation est entièrement traitée par la parton shower. A la différence de *PYTHIA*, *SHERPA* inclut des émissions réelles d'ordre supérieur au niveau des éléments de matrice [130]. Jusqu'à deux partons additionnels peuvent être produits avec la paire de photons. La section efficace totale obtenue avec *SHERPA* est de 36 pb.

Le processus d'hadronisation ainsi que les événements d'empilement et les événements sous-jacent sont modélisés dans les simulations *SHERPA* et *PYTHIA* mais ne le sont pas dans les prédictions partoniques de  $2\gamma$ NNLO et *DIPHOX*. Dans les données, l'isolation est corrigée pour la plupart de ces phénomènes avec l'estimation de l'énergie ambiante, mais il n'est pas certain que ces corrections soient suffisantes pour rendre équivalentes l'isolation expérimentale et l'isolation partonique. Les simulations *SHERPA* et *PYTHIA* sont comparés avec et sans hadronisation et événements sous-jacent. On extrait de cette comparaison des facteurs correctifs (figures 5.33)

qui sont ensuite appliqués aux prédictions à ordre fixe.



**Figure 5.33** – Facteurs correctifs pour les phénomènes non-perturbatifs extraits des simulation *SHERPA* et *PYTHIA* avec et sans hadronisation et underlying events et appliqués aux prédictions à ordre fixe  $2\gamma$ NNLO et *DIPHOX* [117].

### 5.5.3 Comparaison entre prédictions théoriques et données

La section efficace totale mesurée avec  $4.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$  peut être comparée aux sections efficaces prédites par les calculs à ordre fixe, *DIPHOX*+*Gamma2MC* et  $2\gamma$ NNLO, les résultats sont présentés tableau 5.11. Pour *DIPHOX*+*Gamma2MC* comme pour  $2\gamma$ NNLO, l'erreur est dominée par l'erreur d'échelle. Il est attendu que *PYTHIA* et *SHERPA* (LO + parton shower) prédisent une section efficace plus petite que celle mesurée ( $\sim 36 \text{ pb}$ ). Leurs sections efficaces différentielles sont donc pondérées afin de correspondre à la valeur centrale de la section efficace totale mesurée dans les données (facteur 1.2).

	Section efficace totale [pb]
données	$44.0^{+3.2}_{-4.2}$
<i>DIPHOX</i> + <i>Gamma2MC</i>	$39^{+7}_{-6}$
$2\gamma$ NNLO	$44^{+6}_{-5}$

**Tableau 5.11** – Section efficace totale mesurée avec  $4.9 \text{ fb}^{-1}$  de données *ATLAS* à  $\sqrt{s} = 7 \text{ TeV}$  et sections efficaces prédites par *DIPHOX*+*Gamma2MC* et  $2\gamma$ NNLO.

Les comparaisons entre les sections efficaces différentielles prédites et les sections efficaces différentielles mesurées sont présentées figures 5.34 (*PYTHIA*, *SHERPA*) et 5.35 (*DIPHOX*+*Gamma2MC*,  $2\gamma$ NNLO).

Sur tout le spectre  $m_{\gamma\gamma}$  considéré, les prédictions de  $2\gamma$ NNLO présentent un bon accord avec les données. A basse masse invariante,  $2\gamma$ NNLO parvient à décrire correctement les données tandis qu'un désaccord est observé pour l'estimation NLO de DIPHOX. De même, à basse masse, SHERPA correspond mieux aux données que PYTHIA, parce que des émissions réelles d'ordre supérieur sont incluses dans les calculs. On peut donc ici clairement en conclure que les corrections d'ordre supérieur sont nécessaires pour une bonne description des données à basse masse. A plus haute masse, par contre, pour des raisons inconnues les prédictions de SHERPA sont surestimées et les prédictions PYTHIA présentent un meilleur accord avec les données.

La région à bas  $p_{T,\gamma\gamma}$  est sensible à l'émission des gluons mous et pose un problème de divergence des logarithmes pour les calculs à ordre fixe. DIPHOX+Gamma2MC et  $2\gamma$ NNLO surestiment ainsi les données dans cette région. Pour PYTHIA et SHERPA la resommation des gluons mous est prise en compte par la parton shower, et de ce fait la description des données dans cette région à bas  $p_{T,\gamma\gamma}$  est meilleure.

Sur le reste du spectre, la composante NNLO manque clairement à DIPHOX qui sous-estime systématiquement la section efficace. Par ailleurs, on constate que la *Guillet shoulder* (voir section 5.5.1), liée à la composante NLO est correctement décrite dans SHERPA mais pas PYTHIA. Cette différence s'explique par les contributions additionnelles présentes dans SHERPA .

L'effet observé à  $\Delta\phi_{\gamma\gamma}=\pi$  est le même que celui constaté à bas  $p_{T,\gamma\gamma}$ . Les calculs à ordre fixe surestiment les données en raison de la resommation des gluons mous. A bas  $\Delta\phi_{\gamma\gamma}$ , là où la composante de fragmentation est la plus importante, les prédictions  $2\gamma$ NNLO sont sous-estimées mais l'accord avec les données reste meilleur qu'avec DIPHOX.

Pour toutes les prédictions, un désaccord est observé pour les grandes valeurs de  $\cos\theta^*$ . C'est une région particulièrement sensible à la composante de fragmentation ; on s'attend donc à une légère sous-estimation des données par  $2\gamma$ NNLO pour lequel toutes les composantes de fragmentation sont supprimées. Toutefois, même dans cette région,  $2\gamma$ NNLO décrit mieux les données que DIPHOX : la contribution NNLO est donc plus importante que la contribution de la fragmentation.

SHERPA, grâce aux émissions réelles d'ordre supérieur absentes dans PYTHIA, reproduit donc mieux les données, à part à haute masse invariante et à large  $\cos\theta^*$ . La composante NNLO implémentée dans  $2\gamma$ NNLO explique le bon accord de cette prédiction avec les données même si les différences à bas  $p_{T,\gamma\gamma}$  et haut  $\Delta\phi_{\gamma\gamma}$  subsistent en raison des divergences infrarouges.

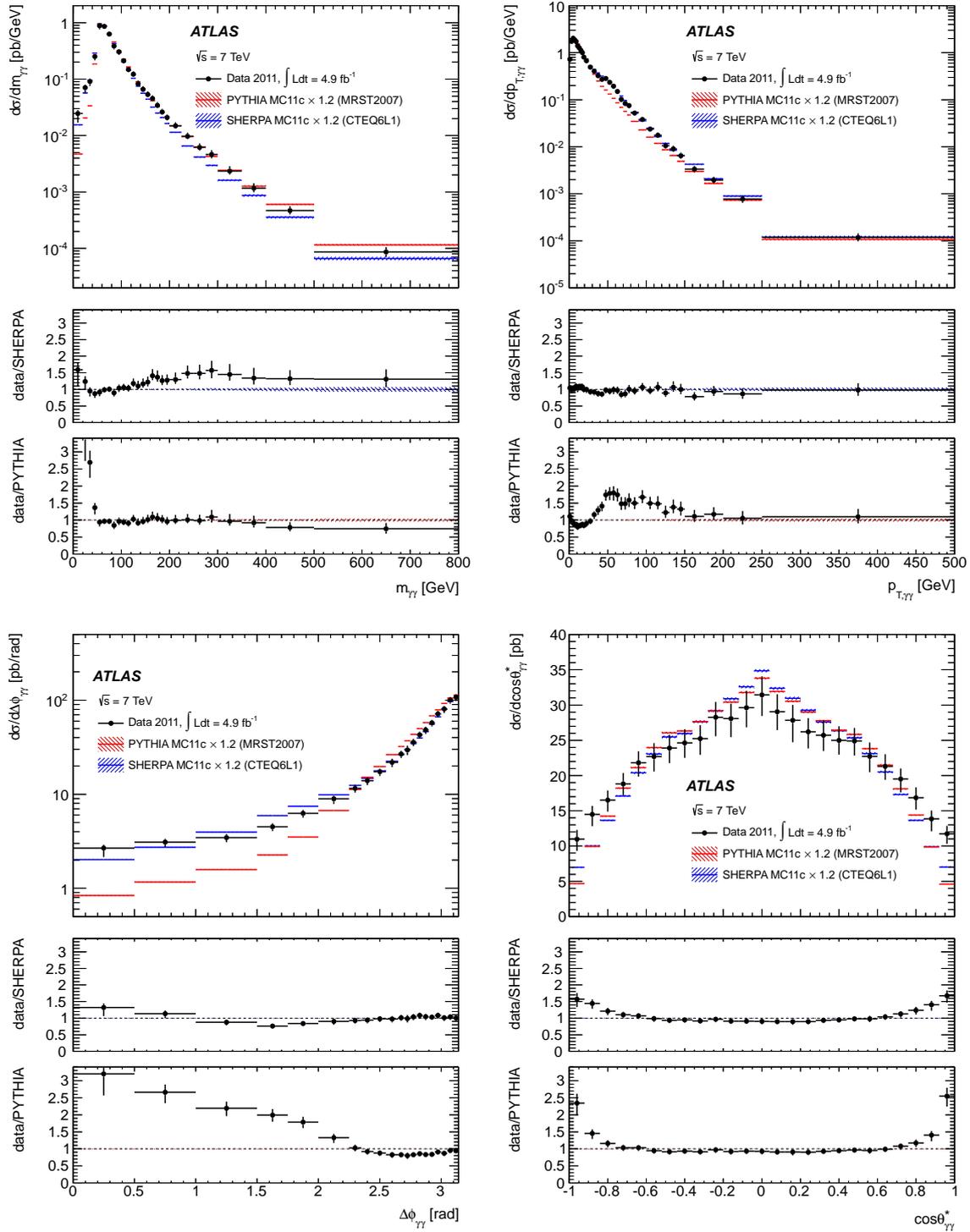
## 5.6 Conclusion

Les sections efficaces différentielles de production de paires de photons en fonction de  $m_{\gamma\gamma}$ ,  $p_{T,\gamma\gamma}$ ,  $\Delta\phi_{\gamma\gamma}$ , et  $\cos\theta^*$  ont été mesurées avec  $4.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$  et comparées aux prédictions théoriques.

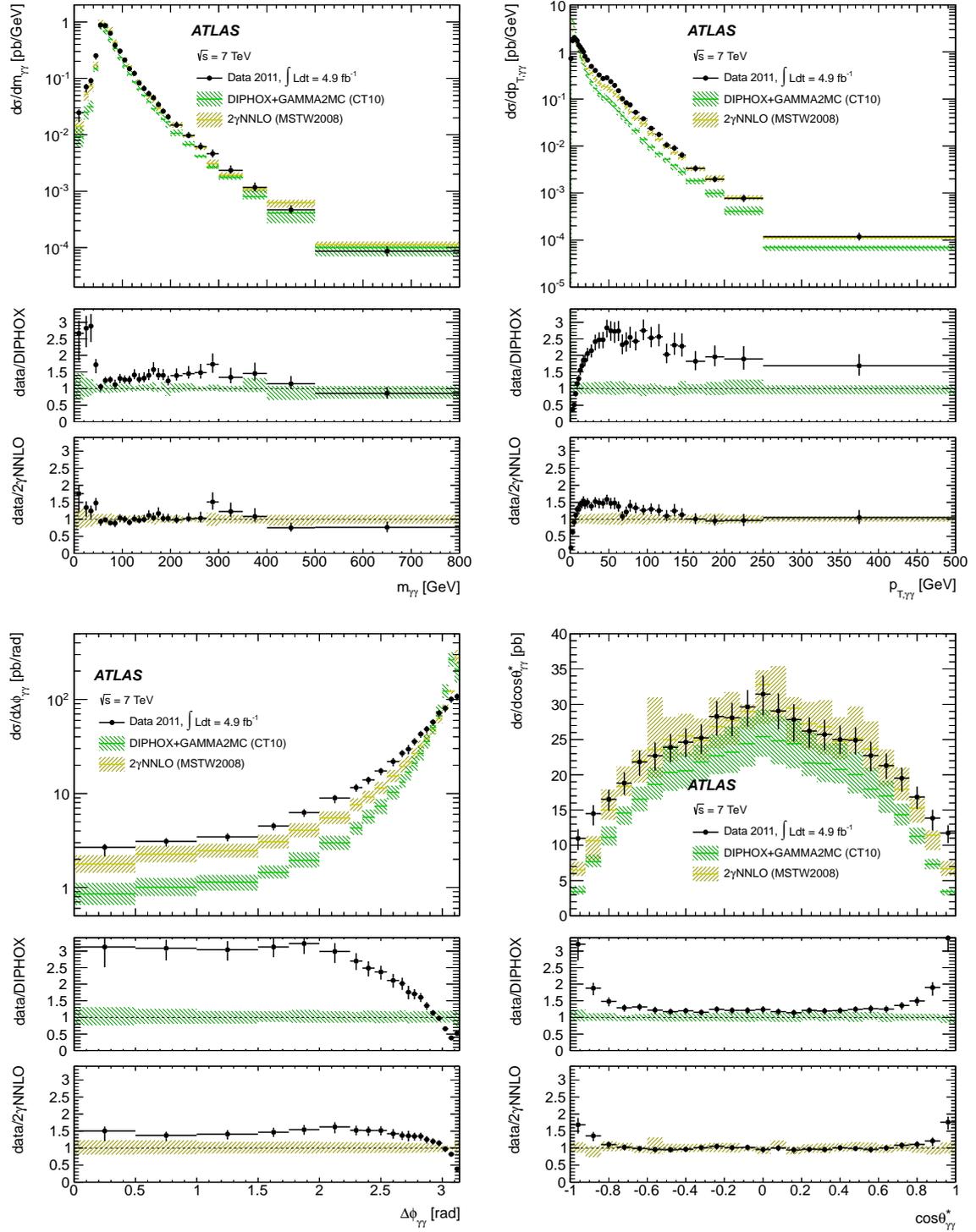
L'erreur sur notre mesure (dominée par l'erreur systématique) est comparable à l'erreur obtenue sur les prédictions théoriques à ordre fixe dans la plupart des régions de l'espace des phases (excepté à très basse masse invariante et à haut  $p_{T,\gamma\gamma}$ ). Les comparaisons avec les prédictions théoriques à NLO et NNLO montrent que la composante NNLO est nécessaire pour une description correcte des données et particulièrement à basse masse invariante. Des conclusions similaires sont tirées de la dernière mesure de sections efficaces différentielles par l'expérience CDF [131]. En outre, comme attendu, les zones à bas  $p_{T,\gamma\gamma}$  et à  $\Delta\phi_{\gamma\gamma} \sim \pi$ , sensibles à l'émission de gluons mous, sont mieux décrites par les générateurs Monte-Carlo avec parton shower. Parmi ces générateurs, SHERPA qui incluent des corrections réelles d'ordre supérieur, semble être le plus recommandé pour les études de photons et notamment pour l'analyse  $H \rightarrow \gamma\gamma$ .

Par ailleurs, la technique de décomposition du bruit de fond utilisée pour l'extraction des événements  $\gamma\gamma$  est validée et sera par la suite utilisée dans l'analyse  $H \rightarrow \gamma\gamma$  afin de s'assurer de la pureté en événements  $\gamma\gamma$  dans les échantillons de données sélectionnées, mais aussi dans l'estimation des incertitudes sur le nombre d'événements de signal.

Plusieurs étapes peuvent être envisagées suite à cette mesure. L'une des plus naturelles serait la mesure de sections efficaces à 8 TeV. Le rapport des mesures à 7 et 8 TeV permettrait de réduire un grand nombre d'incertitudes expérimentales et pourrait être comparé avec d'autant plus de poids au rapport des sections efficaces théoriques. On peut également imaginer une mesure de sections efficaces diphoton + jets afin de contraindre les PDFs des gluons mais également de fournir une meilleure compréhension du bruit de fond dans l'étude de la production du boson de Higgs par fusion de bosons vecteurs dans le canal diphoton.



**Figure 5.34** – Comparaison entre les section efficace différentielles mesurées expérimentalement et les prédictions des générateurs *PYTHIA* et *SHERPA*. Les distributions du nombre d'événements  $\gamma\gamma$  sont obtenues à partir d'une méthode de template fit (détaillée dans [99]). Les deux méthodes template fit et  $2\times 2D$  sideband présentent un accord à 1.5%.



**Figure 5.35** – Comparaison entre les section efficace différentielles mesurées expérimentalement et les prédictions des générateurs *DIPHOX+Gamma2MC* et *2 $\gamma$ NNLO*. Les distributions du nombre d'événements  $\gamma\gamma$  sont obtenues à partir d'une méthode de template (détaillée dans [99]). Les deux méthodes template fit et  $2\times 2D$  sideband présentent un accord à 1.5%.



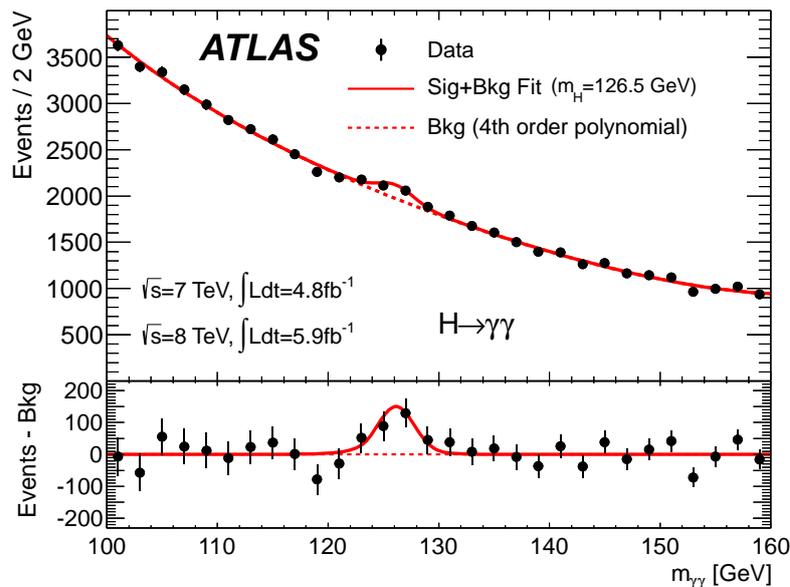
## Chapitre 6

# Observation d'une résonance dans la recherche du boson de Higgs Modèle Standard dans le canal diphoton

La recherche du boson de Higgs, lié au mécanisme de brisure spontanée de la symétrie électrofaible, était une des missions clé du LHC. Dans leur recherche du boson de Higgs, les collaborations ATLAS et CMS ont présenté en décembre 2011 un excès de  $2.8\sigma$  à  $m_{\gamma\gamma} = 126.5$  GeV et de  $3.1\sigma$  à  $m_{\gamma\gamma} = 124$  GeV respectivement [18, 19].

Nous présentons ici l'analyse développée pour la recherche du boson de Higgs dans le canal diphoton avec l'ensemble des données 2011, soit  $4.8 \text{ fb}^{-1}$  à  $\sqrt{s} = 7$  TeV et  $5.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8$  TeV enregistrées en 2012 par ATLAS.

Nous détaillerons, en premier lieu, la sélection des données qui correspondent au spectre de masse invariante présenté figure 6.1 ainsi que la composition du bruit de fond en événements  $\gamma\text{-}\gamma$ ,  $\gamma\text{-jet}$ ,  $\text{jet-jet}$  que j'ai obtenue à partir de la méthode 2x2D sideband développée au chapitre précédent.



**Figure 6.1** – Distribution de la masse invariante pour les données combinées à  $\sqrt{s} = 7$  TeV ( $4.8 \text{ fb}^{-1}$ ) et à  $\sqrt{s} = 8$  TeV ( $5.9 \text{ fb}^{-1}$ ) [132].

Dans un second temps, nous présenterons les différentes étapes dans la recherche d'une

résonance : la catégorisation des événements pour améliorer la signification du signal, la modélisation et paramétrisation du signal et du bruit de fond, les outils statistiques qui permettent de quantifier la déviation observée par rapport au bruit de fond et d'établir une découverte. Les résultats obtenus et l'observation de la résonance à  $m_{\gamma\gamma} \sim 126.5$  GeV seront présentés en dernier lieu.

## 6.1 Sélection des données, juillet 2012

### 6.1.1 Sélection inclusive des événements

La sélection appliquée sur les événements pour la recherche du boson de Higgs en deux photons diffère de celle utilisée pour la mesure de section efficace de paires de photons, notamment pour les coupures sur le  $p_T$  des photons et sur la masse invariante  $m_{\gamma\gamma}$  afin d'augmenter la pureté en événements  $\gamma\gamma$ . Les sélections sont détaillées ci-dessous pour les données 2011 et 2012 :

- Trigger 2011 : deux candidats photons LOOSE avec  $p_T > 20$  GeV (efficacité :  $98.9 \pm 0.2$  (stat)% , interactions par croisement de paquets maximum : 17). La luminosité associée est de  $4.8 \text{ fb}^{-1}$ .
- Trigger 2012 : deux candidats photons LOOSE avec  $p_T^{\gamma^1} > 35$  GeV et  $p_T^{\gamma^2} > 25$  GeV (efficacité :  $99.6 \pm 0.1$  (stat)% , interactions par croisement de paquets maximum : 25). La luminosité associée est de  $5.9 \text{ fb}^{-1}$ . Le seuil du trigger a été augmenté en raison de l'augmentation de la luminosité.
- Acceptance :  $0 < |\eta| < 1.37 \cup 1.52 < |\eta| < 2.37$ .
- Fenêtre de masse invariante : 100-160 GeV.
- Coupure en  $p_t$  :  $p_T^{\gamma^1} > 40$  GeV et  $p_T^{\gamma^2} > 30$  GeV. La coupure en  $p_t$  est asymétrique afin d'augmenter la réjection du bruit de fond tout en gardant une bonne efficacité de signal et une statistique suffisante.
- Isolation : isolation calorimétrique (définie section 4.3.1) `topoEtCone` < 4 GeV.
- Critère d'identification 2011 : il est basé sur un *Neural Network* (section 4.4.3) avec comme variables d'entrée les variables utilisées pour le critère d'identification TIGHT. L'amélioration sur l'efficacité d'identification par rapport à la sélection TIGHT 2011 est de  $\sim 15\%$  [110].
- Critère d'identification 2012 : il s'agit du critère TIGHT pour les deux candidats photons. L'identification TIGHT a été optimisée pour les conditions d'empilement à 8 TeV. L'efficacité d'identification des photons, moyennée en  $\eta$ , est de l'ordre de 85% à bas  $E_T$  ( $\sim 40$  GeV) et atteint 95% à partir de 100 GeV (figures 4.16 et 4.17).
- Sélection du vertex primaire est faite avec une méthode d'ajustement minimale combinant la direction des photons déterminée par le calorimètre, la position moyenne du faisceau et la somme du  $p_T^2$  des traces associées à chaque vertex reconstruit [133]. Pour les données 2011, le vertex de conversion est également utilisé [134]. La bonne identification du vertex primaire est nécessaire pour la détermination de la masse invariante et pour la sélection des jets associés (voir section 6.2.1).

### 6.1.2 Composition du bruit de fond

Afin d'estimer la proportion d'événements  $\gamma\gamma$  dans l'échantillon de données sélectionnées pour l'analyse  $H \rightarrow \gamma\gamma$ , la même technique de décomposition que pour la mesure de section efficace diphoton est appliquée. Cette décomposition est effectuée pour plusieurs raisons :

- S'assurer que le spectre  $m_{\gamma\gamma}$  est dominé par la composante  $\gamma\gamma$  et ainsi monitorer la performance de l'identification des photons.
- Normaliser les contributions  $\gamma$ - $\gamma$ ,  $\gamma$ -*jet* et *jet-jet* issues de la simulation pour construire une modélisation valide du bruit de fond dans le canal  $H \rightarrow \gamma\gamma$ .
- Vérifier sur les données, la paramétrisation du bruit de fond déterminée à partir de la simulation.

Les erreurs systématiques considérées et la façon dont elles sont évaluées sont identiques à ce qui est fait pour la mesure de section efficace différentielle diphoton (section 5.2.3).

#### 2011

L'application de la méthode 2x2D sideband à l'ensemble des données 2011 donne la décomposition suivante :

- $N_{\gamma\gamma} = 18968 \pm 205$  (stat)  $^{+905}_{-875}$  (syst)
- $N_{\gamma j} = 3330 \pm 105$  (stat)  $^{+655}_{-715}$  (syst)
- $N_{j\gamma} = 1115 \pm 64$  (stat)  $^{+193}_{-135}$  (syst)
- $N_{jj} = 417 \pm 28$  (stat)  $^{+127}_{-146}$  (syst)

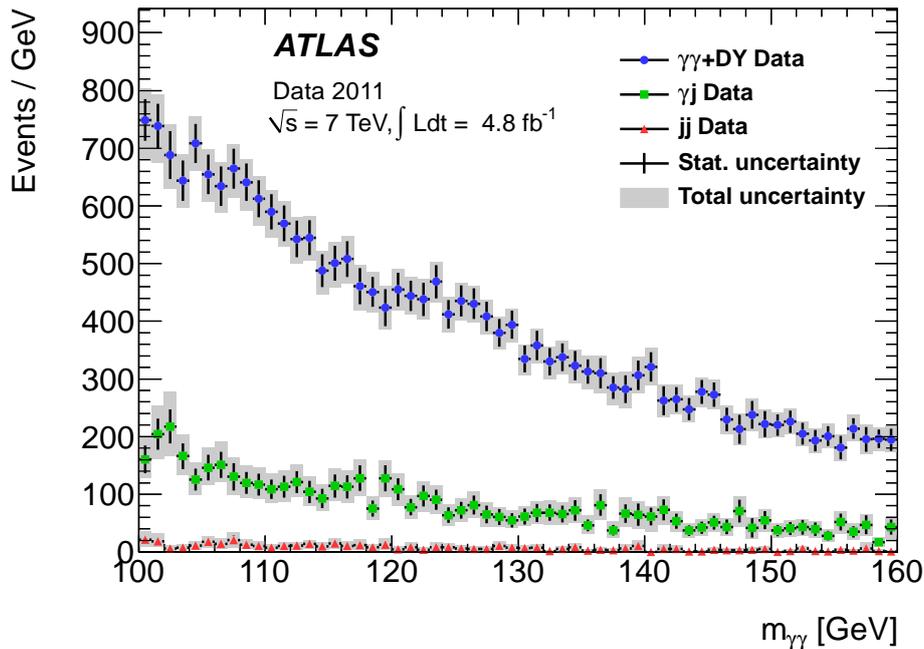
La pureté, c'est à dire le pourcentage d'événements  $\gamma\gamma$ , de l'échantillon sélectionné pour les données 2011 à  $\sqrt{s} = 7$  TeV est de  $79.6 \pm 0.5$  (stat)  $^{+3.8}_{-3.7}$  (syst) %.

Les erreurs systématiques sont données tableau 6.1. La décomposition du spectre de masse invariante est présentée figure 6.2.

Source d'erreurs	$N_{\gamma\gamma}$	$N_{\gamma j}$	$N_{j\gamma}$	$N_{jj}$
systematiques				
Valeur centrale	18968	3330	1115	417
Définition de la région de contrôle du bruit de fond	+557 -515	+274 -391	+151 -63	+70 -102
Efficacité isolation	+695 -696	+592 -592	+109 -106	+105 -104
Efficacité identification	+158 -126	+64 -86	+51 -55	+11 -15
Erreur syst. totale	+905 -875	+655 -715	+193 -135	+127 -146

**Tableau 6.1** –  $4.8 \text{ fb}^{-1}$  à  $\sqrt{s} = 7 \text{ TeV}$  (données 2011) : Nombres d'événements totaux par composante pour l'analyse  $H \rightarrow \gamma\gamma$  obtenus avec la méthode  $2 \times 2D$  sideband. Les différentes erreurs systématiques sont données par composante. L'erreur systématique totale correspond à la somme quadratique des différentes contributions.

#### 2012



**Figure 6.2** – Spectre différentiel  $m_{\gamma\gamma}$  pour les contributions  $\gamma$ - $\gamma$ ,  $\gamma$ -jet+jet- $\gamma$  et jet-jet dans l'analyse  $H \rightarrow \gamma\gamma$  pour  $4.8 \text{ fb}^{-1}$  à  $\sqrt{s} = 7 \text{ TeV}$  (données 2011).

Pour l'échantillon de données 2012, les événements sont répartis tels que :

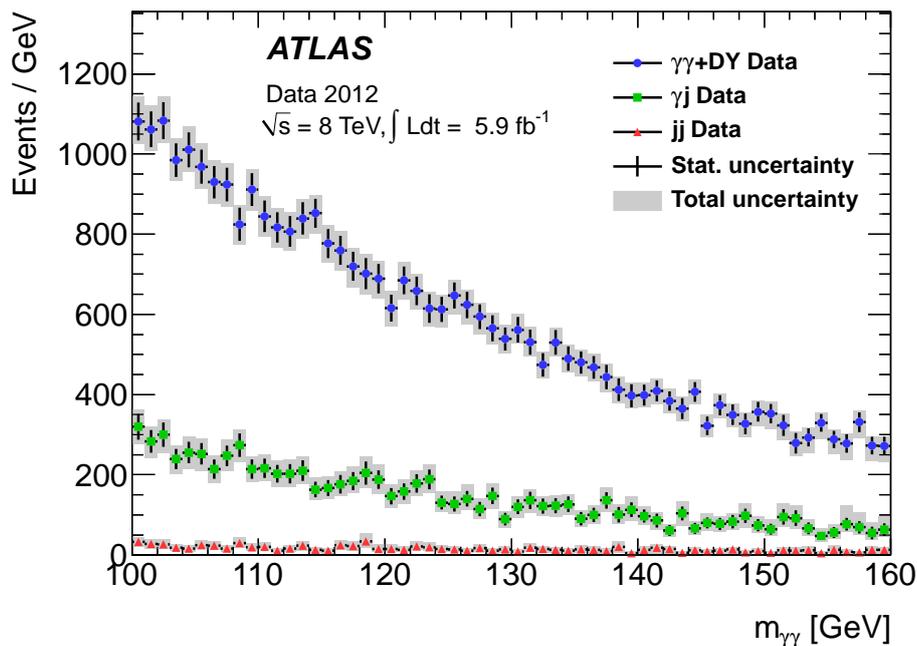
$$\begin{aligned}
 - N_{\gamma\gamma} &= 26567 \pm 259 \text{ (stat)} \begin{matrix} +877 \\ -747 \end{matrix} \text{ (syst)} \\
 - N_{\gamma j} &= 5459 \pm 134 \text{ (stat)} \begin{matrix} +499 \\ -577 \end{matrix} \text{ (syst)} \\
 - N_{j\gamma} &= 2297 \pm 94 \text{ (stat)} \begin{matrix} +239 \\ -197 \end{matrix} \text{ (syst)} \\
 - N_{jj} &= 909 \pm 40 \text{ (stat)} \begin{matrix} +170 \\ -230 \end{matrix} \text{ (syst)}
 \end{aligned}$$

La pureté de l'échantillon sélectionné pour  $5.9 \text{ fb}^{-1}$  de données 2012 à  $\sqrt{s} = 8 \text{ TeV}$  est de  $75.4 \pm 0.4 \text{ (stat)} \begin{matrix} +2.6 \\ -2.2 \end{matrix} \text{ (syst)} \%$ .

Les erreurs systématiques sont données tableau 6.2. La décomposition du spectre de masse invariante est présentée figure 6.3.

Source d'erreurs	$N_{\gamma\gamma}$	$N_{\gamma j}$	$N_{j\gamma}$	$N_{jj}$
systematiques				
Valeur centrale	26567	5459	2297	909
Définition de la région de contrôle du bruit de fond	+690 -508	+254 -380	+202 -156	+65 -170
Efficacité isolation	+494 -494	+403 -403	+116 -116	+154 -154
Efficacité identification	+224 -238	+149 -164	+55 -35	+31 -21
Erreur syst. totale	+877 -747	+499 -577	+239 -197	+170 -230

**Tableau 6.2** –  $5.9 \text{ fb}^{-1}$  à  $\sqrt{s} = 8 \text{ TeV}$  (données 2012) : Nombres d'événements totaux par composante pour l'analyse  $H \rightarrow \gamma\gamma$  obtenus avec la méthode  $2 \times 2D$  sideband. Les différentes erreurs systématiques sont données par composante. L'erreur systématique totale correspond à la somme quadratique des différentes contributions.



**Figure 6.3** – Spectre différentiel  $m_{\gamma\gamma}$  pour les contributions  $\gamma\text{-}\gamma$ ,  $\gamma\text{-jet}+\text{jet-}\gamma$  et  $\text{jet-jet}$  dans l’analyse  $H \rightarrow \gamma\gamma$  pour  $5.9 \text{ fb}^{-1}$  à  $\sqrt{s} = 8 \text{ TeV}$  (données 2012).

L’excellente efficacité sur l’identification des photons ainsi que le critère d’isolation amène à un facteur de réjection des jets de l’ordre  $10^4$  (section 4.4.3) et permet ainsi d’obtenir une pureté en événements  $\gamma\gamma$  de 75 à 80% dans les données sélectionnées.

Dans ces 75 à 80% d’événements  $\gamma\gamma$ , une petite fraction d’événements provient de l’identification en photons des électrons issus du processus Drell-Yan. Cette fraction est estimée dans les données à partir des événements  $Z \rightarrow e^+e^-$ . Pour les données à 7 TeV elle est de  $\sim 2\%$  et de  $\sim 1\%$  pour les données à 8 TeV [133].

La différence de pureté entre les données 2011 et 2012 s’explique principalement par l’utilisation pour les données 2011 d’un Neural Network qui exploite les corrélations entre les variables d’identification et améliore l’efficacité d’identification des photons.

## 6.2 Recherche d’une résonance

### 6.2.1 Catégorisation

Parmi les 59059 événements observés dans les données 2011 et 2012, 191 événements de signaux sont attendus. Afin d’augmenter la sensibilité de la recherche, les événements de données sélectionnés peuvent être classés dans des catégories avec différents rapports signal sur bruit et différentes résolutions sur la masse invariante [91]. Les propriétés des événements diphoton, comme le statut de conversion et la région en  $\eta$  permettent de définir ces sous-ensembles [135]. En plus de ces deux variables, le  $p_{Tt}$ , a également été introduit dans la catégorisation pour augmenter encore un peu plus la sensibilité de l’analyse [136]. Le  $p_{Tt}$  du système diphoton est défini comme la composante perpendiculaire à l’axe principal de l’impulsion transverse du système diphoton ([137] et [138]), comme illustré figure 6.4.

La figure 6.6 montre la distribution en  $p_{Tt}$  pour les événements de bruit de fond et de signal issus de la simulation. Le bruit de fond  $\gamma\gamma$  est obtenu à partir de la simulation SHERPA diphoton et le bruit de fond  $\gamma\text{-jet}$  à partir de la simulation photon+jet ALPGEN. Leurs contributions respectives ont été normalisées à celles extraites des données en utilisant la méthode 2x2D

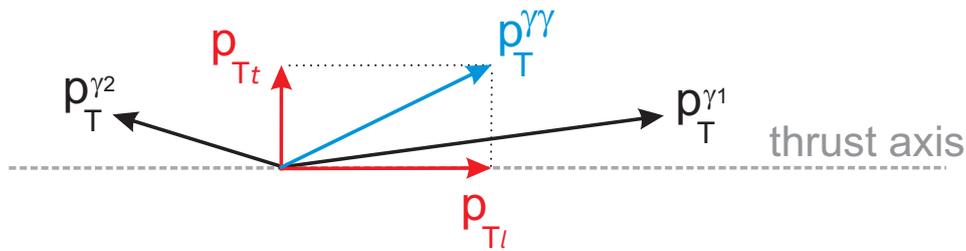
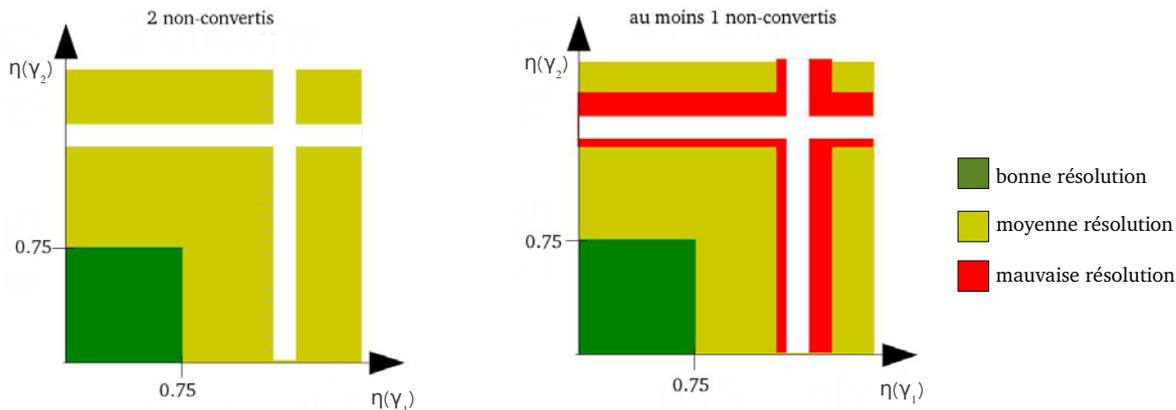


Figure 6.4 – Illustration de la définition de la variable  $p_{Tt}$ .

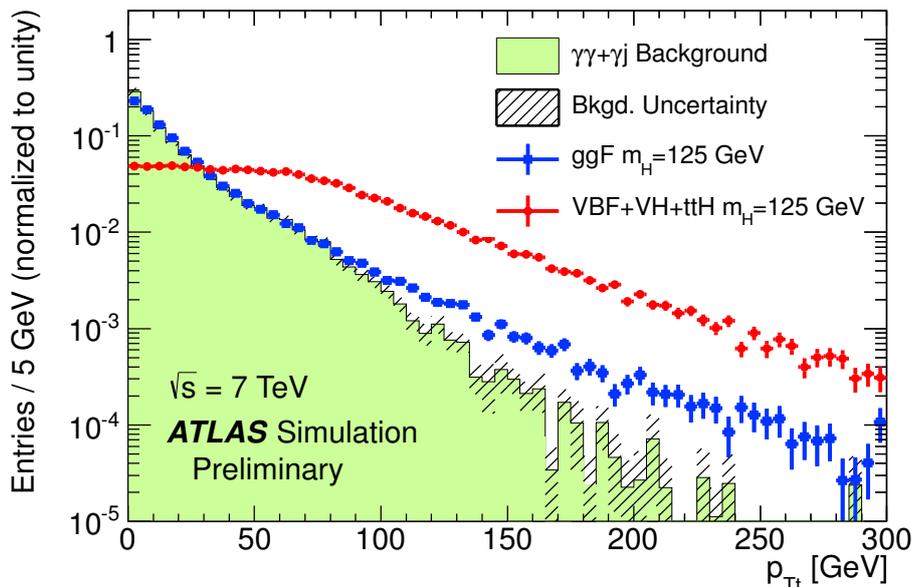
sideband (section 6.1.2), la contribution *jet-jet* a été négligée. La distribution du  $p_{Tt}$  pour le signal lié à la production du boson de Higgs par fusion de gluons est beaucoup plus proche de celle du bruit de fond que pour les autres modes de production du signal. Une coupure à 60 GeV sur le  $p_{Tt}$  est choisie pour la catégorisation [133].

Neuf catégories sont définies en fonction de  $p_{Tt}$ , de  $\eta$  et du statut de conversion (figure 6.5) :

1. Non-convertis central, bas  $p_{Tt}$  ou CP1 : Les deux candidats photons sont reconstruits comme des photons non-convertis.  $|\eta| < 0.75$  pour les deux candidats.  $p_{Tt} < 60$  GeV.
2. Non-convertis central, haut  $p_{Tt}$  ou CP2 : Les deux candidats photons sont reconstruits comme des photons non-convertis.  $|\eta| < 0.75$  pour les deux candidats.  $p_{Tt} > 60$  GeV.
3. Non-converti non-central, bas  $p_{Tt}$  ou CP3 : Les deux candidats photons sont reconstruits comme des photons non-convertis.  $|\eta| > 0.75$  pour au moins l'un des deux candidats.  $p_{Tt} < 60$  GeV.
4. Non-converti non-central, haut  $p_{Tt}$  ou CP4 : Les deux candidats photons sont reconstruits comme des photons non-convertis.  $|\eta| > 0.75$  pour au moins l'un des deux candidats.  $p_{Tt} > 60$  GeV.
5. Converti central, bas  $p_{Tt}$  ou CP5 : Au moins un des deux candidats photons est reconstruit comme photon converti.  $|\eta| < 0.75$  pour les deux candidats.  $p_{Tt} < 60$  GeV.
6. Converti central, haut  $p_{Tt}$  ou CP6 : Au moins un des deux candidats photons est reconstruit comme photon converti.  $|\eta| < 0.75$  (pour les deux candidats) et  $p_{Tt} > 60$  GeV.
7. Converti non-central, bas  $p_{Tt}$  ou CP7 : Au moins un des deux candidats photons est reconstruit comme photon converti. Les deux candidats sont tels que  $|\eta| < 1.3$  ou  $|\eta| < 1.75$ , mais  $|\eta| > 0.75$  pour au moins un des deux candidats.  $p_{Tt} < 60$  GeV.
8. Converti non-central, haut  $p_{Tt}$  ou CP8 : Au moins un des deux candidats photons est reconstruit comme photon converti. Les deux candidats sont tels que  $|\eta| < 1.3$  ou  $|\eta| < 1.75$ , mais  $|\eta| > 0.75$  pour au moins un des deux candidats.  $p_{Tt} > 60$  GeV.
9. Converti transition ou CP9 : Au moins un des deux candidats photons est reconstruit comme photon converti. Au moins un des deux candidats photons est tel que  $1.3 < |\eta| < 1.37$  ou  $1.52 < |\eta| < 1.75$ .



**Figure 6.5** – Illustration des coupures en  $\eta$  pour la définition des catégories, dans le cas où les deux photons sont non-convertis (gauche) et dans le cas où au moins l'un des photons est converti (droite). La coupe en  $p_{Tt}$  n'est pas représentée ici.

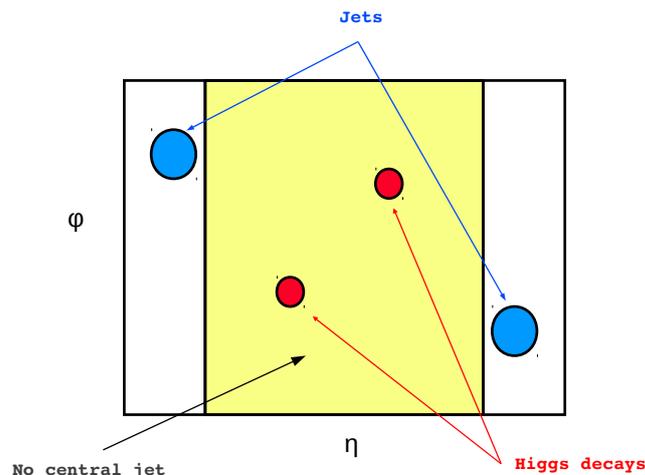


**Figure 6.6** – Distribution en  $p_{Tt}$  pour les événements de bruit de fond. Le bruit de fond  $\gamma\gamma$  est obtenu à partir de la simulation *SHERPA* diphoton et le bruit de fond  $\gamma$ -jet à partir de la simulation *photon+jet ALPGEN*. Leurs contributions respectives ont été normalisées à celles extraites des données à partir de la méthode *2x2D sideband* (section 6.1.2), la contribution *jet-jet* a été négligée [133].

Les différences de distributions du  $p_{Tt}$  pour les différents modes de production peuvent aussi se révéler utiles pour leur séparation.

En plus des catégories exploitant les propriétés du système diphoton, une dernière catégorie dédiée aux événements VBF est définie, la catégorie 2-jets ou CP10. Après la fusion de gluons, la fusion de bosons vecteurs est le plus important mode de production prédit de bosons de Higgs au LHC. Sa signature consiste en deux jets très écartés en  $\eta$  avec une faible activité hadronique centrale en raison de l'absence d'échanges colorés. Les produits de la désintégration du boson de Higgs sont, quant à eux, attendus dans la région centrale du détecteur comme

illustré figure 6.7. Expérimentalement, la catégorie 2-jets VBF est définie par la présence de



**Figure 6.7** – Illustration de la topologie d'un événement VBF dans le plan  $(\eta, \phi)$ .

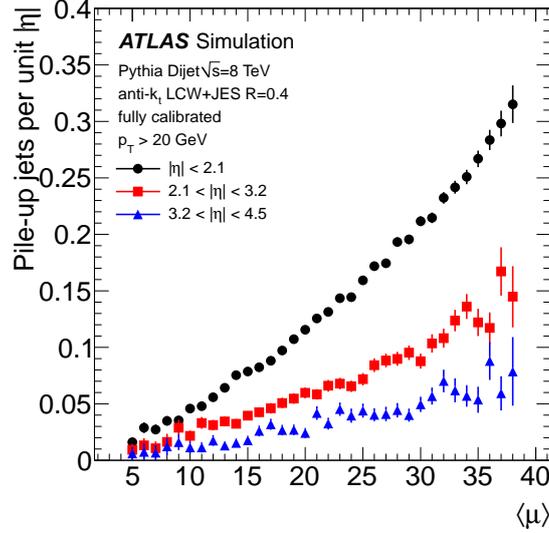
deux jets additionnels dans l'événement. Les jets sont reconstruits avec un algorithme anti- $k_t$  [139] de paramètre  $R=0.4$ , les deux jets les plus énergétiques passant les coupures suivantes sont retenus :

- $|\eta| < 4.5$  et  $p_T^{jet} > 25$  GeV. Pour les données 2012, en raison de l'augmentation des événements d'empilement, la coupure sur  $p_T^{jet}$  est élevée à 30 GeV pour les jets tels que  $2.4 \leq |\eta| < 4.5$ .
- Avec l'augmentation de l'empilement le nombre de jets reconstruits issus des fluctuations de l'empilement augmente comme illustré figure 6.8. La *jet-vertex-fraction* ou JVF est le ratio de la somme en  $p_T$  des traces comprises dans le jet et associées au vertex primaire sur la somme en  $p_T$  de toutes les traces associées au jet. Une coupure sur cette variable permet de diminuer significativement la contribution des jets d'empilements. Pour les jets situés dans la zone d'acceptance du trajectographe ( $|\eta| < 2.5$ ), une coupure minimale de 0.5 sur la JVF est demandée.
- La distance minimale  $\Delta R$  entre chaque jet et les candidats photons sélectionnés doit être de 0.4 afin de limiter la superposition des cônes d'isolation.
- La masse invariante  $m_{jj} > 400$  GeV (figure 6.9).
- La séparation en  $\eta$  entre les deux jets  $|\Delta\eta_{jj}| > 2.8$  (figure 6.9).
- L'angle azimuthal avec le système diphoton  $|\Delta\phi_{\gamma\gamma-jj}| > 2.6$  (figure 6.10).

Le nombre d'événements par catégorie observés dans les données et le nombre d'événements de signal attendus à 7 et 8 TeV sont présentés tableau 6.3.

## 6.2.2 Modélisation et paramétrisation du bruit de fond

Afin d'estimer le nombre d'événements de bruit de fond par catégorie, la distribution en masse invariante du bruit de fond par catégorie est paramétrée par une fonction analytique. Les paramètres de cette fonction sont obtenus par un fit sur les données mais le type de fonction est choisi à partir de la simulation. Différentes fonctions sont testées sur trois simulations de bruit de fond : SHERPA, RESBOS [42] et DIPHOX. SHERPA est également utilisé pour le bruit de fond  $\gamma$ -jet et PYTHIA est choisi pour le bruit de fond jet-jet. Les contributions relatives de chaque composante du bruit de fond sont normalisées aux valeurs extraites des données à partir de la méthode  $2 \times 2D$  sideband (section 6.1.2). La fonction apportant le meilleur compromis entre



**Figure 6.8** – Nombre moyen de jets d’empilement (anti- $k_t$   $R=0.4$  et  $p_T > 20$  GeV) en fonction du nombre moyen d’événements d’empilement par croisement de paquets. Plusieurs régions en  $\eta$  sont considérées,  $|\eta| < 2.1$ ,  $2.1 < |\eta| < 3.2$  et  $3.2 < |\eta| < 4.5$  [140].

le biais introduit et le pouvoir statistique est choisie [133], [141] et [142]. Selon les catégories, trois fonctions sont retenues, une fonction exponentielle pour les catégories dans lesquelles la statistique est faible, un polynôme exponentiel du second degré ou un polynôme de Bernstein d’ordre quatre<sup>1</sup> dans les autres cas.

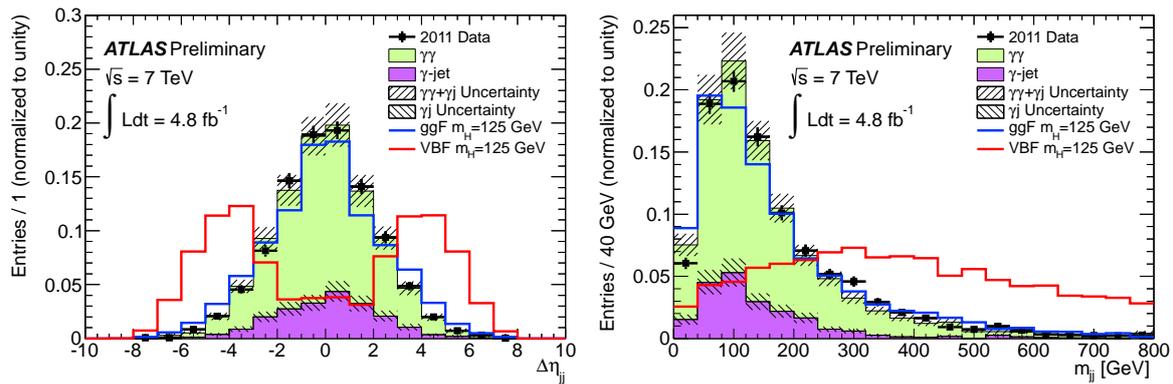
Le nombre d’événements de bruit de fond mais aussi le nombre d’événements de signal sont affectés par la fonction utilisée pour paramétrer le bruit de fond. L’erreur systématique sur le nombre d’événements de signal introduite par le choix de ces fonctions est estimée à partir de la simulation du bruit de fond.

La paramétrisation du bruit de fond est établie à partir de la simulation, la statistique dans les données étant trop faible. On sait, par la mesure de section efficace différentielle diphoton, que la forme de la composante  $\gamma\text{-}\gamma$  est correctement décrite dans la simulation, mais afin de vérifier la validité de la simulation  $\gamma\text{-jet}$  et  $jet\text{-}jet$ , on peut tester les fonctions choisies en utilisant les composantes  $\gamma\text{-jet}$  et  $jet\text{-}jet$  obtenues à partir des données. Pour se faire, on effectue la décomposition du bruit de fond par catégorie. Toutefois, en raison d’une statistique trop faible, certaines catégories sont regroupées. On définit ainsi :

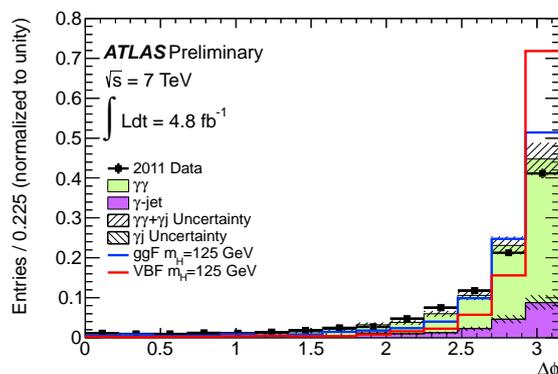
- "crack" : CP9
- "haut  $p_{Tt}$ " : CP2 + CP4 + CP6 + CP8
- "bas  $p_{Tt}$ ,  $|\eta| < 0.75$ " : CP1 + CP5
- "bas  $p_{Tt}$ ,  $|\eta| > 0.75$ " : CP3 + CP7

La décomposition du bruit de fond pour chacune de ces catégories est présentée tableau 6.4. L’erreur statistique sur le nombre d’événements  $\gamma\gamma$  est de l’ordre de  $\pm 0.3\text{-}1.6\%$ , quant à l’erreur systématique, elle est d’environ  $\pm 5\%$ . On constate que la composition du bruit de fond change avec la catégorie, la pureté maximale est obtenue pour la catégorie "bas  $p_{Tt}$ ,  $|\eta| < 0.75$ ". Les distributions  $m_{\gamma\gamma}$  pour les quatre catégories ainsi définies sont présentées figures 6.11 et 6.12. Les composantes  $\gamma\text{-jet}$  et  $jet\text{-}jet$  obtenues à partir des données par catégorie sont additionnées

1. La base de polynômes de Bernstein de degré  $m$  est définie par les polynômes  $B_0^m, \dots, B_m^m$  tels que  $B_i^m(x) = A(i, m)x^i(1-x)^{m-i}$  avec  $A(i, m)$  les coefficients binomiaux. Un polynôme de Bernstein d’ordre  $m$  est une combinaison linéaire des polynômes de la base.



**Figure 6.9** – Différence en  $\eta$  entre le premier et second jet (gauche) et masse invariante du système dijet (droite). Les données sont comparées à la simulation. Le bruit de fond  $\gamma\gamma$  est obtenu à partir de la simulation *SHERPA* diphoton et le bruit de fond  $\gamma$ -jet à partir de la simulation *photon+jet ALPGEN*. Leurs contributions respectives ont été normalisées à celles extraites des données à partir de la méthode *2x2D sideband* (section 6.1.2), la contribution *jet-jet* a été négligée [133].



**Figure 6.10** – Séparation en  $\phi$  entre le système diphoton et le système dijet. Le bruit de fond  $\gamma\gamma$  est obtenu à partir de la simulation *SHERPA* diphoton et le bruit de fond  $\gamma$ -jet à partir de la simulation *photon+jet ALPGEN*. Leurs contributions respectives ont été normalisées à celles extraites des données à partir de la méthode *2x2D sideband* (section 6.1.2), la contribution *jet-jet* a été négligée [133].

à la composante  $\gamma\gamma$  *SHERPA* normalisée à la pureté extraite des données par catégorie. Ces distributions sont comparées à celles obtenues uniquement à partir de la simulation (*SHERPA*  $\gamma$ -jet et *PYTHIA* *jet-jet*) normalisées suivant la composition inclusive du bruit de fond. La composante Drell-Yan mesurée à partir des données est additionnée dans les deux cas. Malgré les différences observées entre la composition du bruit de fond estimée à partir des données par catégorie et celle estimée à partir des données inclusives (et utilisée pour la normalisation des événements *SHERPA+PYTHIA*), particulièrement pour la catégorie "bas  $p_{Tt}$ ,  $|\eta| < 0.75$ ", où un facteur deux est observé sur la composante *jet-jet*, les formes des différentes composantes présentent un accord correct dans les différentes catégories.

### 6.2.3 Modélisation et paramétrisation du signal

Dans chaque catégorie, la fonction analytique choisie pour décrire le signal est la somme d'une fonction *Crystal Ball* [143] de largeur  $\sigma_{CB}$  pour prendre en compte le corps de la distribution et la queue non gaussienne à plus petite masse invariante et d'une fonction Gaussienne de largeur  $\sigma_G$  pour modéliser les événements éloignés du corps de la distribution. Les queues

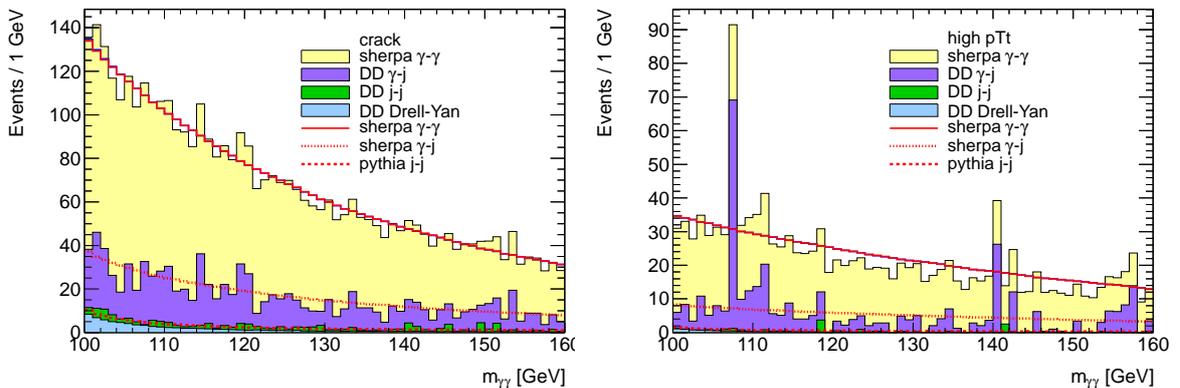
Categorie	$\sqrt{s} = 7$ TeV $\mathcal{L} = 4.8 \text{ fb}^{-1}$		$\sqrt{s} = 8$ TeV $\mathcal{L} = 5.9 \text{ fb}^{-1}$	
	Observé	$N_S$	Observé	$N_S$
CP1 Non-converti central, bas $p_{Tt}$	2054	10.4	2945	14.4
CP2 Non-converti central, haut $p_{Tt}$	97	1.5	173	2.5
CP3 Non-converti non-central, bas $p_{Tt}$	7129	21.6	12136	31.4
CP4 Non-converti non-central, haut $p_{Tt}$	444	2.7	785	5.3
CP5 Converti central, bas $p_{Tt}$	1493	6.7	2021	9.1
CP6 Converti central, haut $p_{Tt}$	77	1.0	113	1.6
CP7 Converti non-central, bas $p_{Tt}$	8313	21.0	11112	27.3
CP8 Converti non-central, haut $p_{Tt}$	501	2.7	708	4.6
CP9 Converti transition	3591	9.5	5149	13.0
CP10 2-jets	89	2.2	129	2.9
Total	23788	79.3	35271	111.6

**Tableau 6.3** – Nombre d'événements observés dans les données et nombre d'événements de signal attendus ( $N_S$ ) par catégorie pour les données à  $\sqrt{s} = 7$  TeV et les données à  $\sqrt{s} = 8$  TeV dans l'intervalle  $m_{\gamma\gamma} [100-160]$  GeV.

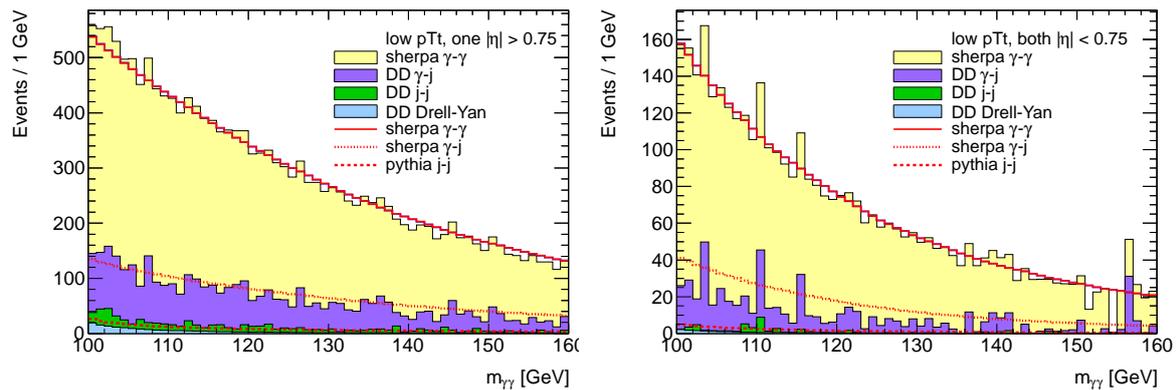
	inclusif/simulation	crack	haut $p_{Tt}$	bas $p_{Tt}$ , $ \eta  < 0.75$	bas $p_{Tt}$ , $ \eta  > 0.75$
$\gamma\gamma$	79.6	77.8	73.3	86.7	79.2
$\gamma$ -jet	18.7	20.0	26.7	12.4	18.0
jet-jet	1.7	2.2	0.0	0.9	2.8

**Tableau 6.4** –  $4.8 \text{ fb}^{-1}$  à  $\sqrt{s} = 7$  TeV (données 2011) : décompositions du bruit de fond obtenues avec la méthode  $2 \times 2D$  sideband pour les quatre catégories définies dans le texte.

non-gaussiennes correspondent à des effets du détecteur (principalement liés à la conversion des



**Figure 6.11** –  $m_{\gamma\gamma}$  pour les catégories "crack" et "haut  $p_{Tt}$ " définies dans le texte.

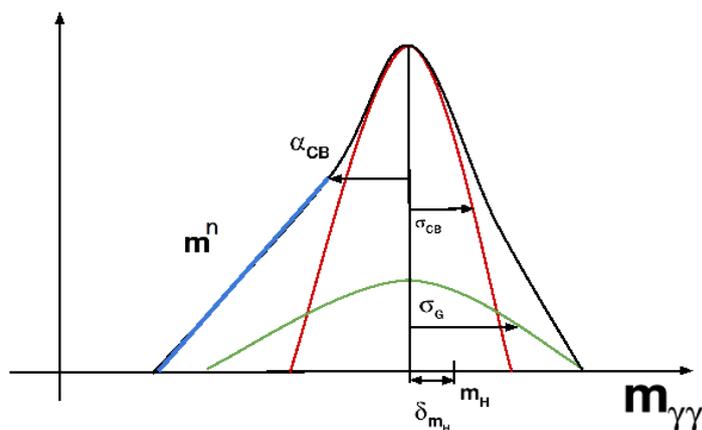


**Figure 6.12** –  $m_{\gamma\gamma}$  pour les catégories "bas  $p_{Tt}$ ,  $|\eta| > 0.75$ " et "bas  $p_{Tt}$ ,  $|\eta| < 0.75$ " définies dans le texte.

photons dans la matière en amont du calorimètre). La fonction Crystal Ball est définie par :

$$N. \begin{cases} e^{-t^2/2} & \text{si } t > -\alpha_{CB} \\ \left(\frac{n}{|\alpha_{CB}|}\right)^n \cdot e^{-|\alpha_{CB}|^2/2} \cdot \left(\frac{n}{|\alpha_{CB}|} - |\alpha_{CB}| - t\right)^{-n} & \text{sinon} \end{cases} \quad (6.1)$$

$t = (m_{\gamma\gamma} - m_h - \delta_{m_h})/\sigma_{CB}$  avec  $\delta_{m_h}$  un décalage dépendant de la catégorie.  $N$  est un facteur de normalisation,  $\sigma_{CB}$  représente la résolution sur la masse invariante diphoton,  $n$  et  $\alpha_{CB}$  sont les paramètres de la queue non-gaussienne comme illustré figure 6.13. Les valeurs de ces différents paramètres sont obtenues à partir de la simulation mais ils dépendent de la masse hypothétique du boson de Higgs. Un fit simultané sur un grand nombre de points de masse permet d'extraire l'évolution de ces paramètres en fonction de la masse  $m_H$ .



**Figure 6.13** – Illustration de la fonction Crystal Ball + gaussienne utilisée pour décrire la forme du signal.

Les échantillons de simulation utilisés pour les productions du boson Higgs par fusion de gluons et par fusion de boson vecteurs sont issues du générateur POWHEG [48, 49] interfacé avec PYTHIA [144] pour la simulation de l'hadronisation et de la parton-shower. Pour la production du boson de Higgs en association avec des bosons vecteurs ou avec une paire de quark top, c'est le générateur PYTHIA qui est utilisé. Plus de détails sur les différentes corrections appliquées aux calculs des sections efficaces peuvent être trouvés dans la section 1.4.1.

Tous les échantillons Monte-Carlo générés sont ensuite injectés dans la simulation GEANT4 du détecteur. Les événements d'empilement sont également pris en compte en ajoutant aux événements de signal des interactions inélastiques  $pp$ . Le nombre d'interactions supplémentaires est

ajusté en fonction du nombre moyen d'interactions par croisement de paquets mesuré dans chaque période de prise de données et pondéré en fonction de la luminosité intégrée.

### 6.2.4 Erreurs systématiques

L'incertitude expérimentale dominante sur le nombre d'événements de signal provient de l'efficacité de reconstruction et d'identification des photons, elle est de  $\pm 8\%$  et  $\pm 11\%$  pour les données à 7 TeV et 8 TeV respectivement [133]. Les erreurs théoriques considérées sont celles sur les sections efficaces de production (liées aux choix des échelles d'énergies et des PDFs) ainsi que celle sur le rapport d'embranchement. Les erreurs systématiques associées sur le nombre d'événements de signal attendu sont résumées tableau 6.5.

L'ensemble des erreurs systématiques est présenté dans l'annexe E et la façon dont elles sont évaluées sera expliquée au chapitre suivant.

**Tableau 6.5** – Résumé des erreurs théoriques sur le nombre d'événements de signal attendus pour l'analyse des données à  $\sqrt{s}=7$  TeV et à  $\sqrt{s}=8$  TeV. Les valeurs présentées sont relatives à un boson de Higgs de masse 125 GeV.

Sources	$\sqrt{s} = 7$ TeV %	$\sqrt{s} = 8$ TeV %
<b>Nbr d'événements de signal</b>		
Section efficace (échelle)	$gg \rightarrow H : \begin{smallmatrix} +12 \\ -8 \end{smallmatrix}$ , VBF : $\pm 0.3$ , WH : $\begin{smallmatrix} +0.2 \\ -0.8 \end{smallmatrix}$ , ZH : $\begin{smallmatrix} +1.4 \\ -1.6 \end{smallmatrix}$ , ttH : $\begin{smallmatrix} +3 \\ -9 \end{smallmatrix}$	$gg \rightarrow H : \begin{smallmatrix} +7 \\ -8 \end{smallmatrix}$ , VBF : $\pm 0.2$ , WH : $\begin{smallmatrix} +0.2 \\ -0.6 \end{smallmatrix}$ , ZH : $\begin{smallmatrix} +1.6 \\ -1.5 \end{smallmatrix}$ , ttH : $\begin{smallmatrix} +4 \\ -9 \end{smallmatrix}$
	$gg \rightarrow H + 2$ jets : $\pm 25$	
Section efficace (PDF+ $\alpha_S$ )	$gg \rightarrow H : \begin{smallmatrix} +8 \\ -7 \end{smallmatrix}$ , VBF : $\begin{smallmatrix} +2.5 \\ -2.1 \end{smallmatrix}$ , VH : $\pm 3.5$ , ttH : $\pm 9$	$gg \rightarrow H : \begin{smallmatrix} +8 \\ -7 \end{smallmatrix}$ , VBF : $\begin{smallmatrix} +2.6 \\ -2.8 \end{smallmatrix}$ , VH : $\pm 3.5$ , ttH : $\pm 8$
Rapport d'embranchement [80]	$\pm 5$	

### 6.2.5 Tests statistiques

L'approche fréquentiste a été préférée pour la recherche du boson de Higgs dans ATLAS. Elle commence généralement avec la définition d'un test statistique afin de quantifier le degré de compatibilité entre les données et l'hypothèse testée. Dans la recherche du boson de Higgs, deux hypothèses sont considérées. L'hypothèse nulle  $H_0$ , dans laquelle les données sont décrites uniquement par le bruit de fond, et l'hypothèse  $H_1$  dans laquelle les données sont constituées de bruit de fond et de signal. Dans notre cas, le paramètre d'intérêt est la "force" du signal  $\mu$ .  $\mu$  est défini comme  $\mu = N_s/N_s^{pred}$ , où  $N_s^{pred}$  est le nombre d'événements de signal prédit par le Modèle Standard. L'hypothèse  $H_0$  est donc caractérisée par  $\mu = 0$  et l'hypothèse  $H_1$  par  $\mu > 0$ .

Le test statistique ici choisi est le test du rapport de vraisemblance ou *Profile Likelihood ratio* basé sur le lemme de Neyman-Pearson [145]. La fonction de vraisemblance non-binnée ou *unbinned likelihood function* [146], [147] est définie telle que :

$$\mathcal{L}(\mu, \theta) = e^{-(\mu N_s^{pred}(\theta_s^{num}) + N_b)} \prod_{k=1}^n \mu N_s^{pred}(\theta_s^{num}) f_s(x_k; m_h, \theta_s) + N_b f_b(x_k; \theta_b) \quad (6.2)$$

Avec :

- $n$  : nombre total d'événements .
- $f_s$  et  $f_b$  : fonctions de densité de probabilité ou pdf (pour *Probability Density Fonction*) pour le signal et le bruit de fond pour la variable  $x$  (typiquement  $m_{\gamma\gamma}$ ).

- $N_b$  : nombre d'événements de bruit de fond.
- $\mu N_s^{pred}$  : correspond au nombre d'événements de signal,  $N_s^{pred}$  est le nombre d'événements prédits par le Modèle Standard et  $\mu$  est le paramètre d'intérêt du fit.
- $\theta_s$  et  $\theta_b$  : paramètres restant dans les fonctions de densité de probabilité du signal et du bruit de fond, il sont appelés également paramètres de nuisance.
- $\theta_s^{num}$  : paramètre de nuisance associé au nombre d'événements de signal.

Le rapport de vraisemblance utilisé au LHC [91] est défini par :

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\theta})}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (6.3)$$

$\hat{\mu}$  et  $\hat{\theta}$  désignent les valeurs de  $\mu$  et  $\theta$  pour lesquelles la fonction de vraisemblance  $\mathcal{L}(\mu, \theta)$  est maximale, ils sont appelés *estimateur du maximum de vraisemblance* ou MLE (pour *maximum likelihood estimator*) de  $\mu$  et  $\theta$ . Quant-à  $\hat{\theta}$ , il s'agit du MLE de  $\theta$  lorsque  $\mu$  est fixé dans le fit. Pour tester  $\mu = 0$  et établir la découverte, on construit le test statistique suivant [148] :

$$q_0 = \begin{cases} -2\ln\lambda_0(\mu) & \hat{\mu} \geq 0 \\ +2\ln\lambda_0(\mu) & \hat{\mu} < 0 \end{cases} \quad (6.4)$$

avec :

$$\lambda_0(\mu) = \frac{\mathcal{L}(\mu = 0, \hat{\theta}(\mu = 0))}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \quad (6.5)$$

Si  $\hat{\mu}$  augmente,  $q_0$  augmente également, signifiant que les données ne sont pas compatibles avec l'hypothèse de bruit de fond uniquement. La condition pour  $\hat{\mu} < 0$  est imposée pour tester les fluctuations négatives du bruit de fond, même si celles-ci ne correspondent pas à la présence d'un signal. Afin de quantifier le niveau de désaccord entre les données et l'hypothèse  $H_0$ , on définit la valeur-p (ou *p-value*) :

$$p_0 = \int_{q_0, obs}^{\infty} f(q_0|0) dq_0 \quad (6.6)$$

$f(q_0|0)$  est la pdf de la statistique  $q_0$ , d'après le théorème de Wilks elle suit une distribution de  $\chi^2$  avec un degré de liberté [149], elle est illustrée figure 6.14.

Suivant la définition de  $q_0$  :  $p_0 < 0.5$  pour  $\hat{\mu} > 0$  et  $0.5 < p_0 < 1$  pour  $\hat{\mu} < 0$ . La valeur-p donne la probabilité que l'excès d'événements observé dans les données (ou un excès plus grand encore) soit compatible avec l'hypothèse  $H_0$  selon laquelle les données ne sont constituées que de bruit de fond. L'hypothèse ici considérée  $H_0$  est donc exclue en dessous d'une certaine valeur de  $p_0$ . La valeur-p est convertie en signficance locale suivant la formule :

$$Z = \Phi^{-1}(1 - p) \quad (6.7)$$

Où  $\Phi$  est la fonction de répartition de la distribution Gaussienne. Afin d'établir la découverte, la signficance doit être au moins égale à 5, ce qui correspond à une valeur-p maximum de  $2.87 \times 10^{-7}$ .

Il est par ailleurs nécessaire de connaître la valeur  $p_0$  attendue dans l'hypothèse du modèle standard (pour  $\hat{\mu} = 1$ ). Pour le  $p_0$  attendu,  $q_0, obs$  devient  $q_0, exp$  dans l'équation 6.6.

$q_0, exp$  correspond à la médiane de la distribution  $f(q_0|1)$  obtenue à partir de pseudo-expériences bruit de fond + signal [134], illustrée figure 6.15.

Toutefois, la génération des pseudo-expériences, nécessaires à la détermination du  $p_0$  attendu, s'avère généralement longue. Une autre approche est ici choisie, un échantillon "Asimov", c'est-à-dire un échantillon parfait sans fluctuations statistiques [148] est préféré.

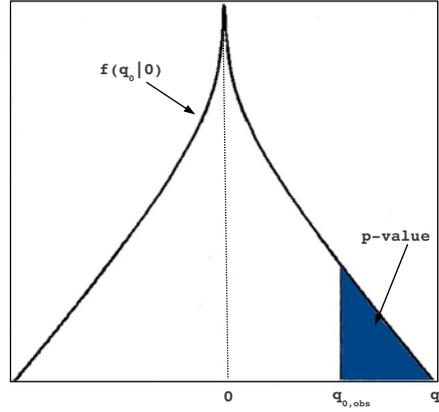


Figure 6.14 – Illustration de la distribution de  $f(q_0|0)$  et relation avec le  $p_0$ .

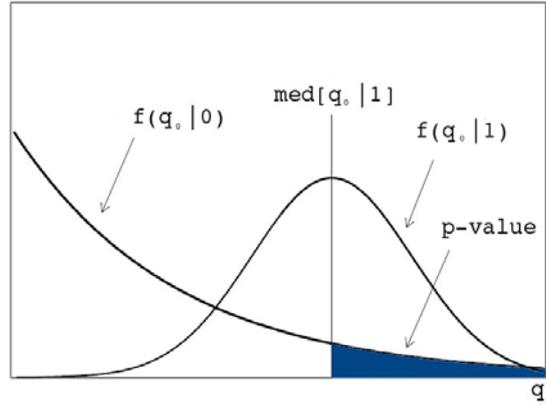


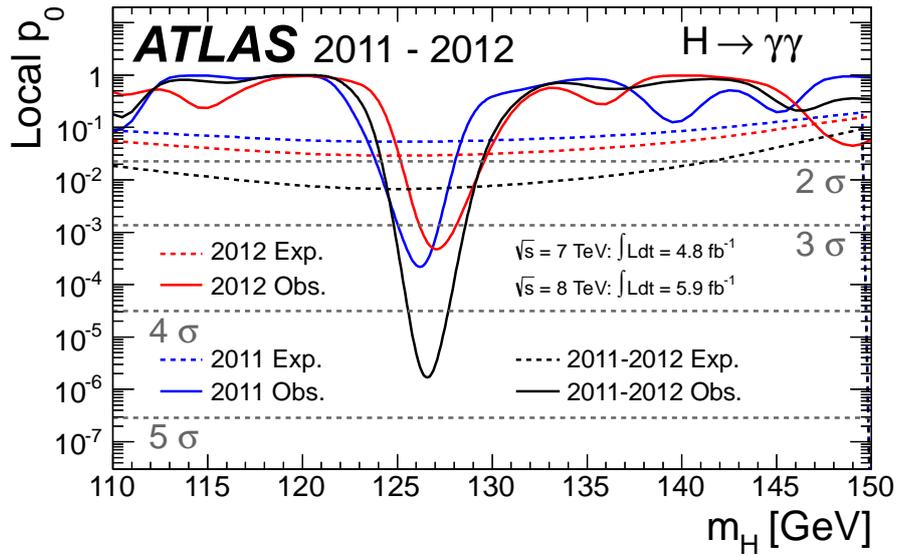
Figure 6.15 – Illustration du  $q_{0,exp}$  correspondant à la médiane de  $q_0$  pour  $\mu'=1$  [148].

### 6.3 Observation d'une nouvelle résonance dans le canal diphoton

La figure 6.16 présente les valeurs de  $p_0$  observée et attendue en fonction de la masse pour l'ensemble des données 2011 et  $5.9 \text{ fb}^{-1}$  de données à  $\sqrt{s}=8 \text{ TeV}$ . Le  $p_0$  observé est de  $2 \times 10^{-6}$  pour  $m_H$  proche de  $126.5 \text{ GeV}$ , ce qui correspond à une signification local de  $4.5\sigma$  (figure 6.16). La valeur préférée par le fit pour la force du signal est  $\hat{\mu} = 1.9 \pm 0.5$  pour  $m_H=126.5 \text{ GeV}$ . Il s'agit de l'observation d'un boson compatible avec le boson de Higgs prédit par le Modèle Standard dans le canal diphoton au LHC.

La recherche du boson de Higgs dans ATLAS ne se limite toutefois pas au canal diphoton. L'observation dans le canal diphoton est confirmée dans le canal  $H \rightarrow ZZ^{(*)} \rightarrow 4l$ , avec un excès de  $3.6\sigma$  à  $m_H=125 \text{ GeV}$ . L'analyse dans le canal  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$  révèle, quant à elle, un excès de  $2.8\sigma$  à  $m_H=125 \text{ GeV}$ . La combinaison des canaux amène à l'observation d'un excès d'événements à  $m_H=126.5 \text{ GeV}$  avec une signification local de  $5.1 \sigma$  [132].

Parallèlement à ATLAS, la collaboration CMS a également mené la recherche du boson de Higgs. La combinaison des canaux  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow ZZ^{(*)} \rightarrow 4l$ ,  $H \rightarrow \tau\tau$ ,  $H \rightarrow b\bar{b}$  et  $H \rightarrow WW^{(*)}$  dans CMS amène à l'observation d'un excès dont la signification locale est de  $5.0$



**Figure 6.16** –  $p_0$  observée et attendue en fonction de la masse hypothétique du boson de Higgs pour  $4.8 \text{ fb}^{-1}$  de données à 7 TeV et  $5.9 \text{ fb}^{-1}$  de données à 8 TeV [132]

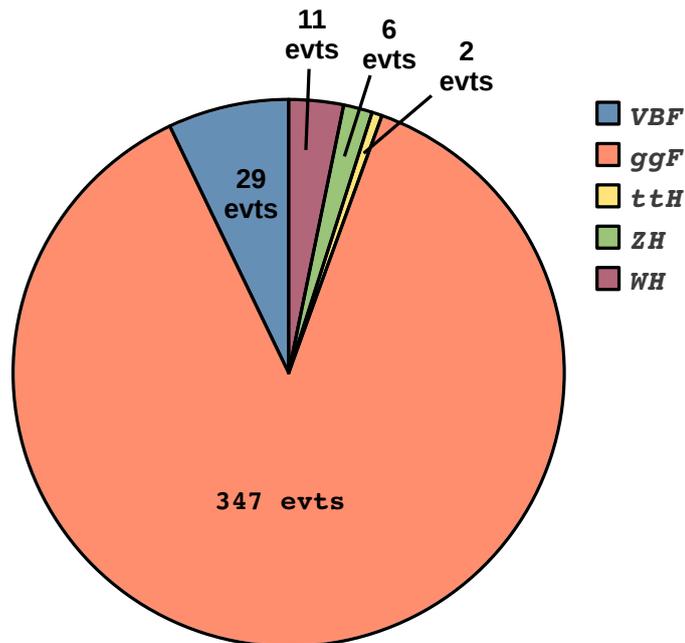
$\sigma$  à  $m_H \sim 125.5 \text{ GeV}$  [150].

Ces résultats amènent ATLAS et CMS à annoncer conjointement le 4 Juillet 2012 la découverte d'un nouveau boson.

## Chapitre 7

# Etudes des couplages de la nouvelle résonance

Cette nouvelle résonance découverte, il est maintenant nécessaire d'étudier ses propriétés. Le boson de Higgs est produit via quatre processus dominants au LHC : la fusion de gluon ( $ggF$ ), la fusion de boson vecteurs ( $VBF$ ), la production associée au boson vecteur  $Z$  et  $W$  ( $VH$ ) et la production associée à une paire de quarks top ( $ttH$ ) (voir chapitre 1). La figure 7.1 présente le nombre d'événements de signal attendus par mode de production pour l'ensemble des données à 8 TeV, soit  $20.7 \text{ fb}^{-1}$ . Dans le chapitre précédent, on a défini la force du signal  $\mu$ , correspondant aux nombres d'événements de signal observés sur le nombre d'événements prédits. De la même façon, on définit la force du signal propre à chaque mode de production. Pour  $i \in \{ggF, VBF, VH, ttH\}$ , on définit  $\mu_i = N_{s,i}/N_{s,i}^{pred}$ . Si auparavant, l'augmentation de la signification globale était le principal but de l'analyse, on se concentre dorénavant sur la précision avec laquelle on mesure les  $\mu_i$ , afin d'évaluer au mieux d'éventuelles tensions entre les sections efficaces de production mesurées et celles prédites par le Modèle Standard. Dans cette optique la classification des événements est revisitée afin d'isoler au mieux les différents modes de production.



**Figure 7.1** – Nombres prédits de bosons de Higgs produits et se désintégrant en deux photons pour  $20.7 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$ .

Dans ce chapitre, nous verrons en premier lieu la sélection des événements diphoton, qui a été améliorée par rapport à celle du chapitre précédent, et la composition du bruit de fond pour l'ensemble des données à 8 TeV. Nous détaillerons ensuite les nouvelles catégories introduites ou redéfinies afin d'apporter plus de précision sur la mesure des  $\mu_i$  et plus particulièrement du  $\mu_{VBF}$  :

- les catégories VH : elles sont basées sur la sélection des objets supplémentaires (leptons, énergie transverse manquante, jets) résultant de la désintégration des bosons  $Z$  et  $W$  produits en association avec le boson de Higgs afin de distinguer les événements VH du bruit de fond QCD.
- les catégories VBF : elles sont basées sur la sélection des deux jets produits en association avec le boson de Higgs. Deux méthodes distinctes sont utilisées pour définir ces catégories et séparer les événements VBF du bruit de fond QCD mais également du bruit de fond constitué par les événements ggF avec deux jets.

Enfin, nous nous attarderons sur la procédure statistique utilisée et les sources d'erreurs systématiques considérées avant de présenter les résultats pour l'ensemble des données 2011 et 2012. J'ai été principalement impliquée dans la définition de la catégorie VBF et dans l'implémentation des erreurs systématiques dans la procédure statistique, ces parties seront donc plus particulièrement développées.

## 7.1 Sélection des données, mars 2013

### 7.1.1 Sélection inclusive des événements

La sélection des événements diphoton pour les données 2011 à  $\sqrt{s} = 7$  TeV reste inchangée mais la sélection des événements diphoton pour l'analyse  $H \rightarrow \gamma\gamma$  des données 2012 à  $\sqrt{s} = 8$  TeV de mars 2013 diffère de celle de juillet 2012. Le trigger est le même, deux candidats photons LOOSE avec  $p_T^{\gamma 1} > 35$  GeV et  $p_T^{\gamma 2} > 25$  GeV sont requis (efficacité :  $99.6 \pm 0.1\%$  (stat)), la luminosité totale correspondante aux données 2012 est de  $20.7 \text{ fb}^{-1}$ . La zone exclue en  $\eta$  correspondante à la transition entre les régions tonneau et bouchons est maintenant  $1.37 < |\eta| < 1.56$  afin de se dédouaner de la mesure d'énergie des photons basée sur les scintillateurs [151]. La sélection du vertex primaire a été améliorée, un Neural Network remplace la méthode d'ajustement minimale, amenant ainsi une plus grande probabilité (+6%) d'identification de la position en  $z$  du vertex [152]. Enfin, le critère d'isolation a été modifié afin de maintenir une bonne efficacité d'isolation malgré les conditions d'empilement [153]. La coupure sur l'énergie d'isolation calorimétrique a été relâchée, de 4 à 6 GeV et elle est combinée avec une coupure sur l'énergie d'isolation des traces, qui doit être inférieure à 2.6 GeV (voir section 4.3.2).

### 7.1.2 Composition du bruit de fond

La décomposition du bruit de fond pour l'ensemble des données 2012 est obtenue à partir de la méthode 2x2D sideband. Les événements sont répartis tels que :

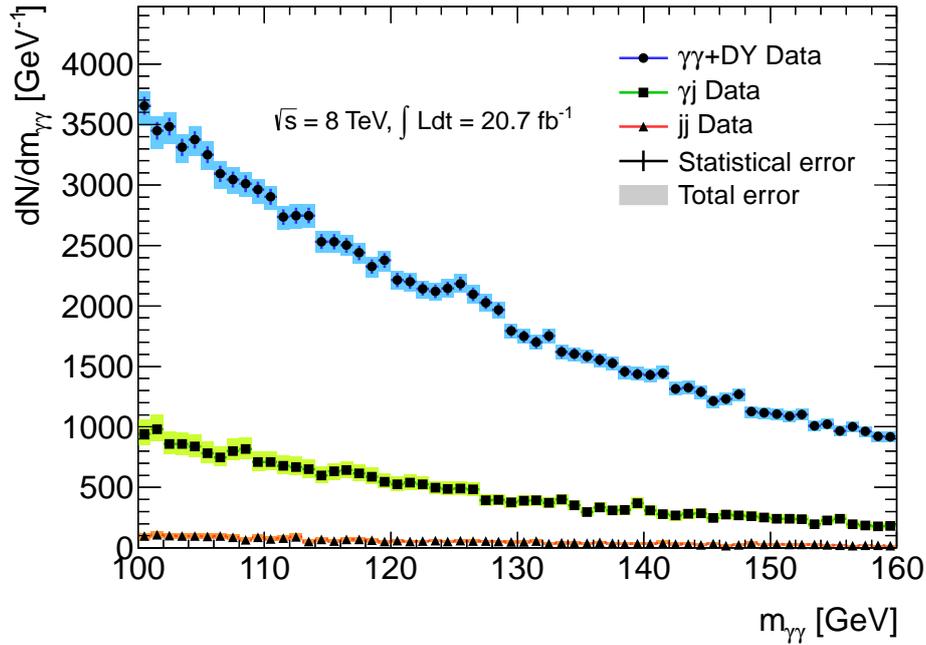
$$\begin{aligned}
 - N_{\gamma\gamma} &= 90364 \pm 440 \text{ (stat)} \begin{matrix} +2716 \\ -3032 \end{matrix} \text{ (syst)} \\
 - N_{\gamma j} &= 18299 \pm 198 \text{ (stat)} \begin{matrix} +2058 \\ -2028 \end{matrix} \text{ (syst)} \\
 - N_{j\gamma} &= 6297 \pm 146 \text{ (stat)} \begin{matrix} +758 \\ -444 \end{matrix} \text{ (syst)} \\
 - N_{jj} &= 3233 \pm 65 \text{ (stat)} \begin{matrix} +462 \\ -585 \end{matrix} \text{ (syst)}
 \end{aligned}$$

Les erreurs systématiques sont détaillées tableau 7.1. La décomposition du spectre de masse invariante est présentée figure 7.2. La pureté de l'échantillon sélectionné pour l'ensemble des données 2012 est de  $76.5 \pm 0.3$  (stat)  $\pm 2.6$  (syst)% et est cohérente avec celle obtenue pour le

sous-ensemble de  $5.9 \text{ fb}^{-1}$  de données à  $\sqrt{s}=8 \text{ TeV}$ .

Sources d'erreurs systématiques	$N_{\gamma\gamma}$	$N_{\gamma j}$	$N_{j\gamma}$	$N_{jj}$
Valeur centrale	90364	18299	6297	3233
Définition de la région de contrôle du bruit de fond	+2040 -2445	+1406 -1361	+622 -100	+391 -531
Efficacité isolation	$\pm 1355$	$\pm 1485$	$\pm 425$	$\pm 242$
Efficacité identification	$\pm 1175$	$\pm 238$	$\pm 82$	$\pm 42$
Erreur syst. totale	+2716 -3032	+2058 -2028	+758 -444	+462 -585

**Tableau 7.1** –  $20.7 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$  (2012) : Nombres d'événements totaux par composante pour l'analyse  $H \rightarrow \gamma\gamma$  obtenus avec la méthode  $2 \times 2D$  sideband. Les différentes erreurs systématiques sont détaillées par composante, l'erreur systématique totale correspond à la somme quadratique des différentes contributions.



**Figure 7.2** – Spectre différentiel  $m_{\gamma\gamma}$  pour les contributions  $N_{\gamma\gamma}$ ,  $N_{\gamma j+j\gamma}$  et  $N_{jj}$  dans l'analyse  $H \rightarrow \gamma\gamma$  pour  $20.7 \text{ fb}^{-1}$  à  $\sqrt{s} = 8 \text{ TeV}$  (2012).

## 7.2 Catégories VH

Les catégories VH sont introduites dans l'analyse afin d'augmenter la sensibilité à la production du boson de Higgs associé à un boson vecteur et ainsi de diminuer l'erreur statistique sur la mesure de  $\mu_{VH}$  [152], [151]. Le nombre d'événements VH attendus est très faible il est donc primordial d'optimiser les coupures suivant les caractéristiques des produits de désintégrations des bosons vecteurs  $Z$  et  $W$  produits en association avec le boson de Higgs afin de rejeter le bruit de fond principal, le bruit de fond QCD.

### 7.2.1 Catégorie leptonique

Les désintégrations leptoniques des bosons  $W$  et  $Z$  fournissent une topologie claire pour distinguer la production du boson de Higgs associée à un boson vecteur. La catégorie leptonique définie en sélectionnant les événements qui contiennent au moins un candidat lepton a pour but de capturer les événements  $WH \rightarrow \ell\nu\gamma\gamma$  et  $ZH \rightarrow \ell\ell\gamma\gamma$ .

Les candidats électrons sont reconstruits à partir des clusters EM comme décrit section 4.1.1. Une énergie transverse supérieure à 15 GeV et  $|\eta| < 2.47$  sont requis. Les critères d'identification appliqués aux électrons sont proches de ceux appliqués aux photons, ils sont décrits dans [103]. Les candidats muons sont reconstruits à partir des traces du détecteur interne et du spectromètre à muons [90]. Pour la région en dehors de la zone d'acceptance couverte par le trajectographe,  $2.5 < |\eta| < 2.7$ , seul le spectromètre à muon est utilisé. Les candidats muons doivent, par ailleurs, avoir une impulsion transverse supérieure à 10 GeV.

Afin d'améliorer la réjection du bruit de fond et l'efficacité du signal, les coupures d'isolation dépendent du  $p_T$  (ou  $E_T$ ) [154] :

- $E_T^{iso,trace}/p_T < 0.15$  pour les muons et  $E_T^{iso,trace}/E_T < 0.15$  pour les électrons.
- $E_T^{iso,calo}/p_T < 0.2$  pour les muons et  $E_T^{iso,calo}/E_T < 0.2$  pour les électrons.

Dans le but de réduire la contamination des événements  $Z(ee)\gamma$ , où l'un des électrons est identifié comme un photon, une coupure sur la masse reconstruite du candidat électron et de l'un des candidats photon est appliquée, l'intervalle de masse  $84 < |M_{e\gamma}| < 94$  GeV est exclu. En outre, si l'un des candidats photons passe les coupures d'identification pour les électrons, l'événement n'est pas retenu.

### 7.2.2 Catégorie avec énergie transverse manquante

Un boson de Higgs produit en association avec un boson vecteur se désintégrant en lepton peut être caractérisé par une énergie manquante significative due à la présence d'un neutrino dans l'état final. Afin de sélectionner les événements  $ZH \rightarrow \nu\nu\gamma\gamma$  mais aussi  $WH \rightarrow \ell\nu\gamma\gamma$ , une catégorie avec énergie transverse manquante,  $E_T^{miss}$ , est définie.

La mesure de l'énergie transverse manquante est basée sur l'énergie transverse de tous les candidats photons, électrons et muons, de tous les jets et de tous clusters EM non-associés à ces objets.

La résolution de  $E_T^{miss}$  est proportionnelle à l'énergie transverse totale. Pour prendre en compte cette dépendance et ainsi identifier la présence d'énergie manquante de fluctuation dues à la résolution, on définit la signification de l'énergie manquante par [155] :

$$S_{E_T^{miss}} = \frac{E_T^{miss}}{0.67\sqrt{\sum E_T}} \quad (7.1)$$

La catégorie  $E_T^{miss}$  est définie avec les événements pour lesquels  $S_{E_T^{miss}} > 5$ . Par ailleurs, les événements avec au moins un photon passant les coupures d'identification électron sont rejetés pour réduire le bruit de fond provenant du processus  $W \rightarrow e\nu\gamma$  où l'électron est identifié comme un photon.

### 7.2.3 Catégorie hadronique

Une catégorie appelée "2 jets à basse masse" est définie pour capturer les événements VH pour lesquels le boson vecteur se désintègre en hadrons.

La sélection appliquée sur les jets est similaire à celle appliquée pour l'analyse de juillet 2012 et décrite section 6.2.1. En raison de l'augmentation des événements d'empilement, la coupure

sur la JVF permettant de rejeter les événements avec des jets reconstruits dus à l'empilement a été modifiée, elle est telle que  $JVF > 0.25$  (0.5 précédemment).

Afin de sélectionner les événements avec deux jets issus de la désintégration d'un boson vecteur, et ainsi de diminuer le bruit de fond QCD, une coupure sur la masse invariante du système dijet,  $60 < m_{jj} < 110$  GeV, ainsi qu'une coupure sur l'écart en  $|\eta|$  entre les deux jets,  $|\Delta\eta_{jj}| < 3.5$ , sont appliquées. En outre, la différence en  $|\eta|$  entre le système diphoton et le système dijet doit être inférieure à 1 et le  $p_{Tt}$  du système diphoton doit être supérieur à 70 GeV [154, 151].

## 7.3 Catégorie VBF

Après la fusion de gluon, la fusion de boson vecteur est le deuxième mode de production dominant du boson de Higgs au LHC. La catégorie VBF, déjà définie pour l'analyse de juillet 2012 (section 6.2.1) a été optimisée après la découverte, dans le but cette fois de diminuer l'erreur statistique sur la mesure de  $\mu_{VBF}$  et non d'augmenter la signification du signal. Outre l'incertitude sur  $\mu_{VBF}$ , il est également essentiel de conserver une bonne pureté en signal VBF (définie comme le nombre d'événements VBF sur le nombre d'événements total) dans la catégorie VBF. L'incertitude liée à l'échelle d'énergie sur la section efficace de production d'un boson de Higgs ggF avec deux jets est en effet l'erreur systématique dominante dans la catégorie VBF. L'optimisation des coupures pour cette catégorie doit donc permettre une bonne réjection du bruit de fond QCD, comme pour les catégories VH, mais également une bonne réjection des événements ggF + 2 jets.

Deux optimisations ont été menées en parallèle, l'une est basée sur la définition de coupures, l'autre sur l'utilisation d'une analyse multivariée. Dans les deux cas, les études ont été faites avec  $13.0 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8$  TeV.

### 7.3.1 Optimisation des coupures

Afin d'optimiser les coupures, l'ensemble de la procédure statistique décrite section 7.5 est utilisée. L'erreur attendue sur le  $\mu_{VBF}$  pour  $m_H = 125$  GeV est obtenue par un fit simultané sur les différentes catégories.

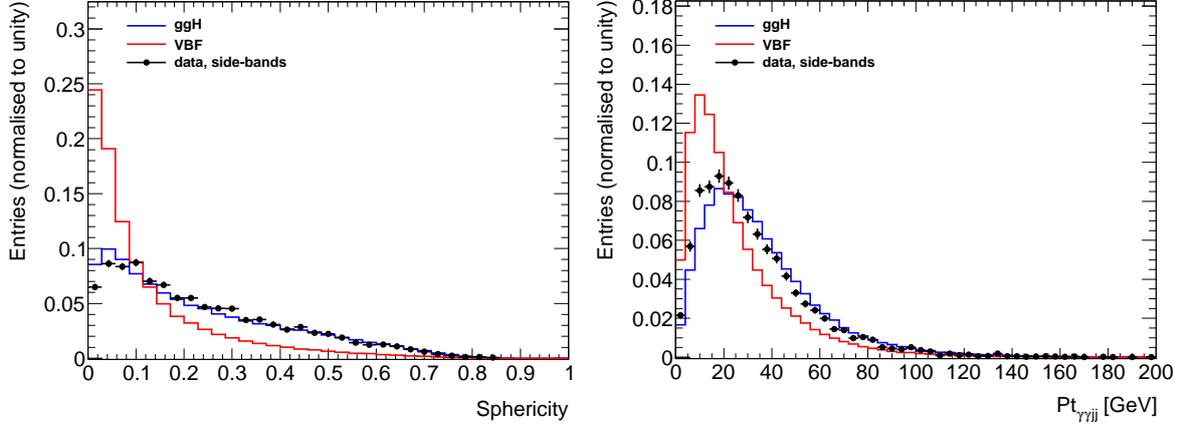
Trois principales options ont été étudiées :

- Conserver uniquement les variables utilisées pour l'analyse de juillet 2012 ( $m_{jj}$ ,  $\Delta\eta_{jj}$ ,  $\Delta\phi_{\gamma\gamma-jj}$ ), mais redéfinir les coupures.
- Introduire de nouvelles variables.
- Définir deux catégories VBF : l'une avec des coupures dures et une très bonne pureté, elle sera par la suite appelée "tight", l'autre avec des coupures plus lâches et une pureté plus faible, elle sera par la suite appelée "loose".

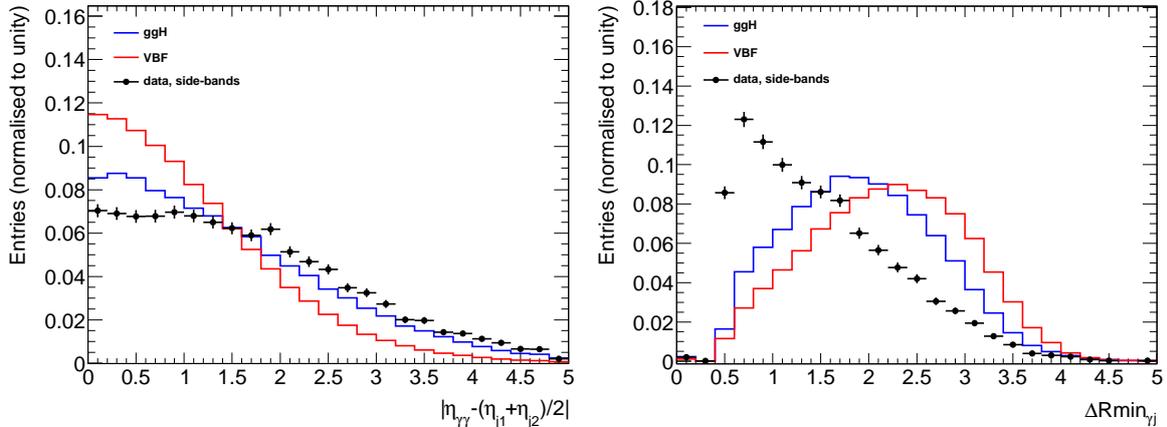
Les coupures pour la catégorie VBF de l'analyse de juillet 2012 étaient telles que :  $m_{jj} > 400$  GeV,  $\Delta\phi_{\gamma\gamma-jj} > 2.6$  et  $\Delta\eta_{jj} > 2.8$ ; ce jeu de coupures sera par la suite appelé "standard". Une coupure plus dure sur  $\Delta\phi_{\gamma\gamma-jj}$  n'est pas envisageable car cela impliquerait une augmentation significative des incertitudes liées aux échelles d'énergie sur la section efficace de production d'un boson de Higgs ggF avec deux jets [156], [157]. En conservant une coupure  $\Delta\phi_{\gamma\gamma-jj} > 2.6$ , l'erreur sur  $\mu_{VBF}$  est minimisée pour  $m_{jj} > 700$  GeV et  $\Delta\eta_{jj} > 3.6$ . Ce point sera par la suite appelé "point 2".

Plusieurs nouvelles variables ont été testées pour la catégorie VBF, comme la sphéricité ou l'impulsion transverse du système dijet-diphoton,  $p_{T\gamma\gamma jj}$  (figure 7.3), mais les deux variables les plus discriminantes sont  $\Delta\eta^* = |\eta_{\gamma\gamma} - (\eta_{jet1} + \eta_{jet2})/2|$  [158] et  $\Delta R_{min}$  qui correspond à la plus petite distance entre un photon et un jet. La distribution de ces variables pour le processus ggF, VBF et pour les données sideband (100-120 et 130-160 GeV) sont présentées figure 7.4.

La variable  $\Delta\eta^*$  pique à zéro dans le cas du processus VBF, les photons sont attendus entre les jets, comme illustré figure 6.7. Avec ces nouvelles variables, le meilleur jeu de coupures, appelé par la suite "point 3" est :  $\Delta\eta_{jj} > 2.8$ ,  $m_{jj} > 520$  GeV,  $\Delta\phi_{\gamma\gamma-jj} > 2.6$ ,  $\Delta\eta^* < 2.4$  et  $\Delta R_{min} > 2.0$ .



**Figure 7.3** – Distributions normalisées à l'unité de la sphéricité (à gauche) et de l'impulsion transverse du système dijet-diphoton  $p_{T,\gamma jj}$  (à droite) pour les événements VBF, ggF et pour les données side-bands (100-120 et 130-160 GeV). La proportion de signal incluse dans les données side-bands est négligeable, leur distribution est donc interprétée comme celle du bruit de fond QCD.



**Figure 7.4** – Distributions normalisées à l'unité de la variable  $\Delta\eta^*$  (à gauche) et de  $\Delta R_{min}$  (à droite) pour les événements VBF, ggF et pour les données side-bands (100-120 et 130-160 GeV). La proportion de signal incluse dans les données side-bands est négligeable, leur distribution est donc interprétée comme celle du bruit de fond QCD.

La pureté en signal VBF dans la catégorie VBF est comprise entre 75 et 80%. Les événements VBF restants sont répartis entre les neuf catégories définies en fonction du  $p_{T,t}$ , de  $\eta$  et du statut de conversion. Si, en plus de cette catégorie VBF "tight", l'on crée une seconde catégorie VBF "loose" avec des coupures plus lâches que la première, une pureté VBF supérieure à celle des neuf catégories restantes ( $p_{T,t}$ ,  $\eta$ , statut de conversion) peut être atteinte et permet de contraindre un peu plus la composante VBF dans le fit. Pour l'analyse VBF basée sur les coupures, la catégorie loose VBF correspond à la catégorie VBF définie pour l'analyse de juillet :  $m_{jj} > 400$  GeV,  $\Delta\phi_{\gamma\gamma-jj} > 2.6$  et  $\Delta\eta_{jj} > 2.8$ .

Le tableau 7.2 présente l'erreur attendue (statistique+systematique),  $\delta\mu_{VBF}$ , sur  $\mu_{VBF}$  pour les différentes catégorisations VBF mentionnées. La meilleure amélioration correspond au point

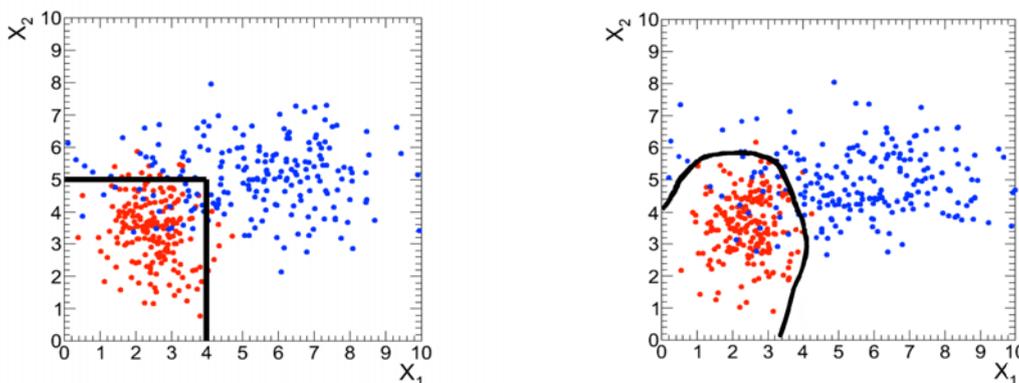
3 avec deux catégories VBF.

		pureté VBF	$\delta\mu_{VBF}$	amélioration rel.
1 catégorie VBF	Standard	69.8	1.40	
1 catégorie VBF	Point 2	80.2	1.29	7.8%
1 catégorie VBF	Point 3	75.7	1.17	16.4%
2 catégories VBF	Point 2	80.2/60.5	1.19	15%
2 catégories VBF	Point 3	75.7/61	1.11	20.7%

**Tableau 7.2** – Pureté en signal VBF dans la/les catégorie(s) VBF et amélioration sur le  $\delta\mu_{VBF}$  (stat+syst) attendu pour chaque scénario d'optimisation des coupures. Les trois catégories VH, ainsi que les neuf catégories usuelles (définies section 6.2.1) sont également considérées.

### 7.3.2 Analyse multi-variée

Les définitions des catégories VBF précédentes sont basées sur des coupures rectangulaires et ne prennent pas en compte les corrélations entre les variables utilisées. L'exploitation de ces corrélations permet d'être plus efficace pour la séparation des événements VBF du bruit de fond. L'analyse multi-variée ou MVA présentée ici permet d'analyser les différentes variables simultanément ; elle utilise l'ensemble des corrélations entre les variables et, dans notre cas, définit des coupures non linéaires comme illustré figure 7.5.



**Figure 7.5** – Illustration de coupures rectangulaires (à gauche) et de coupures non linéaires (à droite) pour discriminer le signal du bruit de fond [159].

La méthode choisie est un arbre de décision "boosté" ou BDT pour *boosted decision tree*. Son implémentation est faite dans la structure *TMVA Toolkit for Multivariate Analysis* documentée dans [109]. Un arbre de décision est un arbre binaire. A chaque étape, ou noeud, un algorithme permet de trouver la meilleure variable et la meilleure coupure sur cette variable afin de distinguer le signal du bruit de fond. On définit la pureté  $p$  comme la fraction d'événements de signal sur le nombre d'événements total. Le critère  $p(1-p)$  est maximum quand signal et bruit sont à mêmes proportions, il est minimal quand la pureté tend vers 1 ou vers 0. L'algorithme mis en jeu a donc pour but de minimiser  $p(1-p)$  à chaque noeud en scannant les variables avec une granularité fixée (ici 30). La construction de l'arbre s'arrête lorsque le nombre minimum d'événements à un noeud est atteint (ici 100) comme illustré figure 7.6. Le noeud devient alors une feuille qui est classée en signal ou bruit de fond. Si une feuille a une pureté supérieure à 50%, elle correspond à du signal.

Le "boosting" étend ce principe à plusieurs arbres (ici 700), constituant ainsi une forêt d'arbres de décision. La variable discriminante ou la réponse du BDT correspond à la moyenne pondérée des arbres individuels [109].

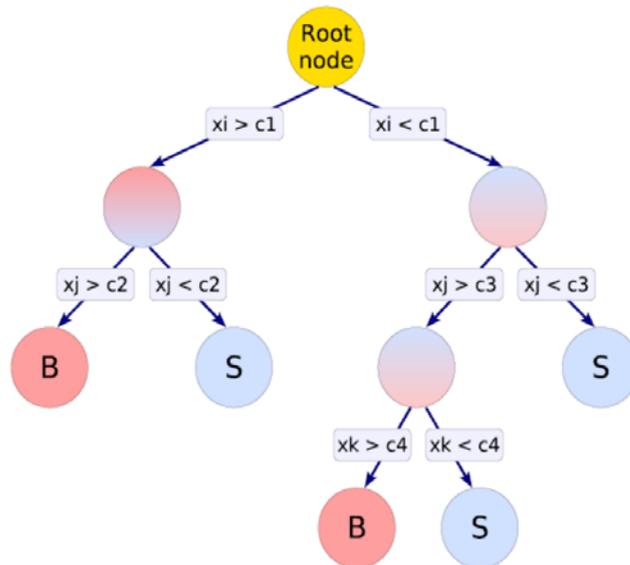


Figure 7.6 – Illustration d'un arbre de décision [109].

Deux échantillons indépendants de données sont nécessaires, l'un pour entraîner les arbres, l'autre pour tester que l'entraînement a convergé vers une solution qui n'est pas dépendante de l'ensemble de données utilisé pour l'entraînement. L'un des principaux risques est le sur-entraînement, le BDT décrit trop bien l'échantillon d'entraînement et prend en compte les fluctuations statistiques, si bien que la performance du BDT est excellente sur l'échantillon d'entraînement mais elle est beaucoup moins bonne sur l'échantillon de test. Le boosting diminue la dépendance de la réponse du BDT en fonction des fluctuations statistiques de l'échantillon d'entraînement.

Dans notre cas, le BDT est entraîné pour distinguer les événements VBF ( $\sim 160000$  événements simulés) du bruit de fond diphoton, constitué de vraies paires de photons (issus de processus durs ou de fragmentation) produit en association avec deux jets, des événements avec un photon direct et au moins trois jets (dont l'un est identifié comme un photon) et des événements multi-jets où au moins deux jets sont identifiés comme des photons. Toutefois, les événements de ce bruit de fond et les événements  $ggF$  avec deux jets présentent des caractéristiques cinématiques proches, exceptés pour le  $p_{Tt}$  et  $\Delta R_{min}$  (figure 7.7 et 7.8) ; la réjection du bruit de fond s'accompagnera ainsi d'une bonne réjection des événements  $ggF$  et permet de maintenir une bonne pureté en signal VBF.

Pour l'entraînement (et le test) du BDT, la composante  $\gamma\gamma$  du bruit de fond est modélisée par un échantillon diphoton SHERPA tandis que les composantes  $\gamma$ -jet et jet-jet sont tirées des données en sélectionnant les événements pour lesquels l'un des deux ou les deux candidats photons ne passent pas la coupure d'isolation ( $\sim 8000$  événements). Les composantes sont ensuite normalisées suivant la décomposition du bruit de fond présentée section 7.1.2.

Les variables utilisées en entrée du BDT ne doivent pas être corrélées à la masse invariante  $m_{\gamma\gamma}$  afin de ne pas créer de biais dans la sélection des événements, c'est à dire de ne pas sélectionner des événements dans une certaine région de masse invariante. Les variables

choisies sont résumées tableau 7.3 et leurs distributions pour les événements VBF, ggF et pour les données sidebands (100-120 et 130-160 GeV) sont présentées figures 7.7 et 7.8. Pour une efficacité de signal de  $\sim 80\%$ , la réjection du bruit de fond atteinte est de l'ordre de  $80\%$ , et les variables présentant le plus grand pouvoir discriminant sont  $m_{jj}$ ,  $\Delta\eta_{jj}$  et  $\Delta R_{min}$ .

Afin de vérifier que l'échantillon choisi pour l'entraînement (et le test) décrit correctement les données, et ainsi s'assurer de l'efficacité du BDT dans les données, les distributions des variables d'entrée pour les données sidebands et pour l'échantillon d'entraînement sont comparées ; elles présentent un accord correct [156].

Variable	Description
$m_{jj}$	Masse invariante du système dijet.
$\Delta\phi_{\gamma\gamma jj}$	Angle azimuthal entre le système diphoton et le système dijet.
$\Delta\eta_{jj}$	Séparation en $\eta$ entre les deux jets.
$\eta_{jet1}$	Pseudorapacité du premier jet.
$\eta_{jet2}$	Pseudorapacité du second jet.
$p_{Tt}$	Composante perpendiculaire à l'axe "principale" de l'impulsion transverse du système diphoton.
$\Delta R_{min}$	Plus petite distance entre un photon et un jet.
$\Delta\eta^*$	Définie par $ \eta_{\gamma\gamma} - (\eta_{j1} + \eta_{j2})/2 $ .

**Tableau 7.3** – Liste et description des variables utilisées pour l'analyse MVA VBF.

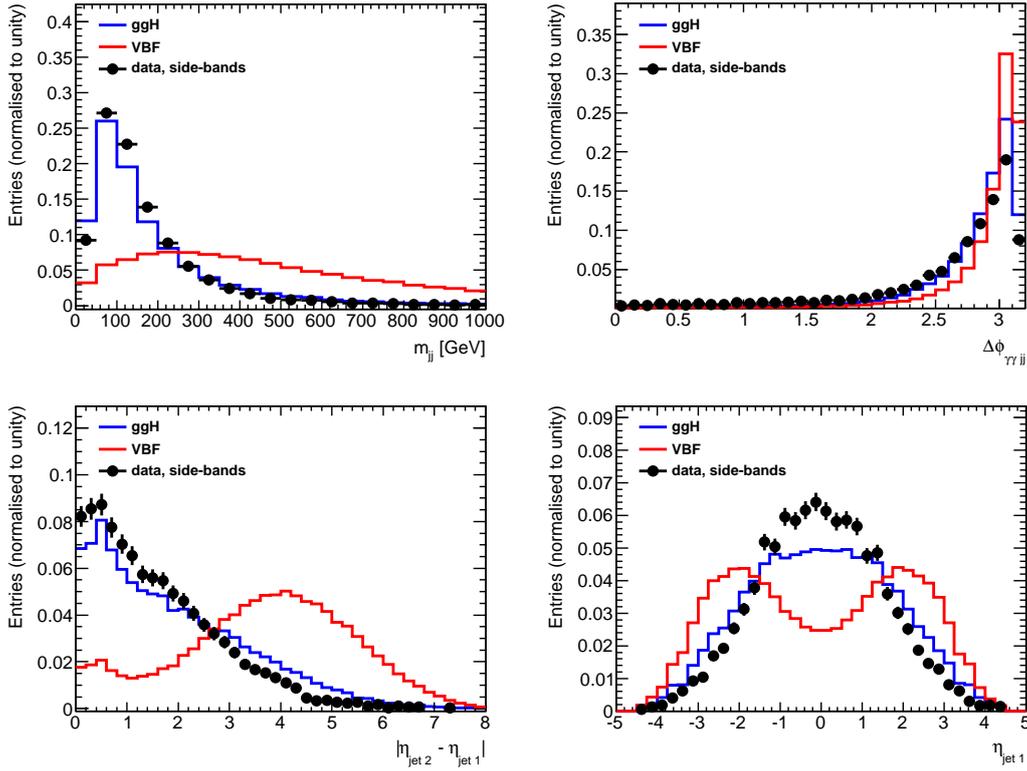
La réponse du BDT sur les données (en excluant la zone de signal) et sur le bruit de fond attendu est présentée figure 7.9. On observe un bon accord entre les réponses du BDT sur les données sidebands et sur le bruit de fond prédit. Le BDT se comporte comme attendu et n'est pas sensible à l'échantillon choisi pour l'entraînement.

La figure 7.10 montre la réponse du BDT en fonction de la masse invariante pour les données sidebands (à gauche) et la forme du spectre  $m_{\gamma\gamma}$  pour différentes coupures sur la réponse du BDT (à droite). Aucune structure n'est observée, aucune zone de masse invariante n'est donc préférée par le BDT. Par ailleurs, la régularité du spectre  $m_{\gamma\gamma}$  quelle que soit la coupure sur la réponse du BDT est une condition nécessaire afin de pouvoir effectuer le fit sur les données pour extraire le nombre d'événements de bruit de fond.

De la même façon que pour l'analyse VBF basée sur les coupures, deux catégories VBF sont définies avec l'analyse MVA en fonction de la coupure appliquée sur la réponse du BDT. Les coupures sur la réponse du BDT sont choisies à partir d'un échantillon statistiquement indépendant de l'échantillon utilisé pour l'entraînement en maximisant la signification du signal VBF définie par :

$$\sigma_{VBF} = \sqrt{2(N_{VBF} + N_{ggF} + N_{background}) \times \ln\left(1 + \frac{N_{VBF}}{N_{ggF} + N_{background}} - N_{VBF}\right)} \quad (7.2)$$

Cette formule est utilisée lorsque le nombre d'événements de signal suit une distribution poissonnienne, elle tend vers  $s/\sqrt{s+b}$  quand le nombre d'événements de signal tend vers une distribution gaussienne. Le nombre d'événements de bruit de fond  $N_{VBF}$ , comme pour l'optimisation des coupures rectangulaires, est obtenu à partir d'un fit sur les données sidebands.



**Figure 7.7** – De gauche à droite, de haut en bas, distributions des variables  $m_{jj}$ ,  $\Delta\phi_{\gamma\gamma-jj}$ ,  $\eta_{jet1}$  et  $\eta_{jet2}$  pour les événements VBF, ggF et pour les données sidebands (100-120 et 130-160 GeV). La proportion de signal incluse dans les données sidebands est négligeable, leur distribution est donc interprétée comme celle du bruit de fond QCD

La catégorie tight contient tous les événements tels que  $BDT \geq 0.74$  et la catégorie loose les événements tels que  $0.44 < BDT < 0.74$ . Les autres événements sont réinjectés dans les neuf catégories restantes ( $p_{Tt}$ ,  $\eta$  et statut de conversion).

La figure 7.11 présente la réponse du BDT pour le bruit de fond et pour les échantillons VBF et ggF. La séparation entre le signal VBF et les autres processus est clairement visible.

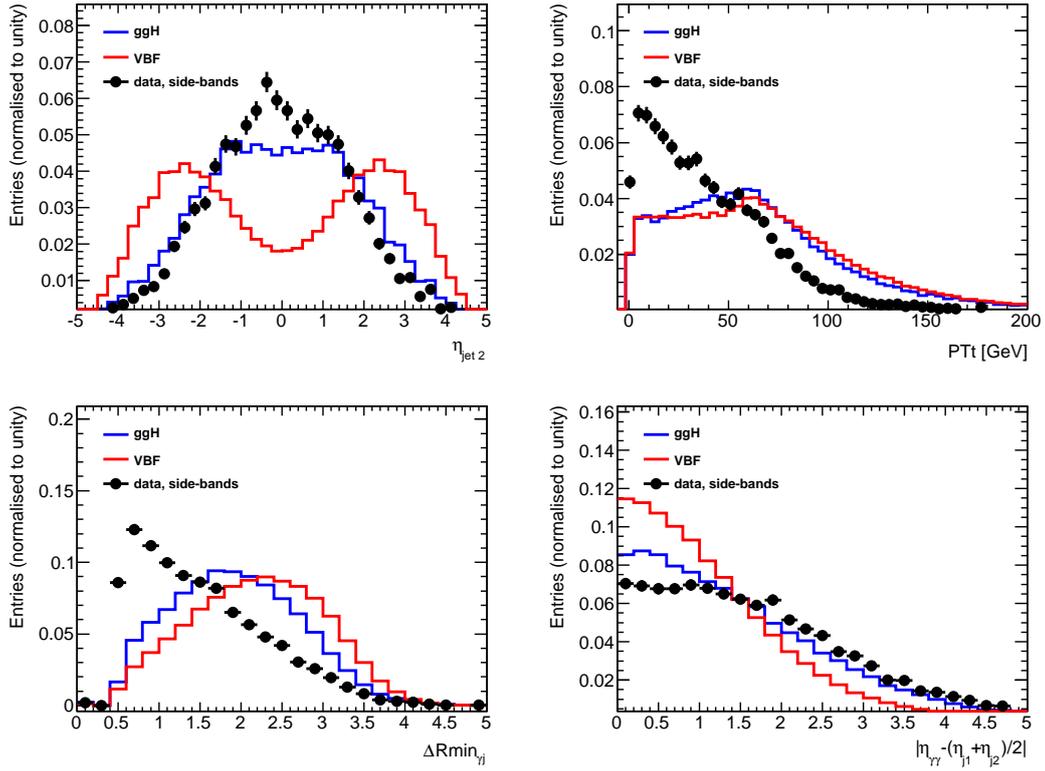
### 7.3.3 Comparaison entre l'analyse multivariée et l'analyse avec coupures

Le tableau 7.4 compare les  $\delta\mu_{VBF}$  attendus dans le cas de l'analyse VBF MVA et de l'analyse VBF basée sur les coupures optimisées.

	pureté VBF %	$\delta\mu_{VBF}$	amélioration rel.
Standard (1 catégorie VBF)	69.8	1.40	
Point 3 (2 catégories VBF)	75.7/61.0	1.11	20.7%
MVA (2 catégories VBF)	77.1/48.0	1.03	26.4%

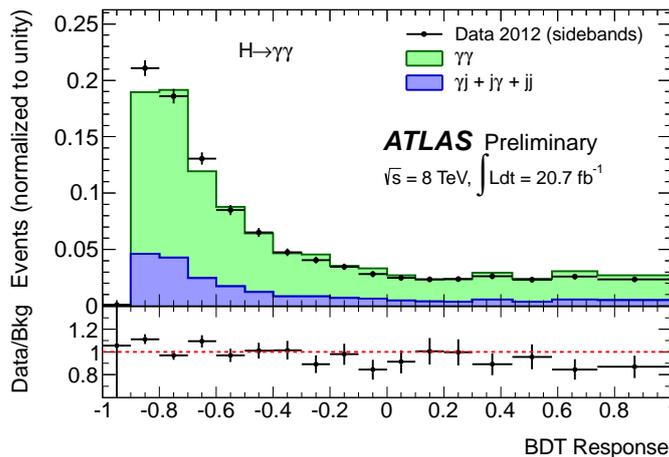
**Tableau 7.4** – Comparaison du  $\delta\mu_{VBF}$  (stat+syst) attendu pour  $13.0 fb^{-1}$  de données à  $\sqrt{s} = 8$  TeV pour l'analyse VBF MVA et pour l'analyse VBF basée sur les coupures optimisées. Les trois catégories VH, ainsi que les neuf catégories usuelles (définies section 6.2.1) sont également considérées.

Parce qu'elle permet d'obtenir une amélioration significative de 26% sur l'erreur de  $\mu_{VBF}$ , l'analyse MVA est choisie et sera utilisée pour tous les résultats détaillés par la suite. L'analyse

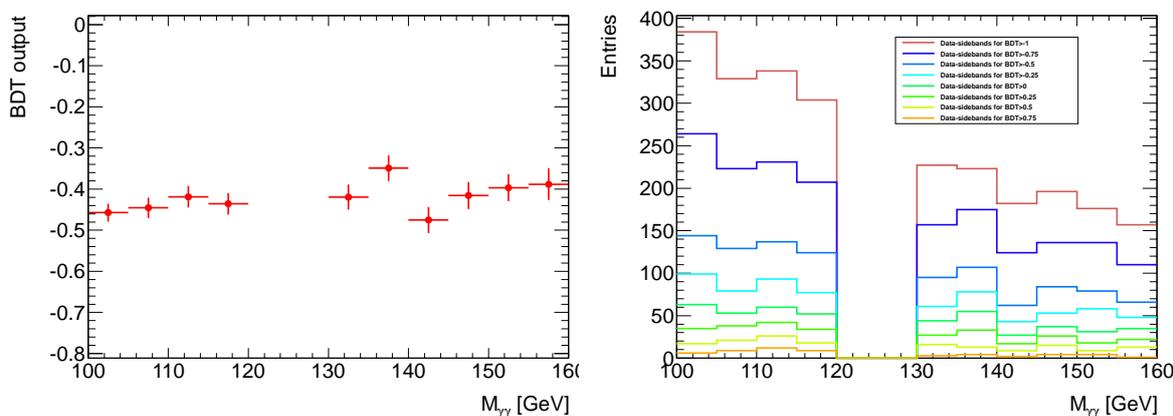


**Figure 7.8** – De gauche à droite, de haut en bas, distributions des variables  $\eta_{jet2}$ ,  $p_{Tt}$ ,  $\Delta R_{min}$  et  $\Delta\eta^*$  pour les événements VBF, ggF et pour les données sidebands (100-120 et 130-160 GeV). La proportion de signal incluse dans les données sidebands est négligeable, leur distribution est donc interprétée comme celle du bruit de fond QCD

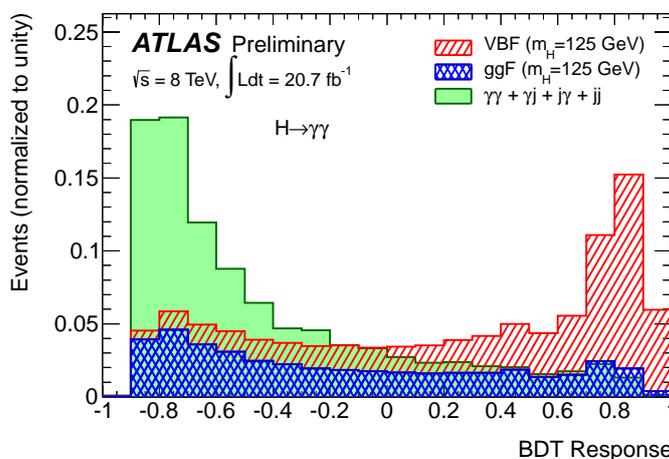
MVA pour la catégorie VBF est appliquée exclusivement aux données 2012.



**Figure 7.9** – Comparaison de la réponse du BDT sur les données sidebands et sur le bruit de fond attendu après les coupures de sélection. Les distributions sont normalisées à 1 [151].



**Figure 7.10** – Distribution de la réponse du BDT en fonction de  $m_{\gamma\gamma}$  pour les données sidebands (à gauche) et distribution de  $m_{\gamma\gamma}$  pour différentes coupures appliquées sur la réponse du BDT (à droite). La région de signal 120-130 GeV est soustraite.



**Figure 7.11** – Réponse du BDT sur l'échantillon de signal VBF, sur l'échantillon ggF et sur le bruit de fond attendu après les coupures de sélection. Les distributions sont normalisées à 1 [151].

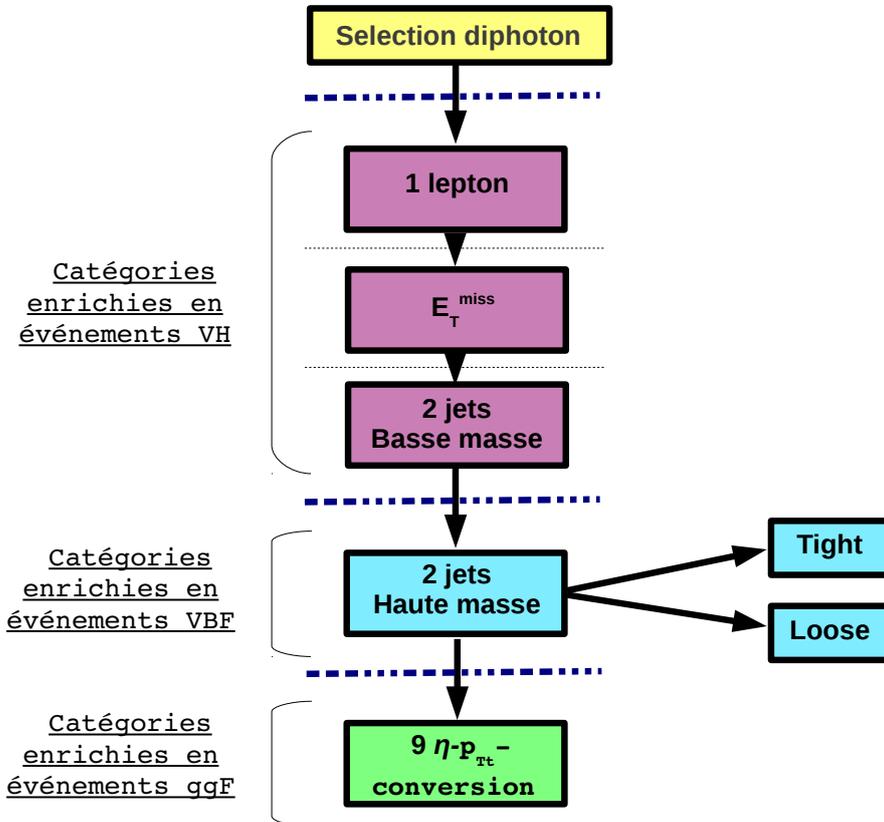
## 7.4 Autres catégories

Le tableau 7.5 rappelle le nombre d'événements de signal attendus pour  $20.7 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$ . Etant donné le faible nombre d'événements ttH attendus, aucune catégorie dédiée à ce mode de production n'a été définie. Toutefois les catégories VH avec un lepton et de l'énergie transverse manquante seront naturellement enrichies en événements ttH.

$\sqrt{s}$		8 TeV			
		$\mathcal{L} = 20.7 \text{ fb}^{-1}$			
$N_{Stot}$	$N_{ggF}$	$N_{VBF}$	$N_{WH}$	$N_{ZH}$	$N_{ttH}$
395	347	29	11	6	2

**Tableau 7.5** – Nombre d'événements de signal attendus pour  $20.7 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$ .

Afin d'éviter tout double-comptage d'événements lors de la catégorisation, un ordre est défini dans l'application des catégories, comme illustré figure 7.12. Les événements ne passant pas les coupures VH ou VBF sont répartis dans les neuf catégories définies en fonction du  $p_{Tt}$ , de  $\eta$  et du statut de conversion, détaillées section 6.2.1. Par ailleurs, des coupures sont appliquées sur la distance entre les objets reconstruits lors de la catégorisation. Un électron tel que  $\Delta R(e, \gamma) = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$ , un jet tel que  $\Delta R(jet, e) < 0.2$  ou  $\Delta R(jet, \gamma) < 0.4$  et un muon tel que  $\Delta R(\mu, jet) < 0.4$  ou  $\Delta R(\mu, \gamma) < 0.4$  ne sont pas considérés dans les catégories correspondantes.



**Figure 7.12** – Organigramme de la catégorisation.

## 7.5 Extraction du signal

### 7.5.1 Modélisation du bruit de fond et du signal

La modélisation du signal et du bruit de fond est obtenue de la même façon que pour l'analyse de juillet 2012.

Le signal est décrit par une fonction Crystal Ball additionnée à une fonction gaussienne. Les nombres d'événements observés dans les données et attendus pour la production d'un boson de Higgs de masse 126.5 GeV sont présentés tableau 7.6 avec le détail de la contribution de chaque mode de production.

Le tableau 7.7 présente la résolution sur la masse du signal ainsi que le rapport signal sur bruit dans une région de masse invariante autour de 126.5 GeV comprenant 90% des événements de signal attendus pour chacune des 14 catégories de l'analyse. La résolution sur la masse invariante varie de 1.40 GeV pour la catégorie "non-converti central, haut  $p_{Tt}$ " à 2.52 GeV pour la catégorie "converti transition".

Categorie	$\sqrt{s} = 8 \text{ TeV}$ $\mathcal{L} = 20.7 \text{ fb}^{-1}$						
	$N_D$	$N_S$	$gg \rightarrow H$ [%]	VBF [%]	$WH$ [%]	$ZH$ [%]	$ttH$ [%]
CP1 Non-conv. central, bas $p_{Tt}$	10900	51.8	93.7	4.0	1.4	0.8	0.2
CP2 Non-conv. central, haut $p_{Tt}$	553	7.9	79.3	12.6	4.1	2.5	1.4
CP3 Non-conv. non-central, bas $p_{Tt}$	41236	107.9	93.2	4.0	1.6	1.0	0.1
CP4 Non-conv. non-central, haut $p_{Tt}$	2558	16.0	78.1	13.3	4.7	2.8	1.1
CP5 Conv. central, bas $p_{Tt}$	7109	33.1	93.6	4.0	1.3	0.9	0.2
CP6 Conv. central, haut $p_{Tt}$	363	5.1	78.9	12.6	4.3	2.7	1.5
CP7 Conv. non-central, bas $p_{Tt}$	38156	97.8	93.2	4.1	1.6	1.0	0.1
CP8 Conv. non-central, haut $p_{Tt}$	2360	14.4	77.7	13.0	5.2	3.0	1.1
CP9 Conv. transition	14864	40.1	90.7	5.5	2.2	1.3	0.2
CP10 Loose haute-masse 2-jets	276	5.3	45.0	54.1	0.5	0.3	0.1
CP11 Tight haute-masse 2-jets	136	8.1	23.8	76.0	0.1	0.1	0.0
CP12 Basse-masse 2-jet	210	3.3	48.1	3.0	29.7	17.2	1.9
CP13 MET	49	1.3	4.1	0.5	35.7	47.6	12.1
CP14 Lepton	123	2.9	2.2	0.6	63.2	15.4	18.6
Inclusif	118893	395.0	88.0	7.3	2.7	1.5	0.5

**Tableau 7.6** – Nombre d'événements observés dans les données ( $N_D$ ) et nombre d'événements de signal attendus ( $N_S$ ) dans chaque catégorie pour  $m_H = 126.5 \text{ GeV}$  dans l'intervalle 100-160 GeV à  $\sqrt{s} = 8 \text{ TeV}$ . La fraction d'événements pour chaque processus est également détaillée. Le processus dominant par catégorie est mis en évidence en rouge.

Comme décrit section 6.2.2, la fonction utilisée pour décrire le bruit de fond dépend du nombre d'événements de la catégorie. Pour les catégories nouvellement définies, une fonction gaussienne est utilisée. Pour l'ensemble inclusif un polynôme de Bernstein de degré quatre est choisi. Les erreurs systématiques liées à la modélisation du bruit de fond sont estimées à partir de la simulation (voir section 6.2.2). Toutefois, en raison du faible nombre d'événements présents dans la catégorie  $E_T^{miss}$ , l'incertitude pour cette catégorie est estimée à partir du spectre  $m_{\gamma\gamma}$  inclusif en utilisant le rapport du nombre d'événements total de signal sur le nombre d'événements de signal dans la catégorie  $E_T^{miss}$ .

La distribution  $m_{\gamma\gamma}$  inclusive pour les données combinées à  $\sqrt{s} = 7 \text{ TeV}$  et à  $\sqrt{s} = 8 \text{ TeV}$  est

Categorie	$\sqrt{s} = 8 \text{ TeV}$ $\mathcal{L} = 20.7 \text{ fb}^{-1}$				
	$\sigma_{CB}$ (GeV)	Observé	$N_S$	$N_B$	$N_S/N_B$
CP1 Non-conv. central, bas $p_{Tt}$	1.50	911	46.6	881	0.05
CP2 Non-conv. central, haut $p_{Tt}$	1.40	49	7.1	44	0.16
CP3 Non-conv. non-central, bas $p_{Tt}$	1.74	4611	97.1	4347	0.02
CP4 Non-conv. non-central, haut $p_{Tt}$	1.69	292	14.4	247	0.06
CP5 Conv. central, bas $p_{Tt}$	1.68	722	29.8	687	0.04
CP6 Conv. central, haut $p_{Tt}$	1.54	39	4.6	31	0.15
CP7 Conv. non-central, bas $p_{Tt}$	2.01	4865	88.0	4657	0.02
CP8 Conv. non-central, haut $p_{Tt}$	1.87	276	12.9	266	0.05
CP9 Conv. transition	2.52	2554	36.1	2499	0.01
CP10 Loose haute-masse 2-jets	1.71	40	4.8	28	0.17
CP11 Tight haute-masse 2-jets	1.64	24	7.3	13	0.57
CP12 Basse-masse 2-jets	1.62	21	3.0	21	0.14
CP13 MET	1.74	8	1.1	4	0.24
CP14 Lepton	1.75	19	2.6	12	0.20
Inclusif	1.77	14025	355.5	13280	0.03

**Tableau 7.7** – Résolution sur la masse du signal ( $\sigma_{CB}$ ), nombre d'événements observés, nombre d'événements de signal attendus ( $N_S$ ) et rapport signal sur bruit ( $N_S/N_B$ ) dans une région de masse invariante autour de 126.5 GeV comprenant 90% des événements de signal attendus pour chacune des 14 catégories de l'analyse à  $\sqrt{s} = 8 \text{ TeV}$ . Le nombre d'événements de bruit de fond est obtenu à partir d'un fit (signal + bruit de fond) sur la distribution  $m_{\gamma\gamma}$  dans les données.

présentée figure 7.13.

## 7.5.2 Incertitudes systématiques

Les erreurs systématiques affectent le nombre attendu d'événements de signal, la résolution du signal ou entraînent une migration d'événements entre les catégories. Les différentes erreurs systématiques considérées sont ici présentées [152] et [133]. Toutes les erreurs systématiques, exceptée celle liée à la luminosité, sont considérées comme corrélées entre les données à  $\sqrt{s} = 7 \text{ TeV}$  et  $\sqrt{s} = 8 \text{ TeV}$ .

### 7.5.2.1 Incertitudes sur le nombre d'événements de signal attendus

Plusieurs sources d'erreurs systématiques, théoriques et expérimentales affectent le nombre d'événements de signal attendu.

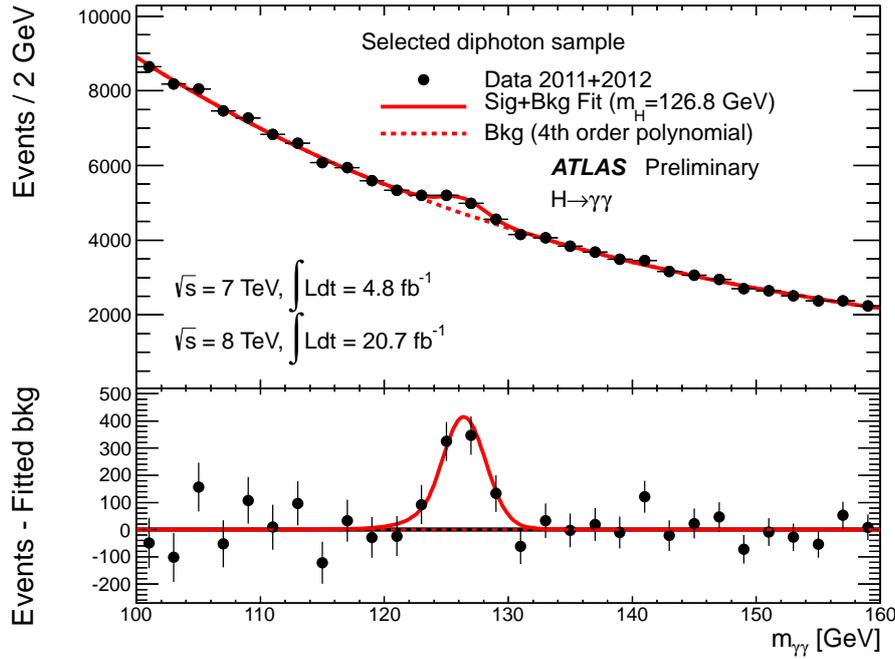
- L'incertitude sur la luminosité est de  $\pm 3.6\%$  pour les données à 8 TeV et de  $\pm 1.8\%$  pour les données à 7 TeV [87].
- L'incertitude sur l'efficacité du trigger est de  $\pm 0.5\%$ . L'efficacité de trigger est calculée à partir des données avec la méthode bootstrap [119] et l'incertitude est obtenue à partir du Monte-Carlo.
- L'incertitude sur l'efficacité d'identification des photons pour les données à 8 TeV a été largement réduite par rapport à l'analyse de juillet 2012, elle passe de 10.8% à 2.4%. Cette

Categorie	Parametrisation	Incertitude [ $N_{\text{evt}}$ ]	
		$\sqrt{s} = 7 \text{ TeV}$	$\sqrt{s} = 8 \text{ TeV}$
Inclusif	4th order pol.	7.3	12.0
CP1 Non-conv. central, low $p_{Tt}$	Pol. exp. du 2nd ordre	2.1	4.6
CP2 Non-conv. central, high $p_{Tt}$	Exponentielle	0.2	0.8
CP3 Non-conv. non-central, low $p_{Tt}$	Pol. 4eme ordre	2.2	11.4
CP4 Non-conv. non-central, high $p_{Tt}$	Exponentielle	0.5	2.0
CP5 Conv. central, low $p_{Tt}$	Pol. exp. du 2nd ordre	1.6	2.4
CP6 Conv. central, high $p_{Tt}$	Exponentielle	0.3	0.8
CP7 Conv. non-central, low $p_{Tt}$	Pol. 4eme ordre	4.6	8.0
CP8 Conv. non-central, high $p_{Tt}$	Exponentielle	0.5	1.1
CP9 Conv. transition	Pol. exp. du 2nd ordre	3.2	9.1
CP10 Loose haute-masse 2-jets	Exponentielle	0.4	1.1
CP11 Tight haute-masse 2-jets	Exponentielle	-	0.3
CP12 Basse-masse 2-jets	Exponentielle	-	0.6
CP13 MET	Exponentielle	-	0.1
CP14 Lepton	Exponentielle	-	0.3

**Tableau 7.8** – Erreur systématique sur le nombre fitté d'événements de signal en raison de la fonction choisie pour la paramétrisation du bruit de fond pour l'analyse des données à  $\sqrt{s} = 7 \text{ TeV}$  et à  $\sqrt{s} = 8 \text{ TeV}$ .

amélioration est principalement liée à la combinaison des différentes méthodes basées sur les données (décrites section 4.4.3).

- L'incertitude sur l'efficacité d'isolation, obtenue en comparant données et Monte-Carlo pour les événements  $Z \rightarrow ee$  est de 1%.
- L'incertitude sur l'échelle d'énergie des photons amène une incertitude de 0.25% sur le nombre d'événements de signal.
- L'incertitude sur le rapport d'embranchement en deux photons du boson de Higgs décroît de façon monotone de  $\pm 5.9\%$  pour  $m_H=110 \text{ GeV}$  à  $\pm 2.1\%$  pour  $m_H=150 \text{ GeV}$  [80].
- L'incertitude liée aux PDFs est estimée en utilisant les PDFs CTEQ [32], MSTW [75], et NNPDF [76]. Les incertitudes sur les échelles d'énergie pour les différents modes de production sont obtenues dans [81] et [23], elles sont détaillées tableau 7.9 pour un boson de Higgs de masse 126 GeV.  
Les erreurs d'échelle pour le processus de fusion de gluons dans les catégories 2-jets sont considérées séparément.
- Les incertitudes théoriques sur les prédictions de sections efficaces de production ggF+2 jets affectent principalement les catégories 2-jets à haute masse (les catégories VBF). Dans la simulation du signal obtenue à partir de POWHEG, l'un des jets est produit par les éléments de matrice calculés à NLO mais l'autre jet est issu de la parton-shower ; l'incertitude liée à cet ordre manquant est estimée à partir des prédictions MCFM [160] à 48% et 28% pour les catégories tight haute masse 2-jets et loose haute masse 2-jets



**Figure 7.13** – Distribution de la masse invariante pour les données combinées à  $\sqrt{s} = 7 \text{ TeV}$  ( $4.8 \text{ fb}^{-1}$ ) et à  $\sqrt{s} = 8 \text{ TeV}$  ( $20.7 \text{ fb}^{-1}$ ) [151].

respectivement. L'incertitude sur la catégorie basse masse 2-jets est de 30%.

Le résumé des incertitudes systématiques sur le nombre d'événements de signal est présenté tableau 7.9.

### 7.5.2.2 Incertitudes sur la résolution en masse du signal

- L'incertitude sur la résolution en énergie du calorimètre (section 2.2.3), combinée à l'incertitude liée à l'extrapolation de la calibration en énergie des électrons vers les photons (estimée à partir de la simulation) (section 4.2.2), évaluée séparément pour chaque catégorie varie entre 14 et 23% en fonction de la catégorie. Cette incertitude est supérieure à celle présentée pour l'analyse de juillet ; la capacité du Monte Carlo à décrire les données  $Z \rightarrow ee$  s'est dégradée avec l'ajout des données de la fin 2012.
- L'incertitude due à la modélisation de l'empilement est de  $\pm 1.5\%$ . Elle est obtenue à partir de la simulation en variant le nombre moyen d'interactions par croisement de paquets.
- L'incertitude associée à la sélection du vertex primaire est de 0.2% et est considérée comme négligeable.

### 7.5.2.3 Incertitudes liées aux migrations entre catégories

- L'estimation du nombre de photons convertis dans la simulation est très sensible à la quantité de matière décrite en amont du calorimètre. L'incertitude liée à cet effet est calculée à partir de la simulation nominale et d'une simulation pour laquelle 5 à 20% de matériel ont été rajoutés en fonction de  $\eta$ . L'impact est de -4% sur le nombre d'événements des catégories non-converties et de +3.5% pour les catégories converties.

- L’incertitude pour les catégories définies en fonction du  $p_{Tt}$ , liée à la modélisation du  $p_{Tt}$  est estimée en variant les échelles d’énergies et les PDFs utilisées [151]. L’impact est de 1.3% pour les catégories à bas  $p_{Tt}$ , de 10.2% pour les catégories à haut  $p_{Tt}$ , de 10.4% (8.5%) pour la catégorie à haute masse 2-jet tight (loose), de 12.5% pour la catégorie basse masse 2-jets, de 2.0% pour la catégorie MET et de 4.0% pour la catégorie lepton.
- L’incertitude liée à l’échelle d’énergie des jets (JES) est obtenue en variant les corrections de  $\pm 1\sigma$ . Ceci induit un effet allant jusqu’à 11.8% (10.7%) pour la catégorie à haute masse 2-jet tight (loose), un effet allant jusqu’à 6.7% pour la catégorie basse masse 2-jets et un effet allant jusqu’à 0.7% pour les autres catégories.  
L’incertitude sur la résolution en énergie des jets (JER) a un impact allant jusqu’à 3.8% (3.4%) pour la catégorie à haute masse 2-jet tight (loose), un impact allant jusqu’à 3.4% sur la catégorie basse masse 2-jets et un impact allant jusqu’à 0.9% pour les autres catégories.
- L’incertitude sur la modélisation des événements sous-jacent est estimée en comparant les simulation avec et sans les interactions multiples entre partons (MPI). La configuration utilisée par défaut est AU2-CT10 [161], tandis que les échantillons sans MPI sont utilisés pour le calcul des erreurs systématiques. Pour la catégorie tight haute masse 2-jets, l’incertitude est de 8.8% pour les processus ggF, VH et ttH et de 2.0% pour le processus VBF. Pour la catégorie loose haute masse 2-jets, ces incertitudes sont de 12.8% et 3.3% respectivement. Enfin pour la catégorie basse masse 2-jets, l’erreur est de 12% pour ggF, VH, ttH et de 3.9% pour le VBF.
- Une incertitude est liée à la modélisation de la variable  $\Delta\phi_{\gamma\gamma-jj}$  pour le processus ggF. La distribution POWHEG est comparée à SHERPA (où les éléments de matrices sont calculés pour jusqu’à trois partons dans l’état final) et pondérée afin de faire coïncider les deux distributions. L’incertitude résultante est de, respectivement, 12.1% et 8.5% pour les catégories tight et loose haute masse 2-jets.
- Une incertitude est liée à la modélisation de la variable  $\Delta\eta^*$  pour le processus ggF. La différence observée entre POWHEG et MCFM est estimée en pondérant la distribution des jets au niveau particule pour la faire coïncider à la distribution MCFM. Cette incertitude est de 7.6% et 6.2% pour les catégories tight et loose haute masse 2-jets respectivement.
- Une incertitude est assignée à la coupure sur la JVF, elle est de 0.3% (1.2%) dans la catégorie loose haute masse 2-jets pour ggF (VBF) et de 2.3% (2.4%) dans la catégorie basse masse 2-jets pour ggF (VBF).
- Des incertitudes inférieures à 1% chacune sont liées à la reconstruction et l’identification des électrons, à l’échelle d’énergie et la résolution des électrons, enfin à la reconstruction et identification des muons.
- L’incertitude liée à la définition de l’énergie transverse manquante est estimée en décalant l’énergie transverse de chaque objet par leur incertitude sur la résolution et l’échelle. L’incertitude résultante est de 66.4% pour le ggF, de 30.7% pour le VBF et de 1.2% pour VH dans la catégorie  $E_T^{miss}$ .

L’ensemble des incertitudes systématiques liées aux migrations d’événements de signal entre catégories sont résumées tableau 7.10.

### 7.5.2.4 Incertitudes sur la mesure de la masse

Trois sources d'erreurs systématiques principales interviennent dans la mesure de la masse du boson de Higgs. Il s'agit de l'extrapolation de la calibration en énergie des électrons vers les photons (0.3%), de la modélisation de la matière en amont du calorimètre (0.3%) et de la calibration de l'échelle d'énergie du pré-échantillonneur (0.1%). D'autres sources d'erreurs ont été investiguées et sont détaillées dans [152]. L'incertitude totale sur la mesure de la masse est de 0.55%, ce qui correspond à une erreur de 0.7 GeV sur la position du pic.

Source	Valeur(%)			Contrainte
Luminosité ( $\theta_{lumi}$ )	$\pm 3.6$			
Trigger	$\pm 0.5$			
Photon identification ( $\theta_{eff}$ )	$\pm 2.4$			Log-normal
Isolation ( $\theta_{iso}$ )	$\pm 1.0$			
Echelle d'énergie des photons ( $\theta_{es}$ )	$\pm 0.25$			
BR ( $\theta_{BR}$ )	$\pm 5.9\% - \pm 2.1\%$ ( $m_H = 110 - 150$ GeV)			Asymétrique Log-normal
Erreur d'échelle ( $\theta_{TH}$ )	ggF : $^{+7.2}_{-7.8}$	VBF : $^{+0.2}_{-0.2}$	WH : $^{+0.2}_{-0.6}$	Asymétrique
	ZH : $^{+1.6}_{-1.5}$	ttH : $^{+3.8}_{-9.3}$		Log-normal
PDF+ $\alpha_s$ ( $\theta_{pdf}$ )	ggF : $^{+7.5}_{-6.9}$	VBF : $^{+2.6}_{-2.7}$	WH : $\pm 3.5$	Asymétrique
	ZH : $\pm 3.6$	ttH : $\pm 7.8$		Log-normal
Section efficace th. pour ggF ( $\theta_{TH,2-jets}$ )	Tight haute-masse 2-jets : $\pm 48$			Log-normal
	Loose haute-masse 2-jets : $\pm 28$			
	Basse-masse 2-jets : $\pm 30$			

**Tableau 7.9** – Résumé des erreurs systématiques sur le nombre d'événements de signal attendus pour l'analyse des données à  $\sqrt{s} = 8$  TeV.

Source	Categorie	Valeur(%)			Contrainte
Evts. sous-jacent ( $\theta_{UE}$ )	Tight haute-masse 2-jets	ggF : $\pm 8.8$	VBF : $\pm 2.0$	VH, ttH : $\pm 8.8$	Log-normal
	Loose haute-masse 2-jets	ggF : $\pm 12.8$	VBF : $\pm 3.3$	VH, ttH : $\pm 12.8$	
	Basse-masse 2-jets	ggF : $\pm 12$	VBF : $\pm 3.9$	VH, ttH : $\pm 12$	
Echelle d'énergie des jets (JES) ( $\theta_{JES}$ )	Bas $p_{Tt}$	ggF : $-0.1$	VBF : $-1.0$	Autres : $-0.1$	Gauss.
	Haut $p_{Tt}$	ggF : $-0.7$	VBF : $-1.3$	Autres : $+0.4$	
	Tight haute-masse 2-jets	ggF : $+11.8$	VBF : $+6.7$	Autres : $+20.2$	
	Loose haute-masse 2-jets	ggF : $+10.7$	VBF : $+4.0$	Autres : $+5.7$	
	Basse-masse 2-jets	ggF : $+6.7$	VBF : $+2.6$	Autres : $1.4$	
	MET	ggF : $0.0$	VBF : $0.0$	Autres : $0.0$	
Jet énergie résolution (JER) ( $\theta_{JER}$ )	Bas $p_{Tt}$	ggF : $0.0$	VBF : $0.2$	Autres : $0.0$	Gauss.
	Haut $p_{Tt}$	ggF : $-0.2$	VBF : $0.2$	Autres : $0.6$	
	Tight haute-masse 2-jets	ggF : $3.8$	VBF : $-1.3$	Autres : $7.0$	
	Loose haute-masse 2-jets	ggF : $3.4$	VBF : $-0.7$	Autres : $1.2$	
	Basse-masse 2-jets	ggF : $0.5$	VBF : $3.4$	Autres : $-1.3$	
	MET	ggF : $0.0$	VBF : $0.0$	Autres : $0.0$	
	Lepton	ggF : $-0.9$	VBF : $-0.5$	Autres : $-0.1$	
$\eta^*$ modélisation ( $\theta_{\eta^*}$ )	Tight haute-masse 2-jets	Tight haute-masse 2-jets : $+7.6$			Gauss.
	Loose haute-masse 2-jets	Loose haute-masse 2-jets : $+6.2$			
$\Delta\phi_{\gamma\gamma jj}$ modélisation ( $\theta_{d\phi}$ )	Tight high-mass two-jet	Tight high-mass two-jet : $+12.1$			Gauss.
	Loose high-mass two-jet	Loose high-mass two-jet : $+8.5$			
Higgs $p_{Tt}$ ( $\theta_{p_{Tmodel}}$ )	Low $p_{Tt}$	Low $p_{Tt}$ : $+1.3$			Gauss.
	High $p_{Tt}$	High $p_{Tt}$ : $-10.2$			
	Tight haute-masse 2-jets	Tight haute-masse 2-jets : $-10.4$			
	Loose haute-masse 2-jets	Loose haute-masse 2-jets : $-8.5$			
	Basse-masse 2-jets	Basse-masse 2-jets : $-12.5$			
	MET	MET : $-2.0$			
Matériel model. ( $\theta_{mat}$ )	Unconv	Unconv : $-4.0$	Conv : $+3.5$		Gauss.
	Conv				
JVF ( $\theta_{JVF}$ )	Loose haute-masse 2-jets	ggF : $-1.2$	VBF : $-0.3$	Autres : $-1.2$	Gauss.
	Basse-masse 2-jets	ggF : $-2.3$	VBF : $-2.4$	Autres : $-2.3$	
MET ( $\theta_{MET}$ )	MET	ggF : $+66.4$	VBF : $+30.7$	VH, ttH : $+1.2$	Gauss.
$e$ reco et ID ( $e_{ID}$ )		Lepton : $< 1$			Gauss.
$e$ ech. énergie et résolution ( $e_{res}$ )		Lepton : $< 1$			Gauss.
$\mu$ reco, ID, résolution ( $\mu_{ID}$ )		Lepton : $< 1$			Gauss.
$\mu$ spectromètre resolution		Lepton : $0$			Gauss.

**Tableau 7.10** – Résumé des erreurs systématiques liées à la migration d'événements entre les catégories pour l'analyse des données à  $\sqrt{s} = 8$  TeV.

### 7.5.3 Procédure statistique

Les outils statistiques utilisés pour établir la découverte ont été introduits au chapitre précédent. On se concentre ici plus précisément sur le traitement des paramètres de nuisance dans la procédure statistique.

L'analyse statistique est basée sur une fonction de vraisemblance non-binnée, dont la forme générale est décrite dans l'équation 6.3. Les paramètres de nuisance sont les paramètres liés aux incertitudes systématiques décrites précédemment mais aussi ceux liés à la forme du signal et à la forme du bruit de fond.

Les différentes catégories sont traitées indépendamment ; un fit simultané est effectué pour extraire les résultats. La fonction de vraisemblance s'écrit :

$$\mathcal{L}(\mu, \theta) = e^{-(\mu N_{s,c}^{pred}(\theta_{s,c}^{num}) + \sum_{c=1}^{n_c} N_{b,c})} \prod_{c=1}^{n_c} \prod_{k=1}^n \mu N_{s,c}^{pred}(\theta_{s,c}^{num}) f_{s,c}(x_k; m_h, \theta_{s,c}) + N_{b,c} f_{b,c}(x_k; \theta_{b,c}) \quad (7.3)$$

- $n$  : nombre total d'événements .
- $n_c$  : nombre total de catégories.
- $f_{s,c}$  et  $f_{b,c}$  : fonctions de densité de probabilité pour le signal et le bruit de fond pour la variable  $x$  (typiquement  $m_{\gamma\gamma}$ ) dans la catégorie  $c$ .
- $N_{b,c}$  : le nombre d'événements de bruit de fond dans la catégorie  $c$ .
- $\mu N_{s,c}^{pred}$  : correspond au nombre d'événements de signal dans la catégorie  $c$ ,  $N_s^{pred}$  est le nombre d'événements prédits par le Modèle Standard dans la catégorie  $c$  et  $\mu$  est la force du signal.
- $\theta_{s,c}$  et  $\theta_{b,c}$  : les paramètres restant dans les fonctions de densité de probabilité du signal et du bruit de fond (paramètres libres dans le fit) et les paramètres utilisés pour prendre en compte les effets systématiques dans la catégorie  $c$ .
- $\theta_{s,c}^{num}$  : le paramètre de nuisance associé au nombre d'événements de signal dans la catégorie  $c$ .

$\mu N_{s,c}^{pred}(\theta_{s,c}^{num})$  peut s'écrire :

$$\begin{aligned} \mu N_{s,c}^{pred}(\theta_{s,c}^{num}) &= \mu N_{s,c}^{pred} \times e^{\sigma_{eff}\theta_{eff}} \times e^{\sigma_{es}\theta_{es}} \times e^{\sigma_{lumi}\theta_{lumi}} \\ &\times e^{\sigma_{BR}\theta_{BR}} \times e^{\sigma_{iso}\theta_{iso}} \times (1 + \sigma_{mat}\theta_{mat}) \end{aligned} \quad (7.4)$$

avec :

$$\mu N_{s,c}^{pred} = \mu [N_c^{ggF}(\theta_c^{ggF}) + N_c^{VBF}(\theta_c^{VBF}) + N_c^{VH}(\theta_c^{VH}) + N_c^{ZH}(\theta_c^{ZH}) + N_c^{ttH}(\theta_c^{ttH})] \quad (7.5)$$

Les  $N_c^x$  sont les nombres d'événements de signal attendus par mode de production dans la catégorie  $c$ , les  $\sigma$  correspondent aux erreurs systématiques détaillées tableaux 7.9 et 7.10 avec les  $\theta$  les paramètres de nuisance associés. Les erreurs systématiques suivent une distribution gaussienne,  $N \rightarrow N(1 + \sigma\theta)$ , ou une distribution log normal dont la variance est approximée par  $\sigma$ ,  $N \rightarrow Ne^{\sigma\theta}$ . La distribution gaussienne est utilisée pour les migrations d'événements (tableau 7.10) tandis que la distribution log normal est utilisée pour les systématiques sur le nombre d'événements de signal (tableau 7.9) afin d'éviter les contributions négatives.

L'erreur systématique sur le trigger est ajoutée quadratiquement à celle sur l'efficacité d'identification des photons.

Les nombres d'événements par mode de production et les paramètres de nuisance associés sont détaillés par catégorie dans l'annexe F.

La forme du signal  $f_{s,c}$  s'écrit :

$$f_{s,c} = r_{CB} CB(m_{\gamma\gamma}, CB_{peak}(m_H, \theta_{masse}), CB_{\sigma}(\theta_{res}), CB_{\alpha}, CB_N) \\ (1 - r_{CB}) G(m_{\gamma\gamma}, CB_{peak}(m_H, \theta_{masse}), G_{\sigma}(\theta_{res})) \quad (7.6)$$

$CB_N$  est fixé à 10,  $r_{CB}$  correspond au rapport de l'aire de la Crystal-Ball sur l'aire totale. Les paramètres de nuisance  $\theta_{masse}$  et  $\theta_{res}$  sont respectivement associés aux systématiques sur la mesure de la masse et sur la résolution du signal décrites section 7.5.2.

La forme du bruit de fond est paramétrée par :

$$f_{b,c} = f(c1, c2, m_{\gamma\gamma}) , \text{ polynôme exp. } 2^{nd} \text{ ordre pour les catégories CP1, CP5 et CP9} \\ f_{b,c} = f(cs1, cs2, cs3, cs4, m_{\gamma\gamma}) , \text{ polynôme Bernstein } 4^{ème} \text{ ordre pour les catégories CP7 et CP3} \\ f_{b,c} = f(\xi, m_{\gamma\gamma}) , \text{ gaussienne pour les autres catégories} \quad (7.7)$$

Les deux paramètres libres du fit sont  $\mu$  et  $m_H$ . Les paramètres liées à la forme du bruit de fond ( $\xi, c_i, c_{si}$ ) ne sont pas contraints et tous les autres paramètres de nuisance sont contraints par une distribution gaussienne ou gaussienne bifurquée si l'erreur est asymétrique.

## 7.6 Résultats

### 7.6.1 $p_0$ et force du signal

#### 2012

La valeur de  $p_0$  observée dans les données 2012 en fonction de  $m_H$  est présentée figure 7.14. La plus grande signification observée est de  $6.5\sigma$  correspondant à  $p_0=3.15\times 10^{-11}$  pour  $m_H$  proche de 126.5 GeV. La signification attendue est de  $3.8\sigma$ .

La meilleure valeur de  $\hat{\mu}$  obtenue par le fit est  $\hat{\mu} = 1.587 \pm 0.274$  (stat)  $\pm 0.205$  (syst) pour  $m_H=126.90 \pm 0.788$  GeV.

#### 2011+2012

Pour la combinaison des données 2011 et 2012, la plus grande signification est de  $7.4\sigma$  (quand la signification attendue est de  $4.1\sigma$ ) pour  $m_H$  autour de 126.5 GeV (figure 7.15).

La meilleure valeur de  $\hat{\mu}$  obtenue par le fit est  $\hat{\mu} = 1.640 \pm 0.259$  (stat)  $\pm 0.198$  (syst) pour  $m_H=126.817 \pm 0.747$  GeV. L'incertitude systématique est également répartie entre incertitude théorique et incertitude expérimentale. L'incertitude liée aux migrations contribue à l'incertitude expérimentale à hauteur de 30%, l'incertitude sur le nombre d'événements de signal à hauteur de 30% et l'incertitude sur la résolution en masse du signal à hauteur de 40%. L'incertitude sur la résolution de la masse utilisé pour extraire ce résultat est  $\sim 20\%$  (comme décrit 7.5.2.2), d'autres études ont été menées sur l'influence de la résolution en masse sur  $\mu$ ; une variation de 40% sur la résolution amène une variation de 10% sur  $\mu$  [162].

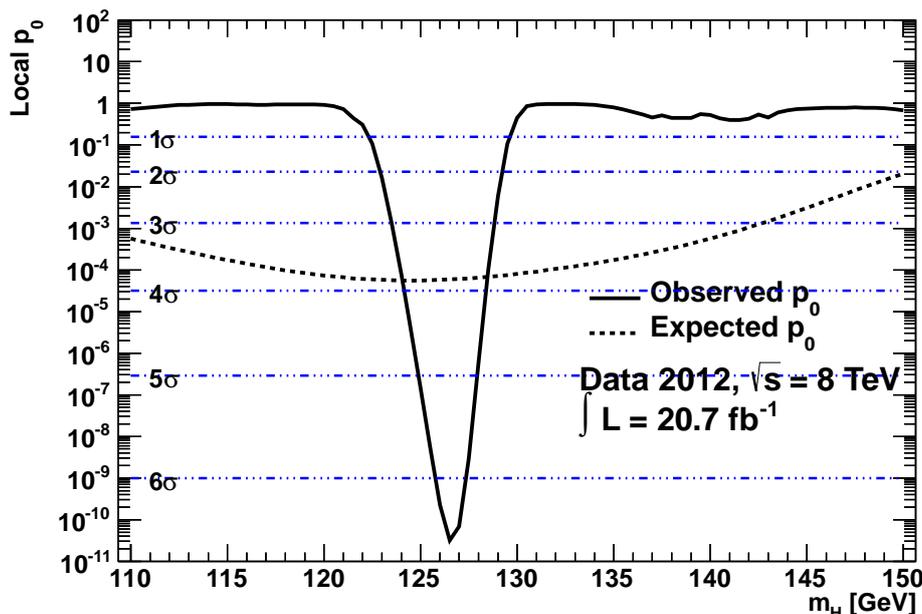


Figure 7.14 –  $p_0$  observée et attendue en fonction de la masse hypothétique du boson de Higgs pour  $20.7 \text{ fb}^{-1}$  de données à 8 TeV.

## 7.6.2 Force du signal par mode de production

Les  $\mu_i$  nous permettent d'estimer la contribution des différents modes de production et ainsi de quantifier d'éventuelles déviations entre les sections efficaces prédites par le Modèle Standard et celles mesurées dans les données. Comme mentionné auparavant, le but des nouvelles catégories VH introduites ainsi que l'optimisation de la catégorie VBF était d'augmenter la précision sur la mesure des  $\mu_i$ . En raison du faible nombre d'événements ttH attendus pour la luminosité dont on dispose, et parce que les modes de production ggF et ttH mettent tous deux en jeu le couplage du boson de Higgs au quark top, une seule force du signal est définie pour ces deux processus  $\mu_{ggF+ttH}$ . La fonction de vraisemblance utilisée est  $\mathcal{L}(\mu_{ggF+ttH}, \mu_{VBF}, \mu_{VH}, \hat{\theta})$ .

Les différentes forces du signal par mode de production sont détaillées pour les données 2012 seulement et pour la combinaison des données 2011 et 2012 (figure 7.16). Les éventuelles modifications de couplage influençant la désintégration du boson de Higgs en deux photons ne sont pas ici prises en compte.

### 2012

$$- \mu_{ggF+ttH} = 1.540_{-0.370}^{+0.389}(\text{stat}) \quad {}_{-0.136}^{+0.255}(\text{syst})$$

$$- \mu_{VBF} = 1.631_{-0.792}^{+0.883}(\text{stat}) \quad {}_{-0.459}^{+0.506}(\text{syst})$$

$$- \mu_{VH} = 1.850_{-1.348}^{+1.505}(\text{stat}) \quad {}_{-0.245}^{+0.419}(\text{syst})$$

### 2012+2011

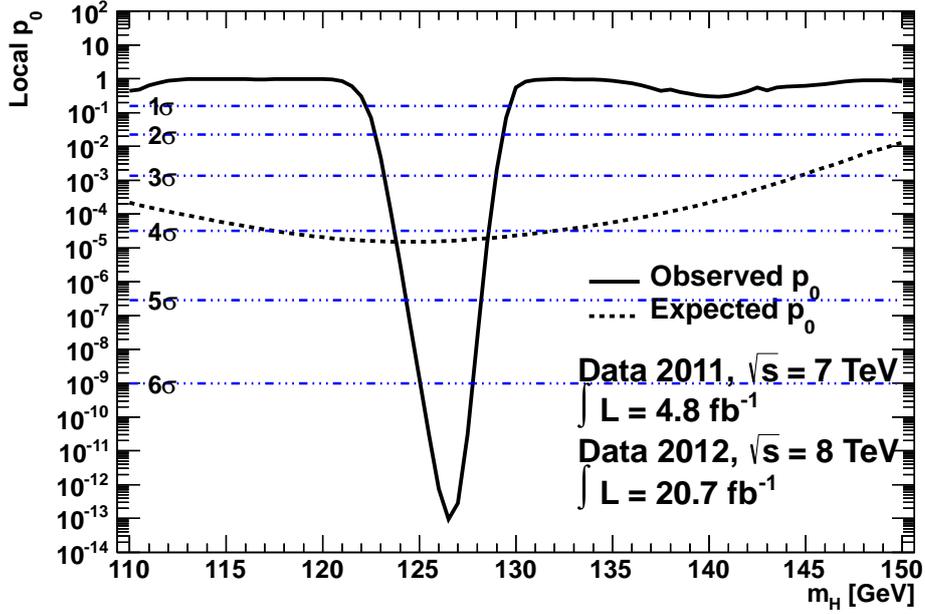


Figure 7.15 –  $p_0$  observée et attendue en fonction de la masse hypothétique du boson de Higgs pour  $4.8 \text{ fb}^{-1}$  de données à 7 TeV et  $20.7 \text{ fb}^{-1}$  de données à 8 TeV.

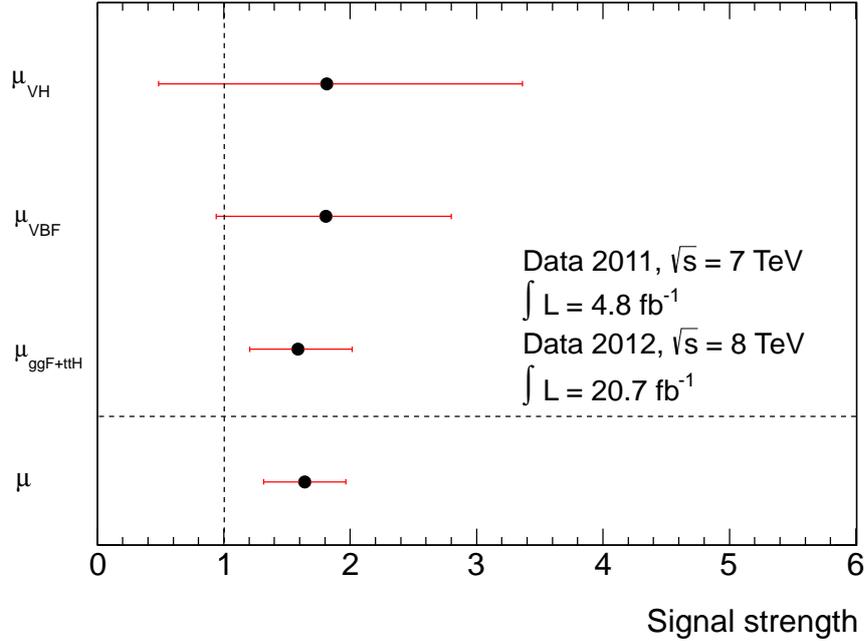
$$\begin{aligned}
 - \mu_{ggF+ttH} &= 1.585^{+0.364}_{-0.340}(\text{stat}) \quad ^{+0.232}_{-0.168}(\text{syst}) \\
 - \mu_{VBF} &= 1.808^{+0.824}_{-0.755}(\text{stat}) \quad ^{+0.549}_{-0.438}(\text{syst}) \\
 - \mu_{VH} &= 1.813^{+1.471}_{-1.315}(\text{stat}) \quad ^{+0.502}_{-0.220}(\text{syst})
 \end{aligned}$$

### 7.6.3 $p_0$ pour le mode de production VBF

Pour mieux apprécier notre sensibilité au mode de production VBF, en plus du  $\mu_{VBF}$ , on calcule la valeur-p pour le mode de production VBF uniquement. Le  $p_0^{VBF}$  est obtenu à partir du même modèle statistique, aucun autre paramètre ni aucune autre systématique n'est introduit. Le calcul est fait de la même façon que pour la valeur-p standard, mais cette fois, c'est uniquement les événements VBF qui sont considérés comme signal, les événements ggF, VH et ttH sont traités comme des bruit de fond additionnels.  $\mu_{VBF}$  est ainsi utilisé comme force du signal et  $\mu_{ggF+ttH}$  et  $\mu_{VH}$  deviennent des paramètres de nuisance.

$$\lambda_{0,VBF}(\mu_{VBF}) = \frac{\mathcal{L}(\mu_{VBF}, \hat{\mu}_t, \hat{\mu}_{VH}, \hat{\theta})}{\mathcal{L}(\hat{\mu}_{VBF}, \hat{\mu}_t, \hat{\mu}_{VH}, \hat{\theta})} \quad (7.8)$$

Le  $p_0$  observé pour le mode de production VBF en fonction de la masse  $m_H$  ainsi que le  $p_0$  attendu pour le mode de production VBF correspondant à l'hypothèse d'un boson de Higgs Modèle Standard plus bruit de fond est présenté figure 7.17 pour l'ensemble des données 2011 et 2012. La plus grande signification  $2.9\sigma$  est obtenue pour  $m_H = 123.5 \text{ GeV}$  quand la signification attendue est  $1.2\sigma$ . Pour  $m_H = 126.8 \text{ GeV}$  la signification observée et attendue sont respectivement de  $1.9\sigma$  et  $1.2\sigma$ .



**Figure 7.16** – Force du signal mesurée pour les différents mode de production,  $\mu_{ggF+ttH}$ ,  $\mu_{VBF}$ ,  $\mu_{VH}$ , et force du signal global  $\mu$  pour  $4.8 \text{ fb}^{-1}$  de données à 7 TeV et  $20.7 \text{ fb}^{-1}$  de données à 8 TeV.

## 7.7 Extraction des couplages

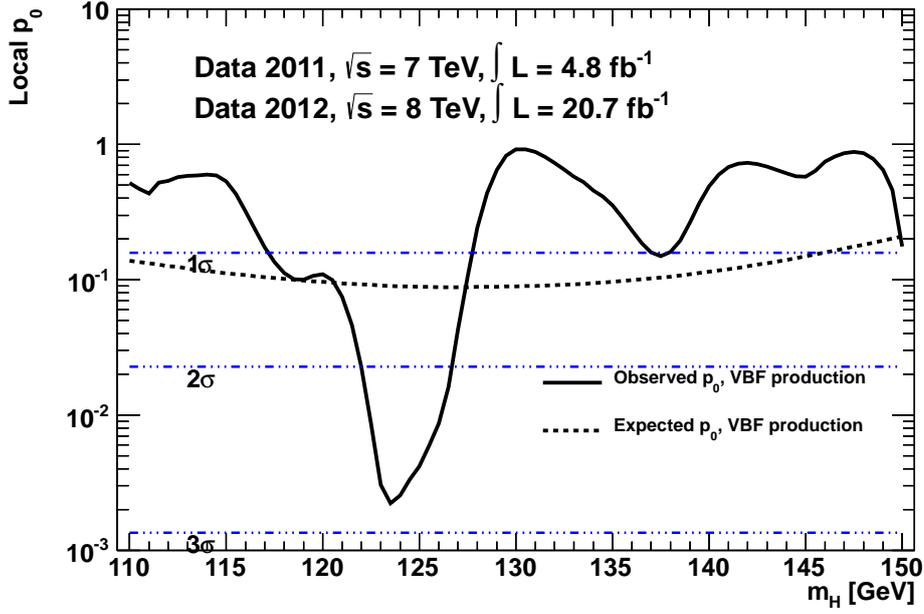
Les propriétés de ce nouveau boson sont également étudiées dans d'autres canaux de désintégration,  $H \rightarrow ZZ^{(*)} \rightarrow 4l$  et  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ . La combinaison de tous les canaux permet d'effectuer des mesures de couplage [163].

Pour chaque mode de production on a introduit une force de signal  $\mu_i$ , de façon équivalente on peut définir un facteur pour chaque mode de désintégration  $\mu_f = B_f/B_{f,SM}$  où  $B_{f,SM}$  est le rapport d'embranchement prédit par le Modèle Standard pour le mode de désintégration  $f$ .

Dans une première approche, on peut paramétrer le nombre d'événements de signal observés par le produit  $\mu_i \mu_f$ . C'est une approche générale dans le sens où elle n'est restreinte par aucune relation entre  $\mu_i$  et  $\mu_f$ .

Il n'est pas possible d'utiliser la combinaison des canaux pour accéder directement au  $\mu_i$  en raison des différents rapports d'embranchement. Il est toutefois possible déliminer cette dépendance en utilisant le rapport des  $\mu_i$  canal par canal,  $\mu_{VBF}/\mu_{ggF+ttH}$  devient le paramètre d'intérêt. Pour cette combinaison, il est simplement nécessaire de supposer que tous les événements de signal observés sont issus du même boson. Avec l'ensemble des données à 7 et 8 TeV [164] et en utilisant les canaux  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow ZZ^{(*)} \rightarrow 4l$  et  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ , la valeur obtenue pour  $\mu_{VBF}/\mu_{ggF+ttH}$  est  $1.4_{-0.3}^{+0.4}$  (stat)  $_{-0.4}^{+0.6}$  (syst). Ce résultat est compatible avec les prédictions du Modèle Standard et constitue la première évidence d'un boson de Higgs produit par fusion de bosons vecteurs.

Dans une seconde approche, afin d'extraire les facteurs de couplage de ce nouveau boson, les modes de production et de désintégration ne sont plus traités indépendamment. Si l'on considère une seule résonance à une masse proche de 125.5 GeV et si on néglige la largeur du boson de Higgs (prédite égale à  $\sim 4$  MeV pour  $m_H \sim 125$  GeV), le nombre d'événements de



**Figure 7.17** –  $p_0$  observée et attendue pour le mode de production VBF en fonction de la masse hypothétique du boson de Higgs pour  $4.8 \text{ fb}^{-1}$  de données à 7 TeV et  $20.7 \text{ fb}^{-1}$  de données à 8 TeV.

signal observés  $\sigma \times BR(ii \rightarrow H \rightarrow ff)$  peut s'écrire :

$$\sigma \times BR(ii \rightarrow H \rightarrow ff) = \sigma_{ii} \frac{\Gamma_{ff}}{\Gamma_H} \quad (7.9)$$

Où  $\Gamma_H$  est la largeur totale de désintégrations du boson de Higgs et  $\Gamma_{ff}$  la largeur partielle de désintégration dans l'état final  $ff$ . On suppose par ailleurs que le nouveau boson observé est un boson de spin  $0^+$  conformément aux prédictions du Modèle Standard. Le facteur de couplage  $\kappa_j$  est défini de façon à ce que la section efficace  $\sigma_j$  et la largeur partielle de désintégration  $\Gamma_j$  associées à la particule  $j$  soit pondérées par  $\kappa_j^2$  lors de la comparaison avec les prédictions du Modèle Standard, soit  $\kappa_j^2 = \Gamma_j/\Gamma_{j,SM}$ . On peut alors écrire :

$$\sigma \times BR(ii \rightarrow H \rightarrow ff) = \sigma_{SM}(ii \rightarrow H) \times BR_{SM}(H \rightarrow ff) \times \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2} \quad (7.10)$$

Pour les processus  $gg \rightarrow H$  et  $H \rightarrow \gamma\gamma$  induits par une boucle, les facteurs  $\kappa_g$  et  $\kappa_\gamma$  sont paramétrés en fonction des facteurs fondamentaux  $\kappa_t, \kappa_b, \kappa_W$ .

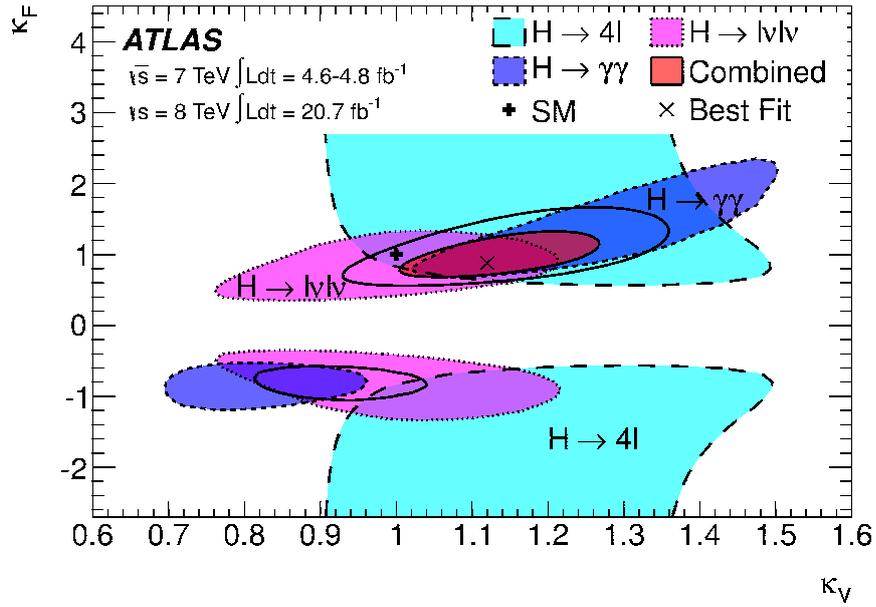
On ne considère que les contributions des particules du Modèle Standard dans la largeur totale et seuls le boson  $W$  et le quark top peuvent contribuer à la boucle de désintégration du boson de Higgs en deux photons. Les constantes de couplage sont identiques pour tous les fermions  $\kappa_F = \kappa_t = \kappa_b = \kappa_\tau = \kappa_g$  et pour les bosons vecteurs  $\kappa_V = \kappa_W = \kappa_Z$ . On peut alors écrire  $\kappa_\gamma$  en fonction de  $\kappa_V$  et  $\kappa_F$  :

$$\kappa_\gamma^2(\kappa_V, \kappa_Z) = 1.59\kappa_V^2 - 0.66\kappa_V\kappa_F + 0.07\kappa_F^2 \quad (7.11)$$

La figure 7.18 montre les résultats obtenus pour  $\kappa_F$  et  $\kappa_V$  pour chacun des trois canaux  $H \rightarrow \gamma\gamma$ ,  $H \rightarrow ZZ^{(*)} \rightarrow 4l$  et  $H \rightarrow WW^{(*)} \rightarrow l\nu l\nu$ , ainsi que pour leur combinaison. A 68% de confiance,  $\kappa_V$  et  $\kappa_F$  appartiennent aux intervalles tels que :

$$\kappa_V \in [1.05, 1.22] \quad (7.12)$$

$$\kappa_F \in [0.74, 1.06] \quad (7.13)$$



**Figure 7.18** – Contours de la fonction de vraisemblance des facteurs de couplages  $\kappa_F$  et  $\kappa_V$  obtenus à partir de fits sur les données pour les trois canaux individuellement et pour leur combinaison. La meilleure valeur obtenue ( $\times$ ) et la valeur attendue par le modèle standard ( $+$ ) sont également indiquées.

Plusieurs autres tests, pour différentes hypothèses sur la largeur totale et sur la boucle de désintégration sont effectués, aucune variation significative par rapport aux prédictions du Modèle Standard n'est observée.



# Conclusion

Depuis 2010, les performances du LHC sont remarquables.  $4.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$  et  $20.7 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$  ont pu être enregistrées par ATLAS et la qualité de ces données, indispensable pour des analyses de physique rigoureuses, est assurée par une procédure d'étalonnage mise en place dans les 48 heures qui suivent la fin de la prise de données.

Avec l'ensemble des données à 7 TeV nous avons pu mesurer la section efficace différentielle de production de paires de photons en fonction de quatre variables cinématiques, la masse invariante  $m_{\gamma\gamma}$  et l'impulsion transverse  $p_{T,\gamma\gamma}$  du système diphoton, l'angle azimuthal  $\Delta\phi_{\gamma\gamma}$  et le cosinus de l'angle polaire  $\cos\theta^*$  entre les deux photons. La précision sur notre mesure, qui n'est plus limitée par la statistique, permet une comparaison significative avec les prédictions théoriques. On constate que la contribution de la composante NNLO est nécessaire pour décrire correctement les données et particulièrement à basse masse invariante. Par ailleurs, dans les régions sensibles à l'émission des gluons mous, qui posent des problèmes de divergences dans les calculs à ordre fixe, le générateur SHERPA apporte la meilleure description des données.

Cette mesure permet de tester les prédictions de QCD perturbative mais elle apporte également une compréhension approfondie du bruit de fond dans le canal  $H \rightarrow \gamma\gamma$ . La méthode de décomposition du bruit fond (réductible et irréductible), développée pour la mesure de section efficace, permet de vérifier la pureté en événements  $\gamma\gamma$  dans les données sélectionnées pour la recherche du boson de Higgs dans le canal diphoton. Elle permet également de construire une simulation Monte-Carlo valide du bruit de fond pour l'estimation de l'incertitude sur le nombre d'événements de signal attendus.

La recherche du boson de Higgs dans le canal diphoton avec  $4.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$  et  $5.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 8 \text{ TeV}$  s'est avérée fructueuse, elle a mené à l'observation d'une nouvelle résonance de masse 126.5 GeV avec une signification locale de  $4.5\sigma$ . Cette observation a été confirmée dans d'autres canaux de recherche dans ATLAS ( $H \rightarrow ZZ^* \rightarrow 4l$  et  $H \rightarrow WW \rightarrow l\nu l\nu$ ) ainsi que dans l'expérience CMS. Le 4 juillet 2012, le CERN annonce la découverte d'une particule compatible avec le boson de Higgs prédit par le Modèle Standard.

La découverte de cette nouvelle particule dans le canal diphoton n'est toutefois pas suffisante pour établir s'il s'agit de la particule prédite par le Modèle Standard. Il est notamment nécessaire de mesurer les taux de production relatifs aux différents modes de production, afin de juger du caractère standard de cette particule. Pour chaque mode de production on définit la force du signal  $\mu_i$  correspondant au rapport du nombre d'événements observés sur le nombre d'événements attendus. Avec l'ensemble des données à  $\sqrt{s} = 8 \text{ TeV}$ , soit  $20.7 \text{ fb}^{-1}$ , l'analyse et la classification des événements diphoton ont été optimisées de façon à isoler au mieux les différents modes de production et ainsi diminuer l'erreur sur la mesure des  $\mu_i$ . L'optimisation de la catégorie dédiée à la production du boson de Higgs par fusion de boson vecteur (VBF) a notamment permis d'améliorer de 25% la précision attendue sur la mesure de  $\mu_{VBF}$ . En groupant les productions du boson de Higgs par fusion de gluons (ggF) et en association avec une paire de quark top (ttH), les  $\mu_i$  mesurés sont :

$$\begin{aligned}
 - \mu_{ggF+ttH} &= 1.585_{-0.340}^{+0.364}(\text{stat}) \quad {}_{-0.168}^{+0.232}(\text{syst}) \\
 - \mu_{VBF} &= 1.808_{-0.755}^{+0.824}(\text{stat}) \quad {}_{-0.438}^{+0.549}(\text{syst}) \\
 - \mu_{VH} &= 1.813_{-1.315}^{+1.471}(\text{stat}) \quad {}_{-0.220}^{+0.502}(\text{syst})
 \end{aligned}$$

La valeur-p pour le mode de production VBF a également été mesurée; la signifi-  
 cance observée à  $m_H = 126.8$  GeV est de  $1.9\sigma$ , quand la signifi-  
 cance attendue est de  $1.2\sigma$ . La combi-  
 naison avec les autres canaux de recherche ( $H \rightarrow ZZ^* \rightarrow 4l$  et  $H \rightarrow WW \rightarrow l\nu l\nu$ ) amène à la  
 première évidence d'un boson de Higgs produit par fusion de bosons vecteurs.

A partir de la mesure des  $\mu_i$  dans les différents canaux de recherche, les constantes de cou-  
 plages de cette nouvelle particule peuvent être calculées. Bien qu'aucune déviation significative  
 par rapport aux prédictions du Modèle Standard ne soit observée pour le moment, il est encore  
 trop tôt pour affirmer que cette particule est le boson de Higgs prédit par le Modèle Standard.  
 Il reste à observer la désintégration fermionique de ce boson dans ATLAS ( $H \rightarrow \tau\tau$ ,  $H \rightarrow b\bar{b}$ )  
 et les données qui seront acquises par le LHC d'ici 2022 ( $300 \text{ fb}^{-1}$ ) permettront une mesure  
 des constantes de couplage avec une précision de l'ordre de 10%; elle sera de l'ordre de 5%  
 avec les  $3000 \text{ fb}^{-1}$  de données attendus en 2030, ce qui devrait nous permettre d'être sensible  
 à d'éventuelles déviations par rapport aux prédictions du Modèle Standard. Par ailleurs, de  
 nombreux modèles au-delà du Modèle Standard prévoient un doublet de bosons de Higgs dont  
 le plus léger présente les mêmes caractéristiques que celui prédit par le Modèle Standard. Il est  
 donc essentiel de continuer la recherche de résonance à plus haute masse pour contraindre ou  
 confirmer ces modèles.

# Annexes



## Annexe A

# Diagrammes de Feynman pour la production du boson de Higgs par fusion de gluons

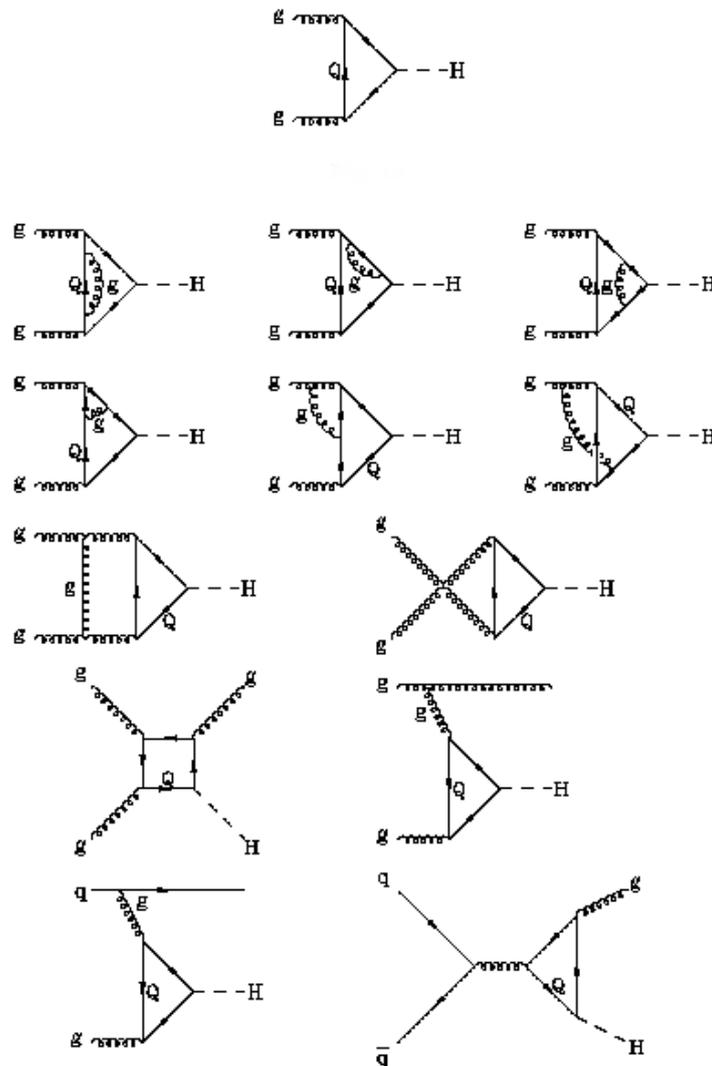
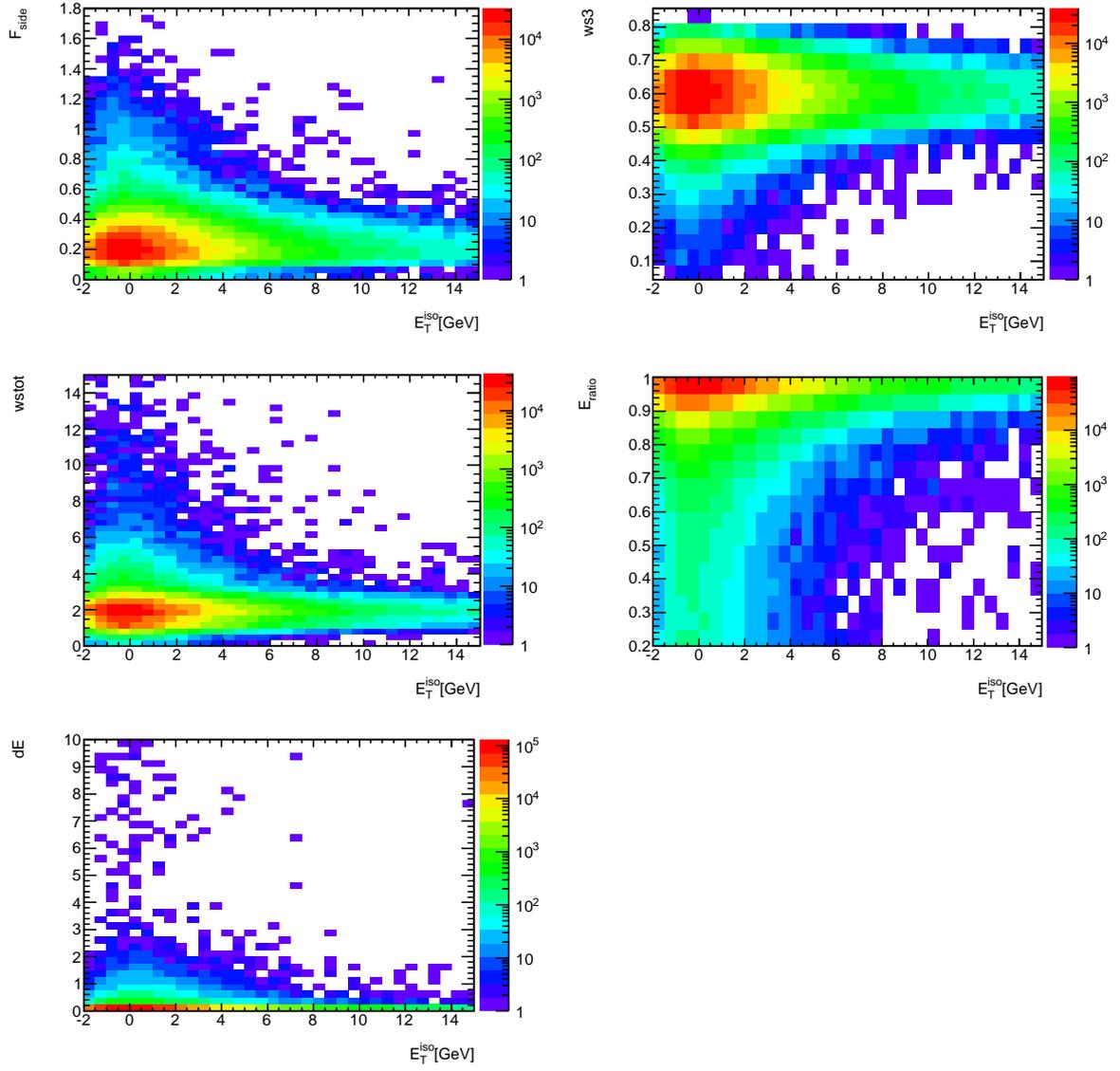


Figure A.1 – Diagramme de Feynman pour la production du boson de Higgs par fusion de gluons à l'ordre dominant et à NLO [52]

## Annexe B

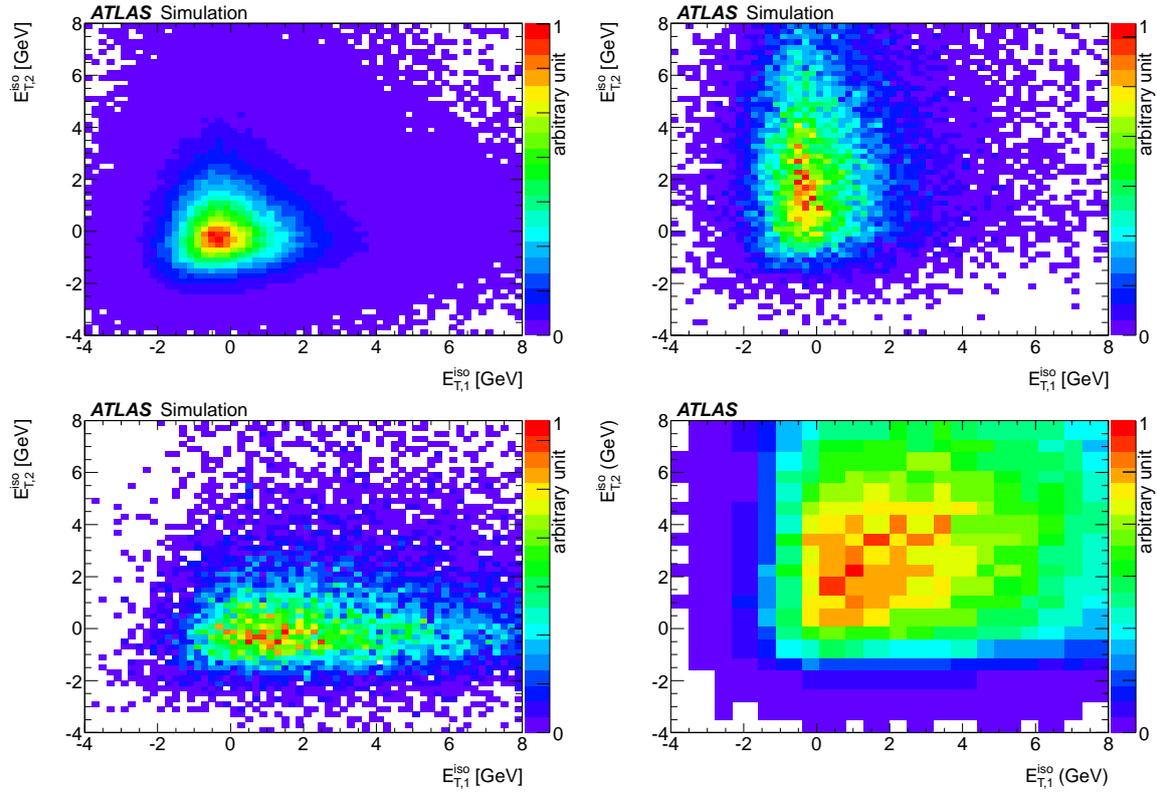
# Corrélations des variables d'identification et d'isolation des candidats photons

La figure B.1 présente les corrélations entre les variables des strips utilisées pour l'identification des photons et l'énergie d'isolation.



**Figure B.1** –  $F_{\text{side}}$ ,  $w_{s3}$ ,  $w_{\text{stot}}$ ,  $E_{\text{ratio}}$ ,  $\Delta E$  en fonction de l'énergie d'isolation pour les photons dans la simulation SHERPA.

La figure B.2 présente les corrélations linéaires d'isolation pour les événements  $\gamma\gamma$ ,  $\gamma$ -jet, jet- $\gamma$  estimées à partir de la simulation Monte-Carlo et pour les événements jet-jet estimée dans les données. Elles sont respectivement de  $3.0\pm 0.2$ ,  $2.1\pm 0.7$ ,  $4.0\pm 1.0$  et  $8.1\pm 0.4\%$ .



**Figure B.2** – Corrélation entre l'énergie d'isolation du premier et second objet. De haut en bas et de gauche à droite : événements  $\gamma$ - $\gamma$  (SHERPA), événements  $\gamma$ -jet (ALPGEN), événements jet- $\gamma$  (ALPGEN) et événements jet-jet (données) [99].

## Annexe C

# Equations utilisées pour la méthode 2×2D sideband

$$\begin{aligned}
 N_{AA} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} f_{I2} f_{T2} \xi_{j2} \\
 &+ W_{j\gamma}^{L'L'} \varepsilon_{I2} \varepsilon_{T2} f_{I1} f_{T1} \xi_{j1} \\
 &+ W_{jj}^{L'L'} f'_{I1} f'_{T1} f'_{I2} f'_{T2} \xi_{Ijj} \xi_{j1} \xi_{j2}
 \end{aligned} \tag{C.1}$$

$$\begin{aligned}
 N_{AB} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} (1 - \varepsilon_{I2}) \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} (1 - f_{I2} \xi_{j2}) f_{T2} \\
 &+ W_{j\gamma}^{L'L'} f_{I1} f_{T1} (1 - \varepsilon_{I2}) \varepsilon_{T2} \xi_{j1} \\
 &+ W_{jj}^{L'L'} f'_{I1} f'_{T1} (1 - f'_{I2} \xi_{Ijj} \xi_{j2}) f'_{T2} \xi_{j1}
 \end{aligned} \tag{C.2}$$

$$\begin{aligned}
 N_{BA} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) \varepsilon_{T1} \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{I1}) \varepsilon_{T1} f_{I2} f_{T2} \xi_{j2} \\
 &+ W_{j\gamma}^{L'L'} (1 - f_{I1} \xi_{j1}) f_{T1} \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{jj}^{L'L'} (1 - f'_{I1} \xi_{Ijj} \xi_{j1}) f'_{T1} f'_{I2} f'_{T2} \xi_{j2}
 \end{aligned} \tag{C.3}$$

$$\begin{aligned}
 N_{BB} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) (1 - \varepsilon_{I2}) \varepsilon_{T1} \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{I1}) (1 - f_{I2} \xi_{j2}) \varepsilon_{T1} f_{T2} \\
 &+ W_{j\gamma}^{L'L'} (1 - \varepsilon_{I2}) (1 - f_{I1} \xi_{j1}) \varepsilon_{T2} f_{T1} \\
 &+ W_{jj}^{L'L'} (1 - f'_{I1} \xi_{j1} - f'_{I2} \xi_{j2} + f'_{I1} f'_{I2} \xi_{Ijj} \xi_{j1} \xi_{j2}) f'_{T1} f'_{T2}
 \end{aligned} \tag{C.4}$$

$$\begin{aligned}
 N_{AC} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} \varepsilon_{I2} (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} f_{I2} (1 - f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} f_{I1} f_{T1} \varepsilon_{I2} (1 - \varepsilon_{T2}) \xi_{j1} \\
 &+ W_{jj}^{L'L'} f'_{I1} f'_{T1} f'_{I2} (1 - f'_{T2} \xi_{j2}) \xi_{Ijj} \xi_{j1}
 \end{aligned} \tag{C.5}$$

$$\begin{aligned}
 N_{CA} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) f_{I2} f_{T2} \xi_{j2} \\
 &+ W_{j\gamma}^{L'L'} f_{I1} (1 - f_{T1} \xi_{j1}) \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{jj}^{L'L'} f'_{I1} (1 - f'_{T1} \xi_{j1}) f'_{I2} f'_{T2} \xi_{Ijj} \xi_{j2}
 \end{aligned} \tag{C.6}$$

$$\begin{aligned}
 N_{AD} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} (1 - \varepsilon_{I2}) (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} \varepsilon_{T1} (1 - f_{I2} - f_{T2} + f_{I2} f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} f_{I1} f_{T1} \xi_{j1} (1 - \varepsilon_{I2}) (1 - \varepsilon_{T2}) \\
 &+ W_{jj}^{L'L'} \xi_{j1} f'_{I1} f'_{T1} (1 - \xi_{Ijj} f'_{I2} - f'_{T2} + f'_{I2} f'_{T2} \xi_{j2} \xi_{Ijj})
 \end{aligned} \tag{C.7}$$

$$\begin{aligned}
 N_{DA} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) (1 - \varepsilon_{T1}) \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{I1}) (1 - \varepsilon_{T1}) f_{I2} f_{T2} \xi_{j2} \\
 &+ W_{j\gamma}^{L'L'} (1 - f_{I1} - f_{T1} + f_{I1} f_{T1} \xi_{j1}) \varepsilon_{I2} \varepsilon_{T2} \\
 &+ W_{jj}^{L'L'} (1 - f'_{I1} \xi_{Ijj} - f'_{T1} + f'_{I1} f'_{T1} \xi_{j1} \xi_{Ijj}) f'_{I2} f'_{T2} \xi_{j2}
 \end{aligned} \tag{C.8}$$

$$\begin{aligned}
 N_{BC} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) \varepsilon_{T1} \varepsilon_{I2} (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{I1}) \varepsilon_{T1} f_{I2} (1 - f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} (1 - f_{I1} \xi_{j1}) f_{T1} \varepsilon_{I2} (1 - \varepsilon_{T2}) \\
 &+ W_{jj}^{L'L'} (1 - f'_{I1} \xi_{Ijj} \xi_{j1}) f'_{T1} f'_{I2} (1 - f'_{T2} \xi_{j2})
 \end{aligned} \tag{C.9}$$

$$\begin{aligned}
 N_{CB} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) (1 - \varepsilon_{I2}) \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} (1 - f_{I2} \xi_{j2}) (1 - \varepsilon_{T1}) f_{T2} \\
 &+ W_{j\gamma}^{L'L'} f_{I1} (1 - \varepsilon_{I2}) (1 - f_{T1} \xi_{j1}) \varepsilon_{T2} \\
 &+ W_{jj}^{L'L'} f'_{I1} (1 - f'_{T1} \xi_{j1}) (1 - f'_{I2} \xi_{Ijj} \xi_{j2}) f'_{T2}
 \end{aligned} \tag{C.10}$$

$$\begin{aligned}
 N_{CC} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) \varepsilon_{I2} (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) f_{I2} (1 - f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} f_{I1} (1 - f_{T1} \xi_{j1}) \varepsilon_{I2} (1 - \varepsilon_{T2}) \\
 &+ W_{jj}^{L'L'} f'_{I1} f'_{I2} (1 - f'_{T1} \xi_{j1} - f'_{T2} \xi_{j2} + f'_{T1} f'_{T2} \xi_{j1} \xi_{j2}) \xi_{Ijj}
 \end{aligned} \tag{C.11}$$

$$\begin{aligned}
 N_{BD} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) \varepsilon_{T1} (1 - \varepsilon_{I2}) (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} ((1 - \varepsilon_{I1}) \varepsilon_{T1} (1 - f_{I2} - f_{T2} + f_{I2} f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} (1 - f_{I1} \xi_{j1}) f_{T1} (1 - \varepsilon_{T2}) (1 - \varepsilon_{I2}) \\
 &+ W_{jj}^{L'L'} f'_{T1} (1 - f'_{I2} - f'_{T2} - f'_{I1} \xi_{j1} + f'_{I2} f'_{T2} \xi_{j2} + f'_{T2} f'_{I1} \xi_{j1} + f'_{I2} f'_{I1} \xi_{Ijj} \xi_{j1} - f'_{I1} f'_{I2} f'_{T2} \xi_{Ijj} \xi_{j1} \xi_{j2})
 \end{aligned} \tag{C.12}$$

$$\begin{aligned}
 N_{DB} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{I1}) (1 - \varepsilon_{T1}) (1 - \varepsilon_{I2}) \varepsilon_{T2} \\
 &+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{I1}) (1 - \varepsilon_{T1}) (1 - f_{I2} \xi_{j2}) f_{T2} \\
 &+ W_{j\gamma}^{L'L'} (1 - f_{I1} - f_{T1} + f_{I1} f_{T1} \xi_{j1}) (1 - \varepsilon_{I2}) \varepsilon_{T2} \\
 &+ W_{jj}^{L'L'} (1 - f'_{I1} - f'_{T1} - f'_{I2} \xi_{j2} + f'_{I1} f'_{T1} \xi_{j1} + f'_{T1} f'_{I2} \xi_{j2} + f'_{I1} f'_{I2} \xi_{Ijj} \xi_{j2} - f'_{I1} f'_{T1} f'_{I2} \xi_{Ijj} \xi_{j1} \xi_{j2}) f'_{T2}
 \end{aligned} \tag{C.13}$$

$$\begin{aligned}
 N_{CD} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) (1 - \varepsilon_{I2}) (1 - \varepsilon_{T2}) \\
 &+ W_{\gamma j}^{L'L'} \varepsilon_{I1} (1 - \varepsilon_{T1}) (1 - f_{I2} - f_{T2} + f_{I2} f_{T2} \xi_{j2}) \\
 &+ W_{j\gamma}^{L'L'} f_{I1} (1 - f_{T1} \xi_{j1}) (1 - \varepsilon_{I2}) (1 - \varepsilon_{T2}) \\
 &+ W_{jj}^{L'L'} f'_{I1} ((1 - f'_{T1} \xi_{j1}) (1 - f'_{I2} \xi_{Ijj}) - f'_{T2} (1 - f'_{T1} \xi_{j1}) (1 - f'_{I2} \xi_{j2} \xi_{Ijj}))
 \end{aligned} \tag{C.14}$$

---


$$\begin{aligned}
N_{DC} &= W_{\gamma\gamma}^{L'L'} \varepsilon_{I2} ((1 - \varepsilon_{T2})(1 - \varepsilon_{I1}) - \varepsilon_{T1}(1 - \varepsilon_{T2})(1 - \varepsilon_{I1})) \\
&+ W_{\gamma j}^{L'L'} f_{I2} ((1 - \varepsilon_{I1})(1 - f_{T2}\xi_{j2}) - \varepsilon_{T1}(1 - \varepsilon_{I1})(1 - f_{T2}\xi_{j2})) \\
&+ W_{j\gamma}^{L'L'} \varepsilon_{I2} ((1 - \varepsilon_{T2})(1 - f_{I1}) - f_{T1}(1 - f_{I1}\xi_{j1})(1 - \varepsilon_{T2})) \\
&+ W_{jj}^{L'L'} ((1 - f'_{T2}\xi_{j2})(1 - f'_{I1}\xi_{Ijj}) - f'_{T1}(1 - f'_{T2}\xi_{j2})(1 - f'_{I1}\xi_{j1}\xi_{Ijj})) f'_{I2}
\end{aligned} \tag{C.15}$$

$$\begin{aligned}
N_{DD} &= W_{\gamma\gamma}^{L'L'} (1 - \varepsilon_{T1} - \varepsilon_{T2} + \varepsilon_{T1}\varepsilon_{T2} + \varepsilon_{I2}(1 - \varepsilon_{T1} - \varepsilon_{T2} + \varepsilon_{T1}\varepsilon_{T2}) + \varepsilon_{I1}(-(1 - \varepsilon_{I2})(1 - \varepsilon_{T1}) + \varepsilon_{T2}(1 - \varepsilon_{I2})(1 - \varepsilon_{T1}))) \\
&+ W_{\gamma j}^{L'L'} (1 - \varepsilon_{T1} - f_{T2} + \varepsilon_{T1}f_{T2} + f_{I2}(1 - \varepsilon_{T1} - f_{T2}\xi_{j2} + \varepsilon_{T1}f_{T2}\xi_{j2}) + \varepsilon_{I1}((1 - \varepsilon_{T1})(1 - f_{I2}) + f_{T2}(1 - f_{I2}\xi_{j2})(1 - \varepsilon_{T1}))) \\
&+ W_{j\gamma}^{L'L'} (1 - \varepsilon_{T2} - f_{T1} + \varepsilon_{T2}f_{T1} + \varepsilon_{I2}(1 - f_{T1} - \varepsilon_{T2} + \varepsilon_{T2}f_{T1}) + f_{I1}((1 - \varepsilon_{I2})(1 - f_{T1}\xi_{j1}) + \varepsilon_{T2}(1 - \varepsilon_{I2})(1 - f_{T1}\xi_{j1}))) \\
&+ W_{jj}^{L'L'} (1 - f'_{T1} - f'_{T2} + f'_{T1}f'_{T2} + f'_{I2}(1 - f'_{T1} - f'_{T2}\xi_{j2} + f'_{T1}f'_{T2}\xi_{j2})) \\
&+ W_{jj}^{L'L'} (f'_{I1}(-(1 - f'_{I2}\xi_{Ijj})(1 - f'_{T1}\xi_{j1}) + f'_{T2}(1 - f'_{I2}\xi_{Ijj}\xi_{j2})(1 - f'_{T1}\xi_{j1})))
\end{aligned} \tag{C.16}$$

## Annexe D

# Isolation Frixione, critère discret et continu

Le critère d'isolation Frixione a été développé afin de supprimer complètement les singularités liées à l'émission d'un parton collinéaire au photon dans l'état final. Ces singularités sont normalement réabsorbées dans des fonctions de fragmentation, l'implémentation du critère de Frixione implique donc une contribution nulle des composantes de fragmentation dans le calcul des sections efficaces de production. Le critère de Frixione est un critère continu et il est impossible à transposer expérimentalement. Une version discrétisée du critère de Frixione, plus proche des considérations expérimentales, a été imaginée. Le critère discret consiste en une série limitée de cônes emboîtés  $C_{j=1,\dots,n}$ , de rayon respectif  $R_1 = R_{min} < R_2 < \dots < R_n = R_{max}$  comme illustré figure D.1. L'énergie transverse d'isolation  $E_{T,j}^{iso}$  dans chacun des cônes  $C_j$  doit être telle que  $0 < E_{T,1}^{iso} < \dots < E_{T,n}^{iso}$ . Contrairement au critère de Frixione continu, le critère discret implique une contribution de la composante de fragmentation, même si elle est supposée petite.

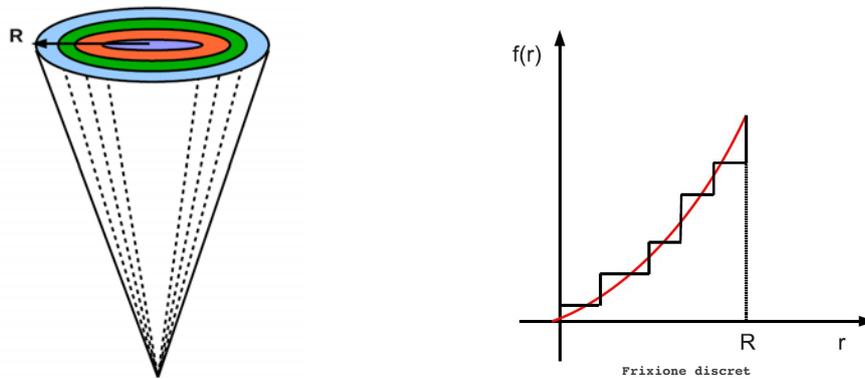
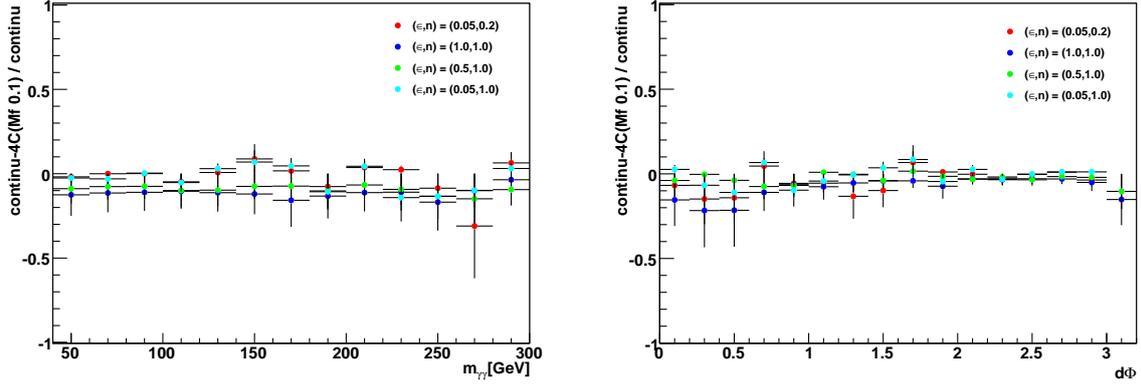


Figure D.1 – Illustration du critère de Frixione discret [165].

Cette étude concerne la production de paires de photon dans des collision  $pp$  à  $\sqrt{s} = 7$  TeV. Des coupures en  $p_T$  et  $\eta$  sont appliquées sur les deux photons,  $p_T^{\gamma,1}, p_T^{\gamma,2} > 20$  GeV et  $|\eta| < 2.5$ . Par ailleurs, comme pour l'analyse des données une distance minimale entre les deux photons,  $\Delta R_{\gamma\gamma} > 0.4$  est requise. Les PDF utilisées sont CTEQ 66 et les fonctions de fragmentation BFG set II [166]. L'échelle de renormalisation  $\mu$  et factorisation  $M$  sont fixées telles que  $\mu = M = \min(p_T^{\gamma,1}, p_T^{\gamma,2})$ . Le but de cette étude est de comparer les critères de Frixione discret et continu implémentés dans le générateur NLO Diphox et ainsi donner une estimation NLO de la contribution de la fragmentation manquante dans les calculs à NNLO de  $2\gamma$  NNLO [125] qui utilisent le critère de Frixione continu.

Le critère discret utilisé ici est basé sur quatre cônes emboîtés de rayon respectif  $R_{min} =$



**Figure D.2** – Comparaison entre les critères de Frixione discret et continu pour la distribution  $m_{\gamma\gamma}$  (gauche) et  $\Delta\phi$  (droite).

0.1,  $R_2 = 0.2$ ,  $R_3 = 0.3$  et  $R_{max} = 0.4$ . L'énergie d'isolation dans chacun d'entre eux est définie par :

$$E_T^{iso}(R) = \epsilon p_T^\gamma \left( \frac{1 - \cos R}{1 - \cos R_0} \right)^n \quad (\text{D.1})$$

Quatre variantes du profil de l'énergie d'isolation sont considérées :  $(\epsilon, n) = (0.05, 0.2)$ ,  $(0.05, 1)$ ,  $(0.5, 1)$  et  $(1, 1)$ .

La figure D.2 de gauche présente la comparaison entre critères discret et continu pour la distribution de la masse invariante des paires de photons dans l'intervalle [40-300] GeV. Avec la configuration  $(\epsilon, n) = (0.5, 1)$ , la suppression de la composante de fragmentation par le critère discret est telle que l'on n'observe pas de différence notable entre critère continu et discret. Pour la configuration  $(\epsilon, n) = (1, 1)$ , le critère discret amène une distribution  $O(10 - 12\%)$  plus grande que celle obtenue avec le critère continu. L'effet obtenu avec  $(\epsilon, n) = (1, 0.5)$  est comparable même si quantitativement moins important.

Une comparaison similaire est faite pour la variable  $d\phi$  D.2 (droite) dont la queue de distribution, à bas  $d\phi$  est très sensible à la composante de fragmentation. Les conclusions sont les mêmes que celles établies pour  $m_{\gamma\gamma}$ , même si l'on constate, à bas  $d\phi$ , un effet qualitatif plus important.

## Annexe E

# Résumé des incertitudes systématiques pour l'analyse $H \rightarrow \gamma\gamma$ de juillet 2012

L'ensemble des incertitudes systématiques considérées pour l'analyse  $H \rightarrow \gamma\gamma$  de juillet 2012 sont présentées tableau E.2 et tableau E.1 [133].

Categorie	Paramétrisation	Incertitude [ $N_{\text{evt}}$ ]	
		$\sqrt{s} = 7\text{TeV}$ $\mathcal{L} = 4.8 \text{ fb}^{-1}$	$\sqrt{s} = 8\text{TeV}$ $\mathcal{L} = 5.9 \text{ fb}^{-1}$
Inclusive	Pol. 4eme ordre	7.3	10.6
CP1 Non-converti central, bas $p_{Tt}$	Pol. exp du 2nd order	2.1	3.0
CP2 Non-converti central, haut $p_{Tt}$	Exponentielle	0.2	0.3
CP3 Non-converti non-central, bas $p_{Tt}$	Pol. 4eme ordre	2.2	3.3
CP4 Non-converti non-central, haut $p_{Tt}$	Exponentielle	0.5	0.8
CP5 Converti central, bas $p_{Tt}$	Pol. exp du 2nd order	1.6	2.3
CP6 Converti central, haut $p_{Tt}$	Exponentielle	0.3	0.4
CP7 Converti non-central, bas $p_{Tt}$	Pol. 4eme ordre	4.6	6.8
CP8 Converti non-central, haut $p_{Tt}$	Exponentielle	0.5	0.7
CP9 Converti transition	Pol. exp du 2nd order	3.2	4.6
CP10 2-jets	Exponentielle	0.4	0.6

**Tableau E.1** – Erreur systématique sur le nombre d'événements de signal fittées en raison de la fonction choisie pour la paramétrisation du bruit de fond pour l'analyse données à  $\sqrt{s} = 7 \text{ TeV}$  et à  $\sqrt{s} = 8 \text{ TeV}$  [133].

**Tableau E.2** – Résumé des erreurs systématiques sur le nombre d'événements de signal attendus et sur la résolution sur la masse invariante pour l'analyse des données à  $\sqrt{s}= 7$  TeV et à  $\sqrt{s}= 8$  TeV. Les valeurs présentées sont relatives à un boson de Higgs de masse 125 GeV [133].

Sources	$\sqrt{s} = 7$ TeV %	$\sqrt{s} = 8$ TeV %
<b>Nbr d'événements de signal</b>		
Photon identification	$\pm 8.4$	$\pm 10.8$
Effet de l'empilement sur photon rec/ID		$\pm 4$
Echelle d'énergie des photons		$\pm 0.3$
Photon isolation	$\pm 0.4$	$\pm 0.5$
Trigger		$\pm 1$
Section efficace Higgs boson (échelle)	$gg \rightarrow H : \begin{smallmatrix} +12 \\ -8 \end{smallmatrix}, \text{VBF} : \pm 0.3,$ $WH : \begin{smallmatrix} +0.2 \\ -0.8 \end{smallmatrix}, ZH : \begin{smallmatrix} +1.4 \\ -1.6 \end{smallmatrix}, ttH : \begin{smallmatrix} +3 \\ -9 \end{smallmatrix}$	$gg \rightarrow H : \begin{smallmatrix} +7 \\ -8 \end{smallmatrix}, \text{VBF} : \pm 0.2,$ $WH : \begin{smallmatrix} +0.2 \\ -0.6 \end{smallmatrix}, ZH : \begin{smallmatrix} +1.6 \\ -1.5 \end{smallmatrix}, ttH : \begin{smallmatrix} +4 \\ -9 \end{smallmatrix}$
Section efficace Higgs boson (PDF+ $\alpha_S$ )	$gg \rightarrow H + 2 \text{ jets} : \pm 25$ $gg \rightarrow H : \begin{smallmatrix} +8 \\ -7 \end{smallmatrix}, \text{VBF} : \begin{smallmatrix} +2.5 \\ -2.1 \end{smallmatrix},$ $VH : \pm 3.5, ttH : \pm 9$	$gg \rightarrow H : \begin{smallmatrix} +8 \\ -7 \end{smallmatrix}, \text{VBF} : \begin{smallmatrix} +2.6 \\ -2.8 \end{smallmatrix},$ $VH : \pm 3.5, ttH : \pm 8$
Higgs boson BR [80]		$\pm 5$
Higgs boson $p_T$ modélisation	low $p_{Tt}$ : $\pm 1.1$ , high $p_{Tt}$ : $\mp 12.5$ , 2-jets : $\mp 9$	
Evts. sous-jacent (2-jets)	VBF : $\pm 6$ , Autres : $\pm 30$	
Luminosité [87]	$\pm 1.8$	$\pm 3.6$
<b>Migration du signal entre catégorie</b>		
Matériel	Non-conv : $\pm 4$ , Conv : $\mp 3.5$	
Effet de l'empilement sur photon rec/ID	Non-conv : $\pm 3$ , Conv : $\mp 2$ ,	Non-conv : $\pm 2$ , Conv : $\mp 2$ ,
	2-jets : $\pm 2$	2-jets : $\pm 12$
Echelle d'énergie des jets	low $p_{Tt}$	
	$gg \rightarrow H : \pm 0.1, \text{VBF} : \pm 2.6,$ Autres : $\pm 0.1$	$gg \rightarrow H : \pm 0.1, \text{VBF} : \pm 2.3,$ Autres : $\pm 0.1$
	high $p_{Tt}$	
	$gg \rightarrow H : \pm 0.1, \text{VBF} : \pm 4,$ Autres : $\pm 0.1$	$gg \rightarrow H : \pm 0.1, \text{VBF} : \pm 4,$ Autres : $\pm 0.1$
	2-jets	
	$gg \rightarrow H : \mp 19, \text{VBF} : \mp 8,$ Autres : $\mp 15$	$gg \rightarrow H : \mp 18, \text{VBF} : \mp 9,$ Autres : $\mp 13$
Jet-vertex-fraction		2-jets : $\pm 13$ , Autres : $\mp 0.3$
Primary vertex selection	négligeable	
<b>Résolution sur la masse du signal</b>		
Résolution en énergie du calorimètre		$\pm 12$
Extrapolation électron photon		$\pm 6$
Effet de l'empilement sur la résolution en énergie		$\pm 4$
Sélection du vertex primaire		négligeable
<b>Mesure de la masse</b>		
Echelle d'énergie des photons		$\pm 0.6$
<b>Modélisation du bruit de fond</b>	voir tableau E.1	

## Annexe F

# Expressions du nombre d'événements de signal par mode de production et par catégorie pour l'analyse $H \rightarrow \gamma\gamma$ de mars 2013

Les nombres d'événements par mode de production et les paramètres de nuisance associés sont détaillés par catégorie.

Pour les **catégories  $c$  allant de 1 à 9** :

Les systématiques sur l'échelle d'énergie des jets, sont divisées en 6 composantes.

$$\begin{aligned}
 N_c^{ggF}(\theta_c^{ggF}) &= N_c^{ggF} \times (1 + \sigma_{pTmodel,c} \theta_{pTmodel,c}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,c}^{ggF} \theta_{JES_i,c}^{ggF}) \\
 &\times (1 + \sigma_{JER,c}^{ggF} \theta_{JER,c}^{ggF}) \\
 &\times e^{\sigma_{TH}^{ggF} \theta_{TH}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}}
 \end{aligned} \tag{F.1}$$

$$\begin{aligned}
 N_c^{VBF}(\theta_c^{VBF}) &= N_c^{VBF} \times (1 + \sigma_{pTmodel,c} \theta_{pTmodel,c}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,c}^{VBF} \theta_{JES_i,c}^{VBF}) \\
 &\times (1 + \sigma_{JER,c}^{VBF} \theta_{JER,c}^{VBF}) \\
 &\times e^{\sigma_{TH}^{VBF} \theta_{TH}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}}
 \end{aligned} \tag{F.2}$$

$$\begin{aligned}
 N_c^{WH}(\theta_c^{WH}) &= N_c^{WH} \times (1 + \sigma_{pTmodel,c} \theta_{pTmodel,c}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,c}^{VH,ttH} \theta_{JES_i,c}^{VH,ttH}) \\
 &\times (1 + \sigma_{JER,c}^{VH,ttH} \theta_{JER,c}^{VH,ttH}) \\
 &\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}}
 \end{aligned} \tag{F.3}$$

$$\begin{aligned}
 N_c^{ZH}(\theta_c^{ZH}) &= N_c^{ZH} \times (1 + \sigma_{pTmodel,c} \theta_{pTmodel,c}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,c}^{VH,ttH} \theta_{JES_i,c}^{VH,ttH}) \\
 &\times (1 + \sigma_{JER,c}^{VH,ttH} \theta_{JER,c}^{VH,ttH}) \\
 &\times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}}
 \end{aligned} \tag{F.4}$$

---


$$\begin{aligned}
N_c^{ttH}(\theta_c^{ttH}) &= N_c^{ttH} \times (1 + \sigma_{pTmodel,c} \theta_{pTmodel,c}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,c}^{VH,ttH} \theta_{JES_i,c}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,c}^{VH,ttH} \theta_{JER,c}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}}
\end{aligned} \tag{F.5}$$

Pour la catégorie "**2 jets haute masse tight**" :

$$\begin{aligned}
N_{vbf1}^{ggF}(\theta_{vbf1}^{ggF}) &= N_{vbf1}^{ggF} \times (1 + \sigma_{pTmodel,vbf1} \theta_{pTmodel,vbf1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf1}^{ggF} \theta_{JES_i,vbf1}^{ggF}) \\
&\times (1 + \sigma_{JER,vbf1}^{ggF} \theta_{JER,vbf1}^{ggF}) \\
&\times (1 + \sigma_{JVF,vbf1}^{ggF} \theta_{JVF,vbf1}^{ggF}) \\
&\times e^{\sigma_{TH,vbf1}^{ggF} \theta_{TH,2-jets}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}} \\
&\times (1 + \sigma_{\eta^*,vbf1} \theta_{\eta^*,vbf1}) \times (1 + \sigma_{d\phi,vbf1} \theta_{d\phi,vbf1}) \\
&\times e^{\sigma_{UE,vbf1}^{XH} \theta_{UE,vbf1}^{XH}}
\end{aligned} \tag{F.6}$$

$$\begin{aligned}
N_{vbf1}^{VBF}(\theta_{vbf1}^{VBF}) &= N_{vbf1}^{VBF} \times (1 + \sigma_{pTmodel,vbf1} \theta_{pTmodel,vbf1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf1}^{VBF} \theta_{JES_i,vbf1}^{VBF}) \\
&\times (1 + \sigma_{JER,vbf1}^{VBF} \theta_{JER,vbf1}^{VBF}) \\
&\times (1 + \sigma_{JVF,vbf1}^{VBF} \theta_{JVF,vbf1}^{VBF}) \\
&\times e^{\sigma_{TH,vbf1}^{VBF} \theta_{TH,2-jets}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}} \\
&\times e^{\sigma_{UE,vbf1}^{VBF} \theta_{UE,vbf1}^{VBF}}
\end{aligned} \tag{F.7}$$

$$\begin{aligned}
N_{vbf1}^{ZH}(\theta_{vbf1}^{ZH}) &= N_{vbf1}^{ZH} \times (1 + \sigma_{pTmodel,vbf1} \theta_{pTmodel,vbf1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf1}^{VH,ttH} \theta_{JES_i,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf1}^{VH,ttH} \theta_{JER,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf1}^{VH,ttH} \theta_{JVF,vbf1}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}} \\
&\times e^{\sigma_{UE,vbf1}^{XH} \theta_{UE,vbf1}^{XH}}
\end{aligned} \tag{F.8}$$

$$\begin{aligned}
N_{vbf1}^{WH}(\theta_{vbf1}^{WH}) &= N_{vbf1}^{WH} \times (1 + \sigma_{pTmodel,vbf1} \theta_{pTmodel,vbf1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf1}^{VH,ttH} \theta_{JES_i,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf1}^{VH,ttH} \theta_{JER,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf1}^{VH,ttH} \theta_{JVF,vbf1}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}} \\
&\times e^{\sigma_{UE,vbf1}^{XH} \theta_{UE,vbf1}^{XH}}
\end{aligned} \tag{F.9}$$


---

$$\begin{aligned}
N_{vbf1}^{ttH}(\theta_{vbf1}^{ttH}) &= N_{vbf1}^{ttH} \times (1 + \sigma_{pTmodel,vbf1} \theta_{pTmodel,vbf1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf1}^{VH,ttH} \theta_{JES_i,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf1}^{VH,ttH} \theta_{JER,vbf1}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf1}^{VH,ttH} \theta_{JVF,vbf1}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}} \\
&\times e^{\sigma_{UE,vbf1}^{XH} \theta_{UE,vbf1}^{XH}}
\end{aligned} \tag{F.10}$$

Pour la catégorie "**2 jets haute masse loose**" :

JVF est une migration mais on ne considère l'effet de cette migration que dans les catégories avec 2 jets, elle est négligeable dans les autres catégories.

$$\begin{aligned}
N_{vbf2}^{ggF}(\theta_{vbf2}^{ggF}) &= N_{vbf2}^{ggF} \times (1 + \sigma_{pTmodel,vbf2} \theta_{pTmodel,vbf2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf2}^{ggF} \theta_{JES_i,vbf2}^{ggF}) \\
&\times (1 + \sigma_{JER,vbf2}^{ggF} \theta_{JER,vbf2}^{ggF}) \\
&\times (1 + \sigma_{JVF,vbf2}^{ggF} \theta_{JVF,vbf2}^{ggF}) \\
&\times e^{\sigma_{TH,vbf2}^{ggF} \theta_{TH,2-jets}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}} \\
&\times (1 + \sigma_{\eta^*,vbf2} \theta_{\eta^*,vbf2}) \times (1 + \sigma_{d\phi,vbf2} \theta_{d\phi,vbf2}) \\
&\times e^{\sigma_{UE,vbf2}^{ggF} \theta_{UE,vbf2}^{ggF}}
\end{aligned} \tag{F.11}$$

$$\begin{aligned}
N_{vbf2}^{VBF}(\theta_{vbf2}^{VBF}) &= N_{vbf2}^{VBF} \times (1 + \sigma_{pTmodel,vbf2} \theta_{pTmodel,vbf2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf2}^{VBF} \theta_{JES_i,vbf2}^{VBF}) \\
&\times (1 + \sigma_{JER,vbf2}^{VBF} \theta_{JER,vbf2}^{VBF}) \\
&\times (1 + \sigma_{JVF,vbf2}^{VBF} \theta_{JVF,vbf2}^{VBF}) \\
&\times e^{\sigma_{TH,vbf2}^{VBF} \theta_{TH,2-jets}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}} \\
&\times e^{\sigma_{UE,vbf2}^{VBF} \theta_{UE,vbf2}^{VBF}}
\end{aligned} \tag{F.12}$$

$$\begin{aligned}
N_{vbf2}^{ZH}(\theta_{vbf2}^{ZH}) &= N_{vbf2}^{ZH} \times (1 + \sigma_{pTmodel,vbf2} \theta_{pTmodel,vbf2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf2}^{VH,ttH} \theta_{JES_i,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf2}^{VH,ttH} \theta_{JER,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf2}^{VH,ttH} \theta_{JVF,vbf2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}} \\
&\times e^{\sigma_{UE,vbf2}^{VH,ttH} \theta_{UE,vbf2}^{VH,ttH}}
\end{aligned} \tag{F.13}$$

$$\begin{aligned}
N_{vbf2}^{WH}(\theta_{vbf2}^{WH}) &= N_{vbf2}^{WH} \times (1 + \sigma_{pTmodel,vbf2} \theta_{pTmodel,vbf2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf2}^{VH,ttH} \theta_{JES_i,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf2}^{VH,ttH} \theta_{JER,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf2}^{VH,ttH} \theta_{JVF,vbf2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}} \\
&\times e^{\sigma_{UE,vbf2}^{VH,ttH} \theta_{UE,vbf2}^{VH,ttH}}
\end{aligned} \tag{F.14}$$

$$\begin{aligned}
N_{vbf2}^{ttH}(\theta_{vbf2}^{ttH}) &= N_{vbf2}^{ttH} \times (1 + \sigma_{pTmodel,vbf2} \theta_{pTmodel,vbf2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vbf2}^{VH,ttH} \theta_{JES_i,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf2}^{VH,ttH} \theta_{JER,vbf2}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vbf2}^{VH,ttH} \theta_{JVF,vbf2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}} \\
&\times e^{\sigma_{UE,vbf2}^{VH,ttH} \theta_{UE,vbf2}^{VH,ttH}}
\end{aligned} \tag{F.15}$$

Pour la catégorie "**2 jets basse masse**" (vh3) :

$$\begin{aligned}
N_{vh3}^{ggF}(\theta_{vh3}^{ggF}) &= N_{vh3}^{ggF} \times (1 + \sigma_{pTmodel,vh3} \theta_{pTmodel,vh3}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh3}^{ggF} \theta_{JES_i,vh3}^{ggF}) \\
&\times (1 + \sigma_{JER,vbf3}^{ggF} \theta_{JER,vbf3}^{ggF}) \\
&\times (1 + \sigma_{JVF,vh3}^{ggF} \theta_{JVF,vh3}^{ggF}) \\
&\times e^{\sigma_{TH,vh3}^{ggF} \theta_{TH,2-jets}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}} \\
&\times e^{\sigma_{UE,vh3}^{ggF} \theta_{UE,vh3}^{ggF}}
\end{aligned} \tag{F.16}$$

$$\begin{aligned}
N_{vh3}^{VBF}(\theta_{vh3}^{VBF}) &= N_{vh3}^{VBF} \times (1 + \sigma_{pTmodel,vh3} \theta_{pTmodel,vh3}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh3}^{VBF} \theta_{JES_i,vh3}^{VBF}) \\
&\times (1 + \sigma_{JER,vbf3}^{VBF} \theta_{JER,vbf3}^{VBF}) \\
&\times (1 + \sigma_{JVF,vh3}^{VBF} \theta_{JVF,vh3}^{VBF}) \\
&\times e^{\sigma_{TH,vh3}^{VBF} \theta_{TH,2-jets}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}} \\
&\times e^{\sigma_{UE,vh3}^{VBF} \theta_{UE,vh3}^{VBF}}
\end{aligned} \tag{F.17}$$

$$\begin{aligned}
N_{vbf2}^{ZH}(\theta_{vbf2}^{ZH}) &= N_{vbf2}^{ZH} \times (1 + \sigma_{pTmodel,vh3} \theta_{pTmodel,vh3}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh3}^{VH,ttH} \theta_{JES_i,vh3}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf3}^{VH,ttH} \theta_{JER,vbf3}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vh3}^{VH,ttH} \theta_{JVF,vh3}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}} \\
&\times e^{\sigma_{UE,vh3}^{VH,ttH} \theta_{UE,vh3}^{VH,ttH}}
\end{aligned} \tag{F.18}$$

$$\begin{aligned}
N_{vh3}^{WH}(\theta_{vh3}^{WH}) &= N_{vh3}^{WH} \times (1 + \sigma_{pTmodel,vh3} \theta_{pTmodel,vh3}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh3}^{VH,ttH} \theta_{JES_i,vh3}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf3}^{VH,ttH} \theta_{JER,vbf3}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vh3}^{VH,ttH} \theta_{JVF,vh3}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}} \\
&\times e^{\sigma_{UE,vh3}^{VH,ttH} \theta_{UE,vh3}^{VH,ttH}}
\end{aligned} \tag{F.19}$$

$$\begin{aligned}
N_{vh3}^{ttH}(\theta_{vh3}^{ttH}) &= N_{vh3}^{ttH} \times (1 + \sigma_{pTmodel,vh3} \theta_{pTmodel,vh3}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh3}^{VH,ttH} \theta_{JES_i,vh3}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vbf3}^{VH,ttH} \theta_{JER,vbf3}^{VH,ttH}) \\
&\times (1 + \sigma_{JVF,vh3}^{VH,ttH} \theta_{JVF,vh3}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}} \\
&\times e^{\sigma_{UE,vh3}^{VH,ttH} \theta_{UE,vh3}^{VH,ttH}}
\end{aligned} \tag{F.20}$$

Pour la catégorie "MET" (vh2) :

La systématique sur JES entraîne une variation sur la MET et donc une variation du nombre d'événements dans la catégorie MET. Même si c'est une migration, l'effet n'est pris en compte que dans la catégorie MET, il est négligeable ailleurs.

$$\begin{aligned}
N_{vh2}^{ggF}(\theta_{vh2}^{ggF}) &= N_{vh2}^{ggF} \times (1 + \sigma_{pTmodel,vh2} \theta_{pTmodel,vh2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{ggF} \theta_{JES_i,vh2}^{ggF}) \\
&\times (1 + \sigma_{JER,vh2}^{ggF} \theta_{JER,vh2}^{ggF}) \\
&\times e^{\sigma_{TH}^{ggF} \theta_{TH}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}} \\
&\times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{ggF} \theta_{MET}^{ggF})
\end{aligned} \tag{F.21}$$

$$\begin{aligned}
N_{vh1}^{VBF}(\theta_{vh2}^{VBF}) &= N_{vh2}^{VBF} \times (1 + \sigma_{pTmodel,vh2} \theta_{pTmodel,vh2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VBF} \theta_{JES_i,vh2}^{VBF}) \\
&\times (1 + \sigma_{JER,vh2}^{VBF} \theta_{JER,vh2}^{VBF}) \\
&\times e^{\sigma_{TH}^{VBF} \theta_{TH}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}} \\
&\times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VBF} \theta_{MET}^{VBF})
\end{aligned} \tag{F.22}$$

$$\begin{aligned}
N_{vh1}^{WH}(\theta_{vh2}^{WH}) &= N_{vh2}^{WH} \times (1 + \sigma_{pTmodel,vh2} \theta_{pTmodel,vh2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{JES_i,vh2}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vh2}^{VH,ttH} \theta_{JER,vh2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}} \\
&\times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{MET}^{VH,ttH})
\end{aligned} \tag{F.23}$$

$$\begin{aligned}
N_{vh1}^{ZH}(\theta_c^{ZH}) &= N_{vh2}^{ZH} \times (1 + \sigma_{pTmodel,vh2} \theta_{pTmodel,vh2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{JES_i,vh1}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vh2}^{VH,ttH} \theta_{JER,vh2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}} \\
&\times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{MET}^{VH,ttH})
\end{aligned} \tag{F.24}$$

$$\begin{aligned}
N_{vh1}^{ttH}(\theta_c^{ttH}) &= N_{vh2}^{ttH} \times (1 + \sigma_{pTmodel,vh2} \theta_{pTmodel,vh2}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{JES_i,vh2}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vh2}^{VH,ttH} \theta_{JER,vh2}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}} \\
&\times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh2}^{VH,ttH} \theta_{MET}^{VH,ttH})
\end{aligned} \tag{F.25}$$

Pour la catégorie "**lepton**" (vh1) :

$$\begin{aligned}
N_{vh1}^{ggF}(\theta_{vh1}^{ggF}) &= N_{vh1}^{ggF} \times (1 + \sigma_{pTmodel,vh1} \theta_{pTmodel,vh1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh1}^{ggF} \theta_{JES_i,vh1}^{ggF}) \\
&\times (1 + \sigma_{JER,vh1}^{ggF} \theta_{JER,vh1}^{ggF}) \\
&\times e^{\sigma_{TH}^{ggF} \theta_{TH}^{ggF}} \times e^{\sigma_{pdf}^{ggF} \theta_{pdf}^{ggF}} \\
&\times (1 + \sigma_{eres} \theta_{eres}) \times (1 + \sigma_{eid} \theta_{eid}) \\
&\times (1 + \sigma_{\mu_{id}} \theta_{\mu_{id}})
\end{aligned} \tag{F.26}$$

$$\begin{aligned}
N_{vh1}^{VBF}(\theta_{vh1}^{VBF}) &= N_{vh1}^{VBF} \times (1 + \sigma_{pTmodel,vh1} \theta_{pTmodel,vh1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh1}^{VBF} \theta_{JES_i,vh1}^{VBF}) \\
&\times (1 + \sigma_{JER,vh1}^{VBF} \theta_{JER,vh1}^{VBF}) \\
&\times e^{\sigma_{TH}^{VBF} \theta_{TH}^{VBF}} \times e^{\sigma_{pdf}^{VBF} \theta_{pdf}^{VBF}} \\
&\times (1 + \sigma_{eres} \theta_{eres}) \times (1 + \sigma_{eid} \theta_{eid}) \\
&\times (1 + \sigma_{\mu_{id}} \theta_{\mu_{id}})
\end{aligned} \tag{F.27}$$

$$\begin{aligned}
N_{vh1}^{WH}(\theta_{vh1}^{WH}) &= N_{vh1}^{WH} \times (1 + \sigma_{pTmodel,vh1} \theta_{pTmodel,vh1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh1}^{VH,ttH} \theta_{JES_i,vh1}^{VH,ttH}) \\
&\times (1 + \sigma_{JER,vh1}^{VH,ttH} \theta_{JER,vh1}^{VH,ttH}) \\
&\times e^{\sigma_{TH}^{WH} \theta_{TH}^{WH}} \times e^{\sigma_{pdf}^{WH} \theta_{pdf}^{WH}} \\
&\times (1 + \sigma_{eres} \theta_{eres}) \times (1 + \sigma_{eid} \theta_{eid}) \\
&\times (1 + \sigma_{\mu_{id}} \theta_{\mu_{id}})
\end{aligned} \tag{F.28}$$

$$\begin{aligned}
 N_{vh1}^{ZH}(\theta_c^{ZH}) &= N_{vh1}^{ZH} \times (1 + \sigma_{pTmodel,vh1} \theta_{pTmodel,vh1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh1}^{VH,ttH} \theta_{JES_i,vh1}^{VH,ttH}) \\
 &\quad \times (1 + \sigma_{JER,vh1}^{VH,ttH} \theta_{JER,vh1}^{VH,ttH}) \\
 &\quad \times e^{\sigma_{TH}^{ZH} \theta_{TH}^{ZH}} \times e^{\sigma_{pdf}^{ZH} \theta_{pdf}^{ZH}} \\
 &\quad \times (1 + \sigma_{eres} \theta_{eres}) \times (1 + \sigma_{eid} \theta_{eid}) \\
 &\quad \times (1 + \sigma_{\mu_{id}} \theta_{\mu_{id}})
 \end{aligned} \tag{F.29}$$

$$\begin{aligned}
 N_{vh1}^{ttH}(\theta_c^{ttH}) &= N_{vh1}^{ttH} \times (1 + \sigma_{pTmodel,vh1} \theta_{pTmodel,vh1}) \times \prod_{i=1}^6 (1 + \sigma_{JES_i,vh1}^{VH,ttH} \theta_{JES_i,vh1}^{VH,ttH}) \\
 &\quad \times (1 + \sigma_{JER,vh1}^{VH,ttH} \theta_{JER,vh1}^{VH,ttH}) \\
 &\quad \times e^{\sigma_{TH}^{ttH} \theta_{TH}^{ttH}} \times e^{\sigma_{pdf}^{ttH} \theta_{pdf}^{ttH}} \\
 &\quad \times (1 + \sigma_{eres} \theta_{eres}) \times (1 + \sigma_{eid} \theta_{eid,vh1}) \\
 &\quad \times (1 + \sigma_{\mu_{id}} \theta_{\mu_{id}})
 \end{aligned} \tag{F.30}$$

# Bibliographie

- [1] S.L. Glashow. Partial symmetries of weak interactions. *Nuclear Physics 22*, pages 579–588, 1961.
- [2] A. Salam and J.C. Ward. A model of leptons. *Phys. Lett. 13*, page 1264, 1967.
- [3] A. Salam and J.C. Ward. Electromagnetic and Weak Interactions. *Phys. Lett. 13*, page 168, 1964.
- [4] C.N. Yang and R.L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys. Rev. 96*, pages 191–195, 1954.
- [5] [http://fr.wikipedia.org/wiki/Modele\\_standard\\_physique\\_des\\_particules](http://fr.wikipedia.org/wiki/Modele_standard_physique_des_particules).
- [6] W. Heisenberg. Theory of Ferromagnetism. *Ztschr. Physik*, pages 619–636, 1928.
- [7] L.D. Landau and B.L. Ginzburg. *JETP 20*, page 1064, 1950.
- [8] Y. Nambu. Quasi-Particles and Gauge Invariance in the Theory of Superconductivity. *Phys. Rev. 117*, pages 648–663, 1960.
- [9] J. Goldstone. Field theories with superconductor solutions. *Nuovo Cim. 19*, pages 154–164, 1961.
- [10] F. Englert and R. Brout. Broken Symmetry and the Masses of Gauge Vector Mesons. *Phys. Rev. Lett. 13*, pages 321–323, Aug. 1964.
- [11] P.W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett. 13*, pages 508–509, Oct. 1964.
- [12] G.S. Guralnik, C.R. Hagen, and T.W. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett. 13*, pages 585–587, Nov. 1964.
- [13] D. Dicus and V. Mathur. Upper bounds on the values of masses in unified gauge theories. *Phys. Rev. D7*, pages 3111–3114, 1973.
- [14] J.F. Gunion, H.E. Haber, G. Kane, and S. Dawson. *The Higgs hunter guide*.
- [15] T. Hambye and K. Riessellmann. *SM Higgs Mass Bounds from Theory*.
- [16] LEP Working Group for Higgs boson searches, ALEPH, DELPHI, L3, and OPAL Collaborations. Search for the standard model Higgs boson at LEP. *Phys. Lett. B656*, pages 61–75, 2003.
- [17] The TEVNPH Working Group for the CDF and D0 Collaborations. Updated Combination of CDF and D0’s Searches for Standard Model Higgs Boson Production with up to 10.0 fb<sup>-1</sup> of Data. *FERMILAB-CONF-12-318-E*, 2012.
- [18] ATLAS Collaboration. Combined search for the Standard Model Higgs boson in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector. *Phys.Rev.D86 :032003*, 2012.
- [19] CMS Collaboration. Combination of SM, SM4, FP Higgs boson searches. *CMS-PAS-HIG-12-008*, 2012.
- [20] Gfitter Collaboration. <http://gfitter.desy.de/>.
- [21] <http://lepewwg.web.cern.ch/LEPEWWG/>.
- [22] <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSectionsFigures>.

- 
- [23] LHC Higgs Cross Section Working Group and S. Dittmaier and C. Mariotti and G. Passarino and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections : 2. Differential Distributions. *CERN-2012-002*, CERN, Geneva, 2012.
- [24] C.N. Yang. Selection Rules for the dematerialization of a Particle into two photons. *Phys. Rev.* *77*, pages 242–245, 1950.
- [25] L. Landau. On the angular momentum of a two-photon system. *Dokl. Akad. Nauk Ser. Fiz.* *60*, pages 207–209, 1948.
- [26] T. Binoth, J.Ph. Guillet, E. Pilon, and M. Werlen. A full next-to-leading order study of direct photon pair production in hadronic collisions. *Eur. Phys. J. C* *16*, pages 311–330, 2000.
- [27] D.H. Perkins. *Introduction to High Energy Physics*. 1987.
- [28] R.K. Ellis, W.J. Stirling, and B.R. Webber. *QCD and Collider Physics*. 1996.
- [29] A. Pitch. Aspects of Quantum Chromodynamics . *hep-ph/0001118v1*, 2000.
- [30] S.D. Drell and T.M. Yan. Partons and their applications at high energies. *Ann. Phys.* *66* :578, 1, 1971.
- [31] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys. B* *126*, page 298, 1977.
- [32] D. Stump, J. Huston, J. Pumplin, W.K. Tung, H.L. Lai, S. Kuhlman, and J.F. Owens. *JHEP* *0310*, 046, 2003.
- [33] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne. *Phys. Lett. B* *604*, 61, 2001.
- [34] ALEPH collaboration. First measurement of the quark-to-photon fragmentation function. *Z. Phys. C* *629*, pages 365–378, 1996.
- [35] E.W.N. Glover and A.G. Morgan. Measuring the photon fragmentation function at LEP. *Z. Phys. C* *62*, pages 311–322, 1994.
- [36] P. Aurenche, M. Fontannaz, J.Ph. Guillet, A. Kotikov, and E. Pilon. Is factorization for isolated photon cross sections broken? *Phys. Rev. D* *55*, pages 1124–1126, 1996.
- [37] P. Chiappeta, R. Fergani, and J.Ph. Guillet. *Z. Phys. C* *69*, page 443, 1996.
- [38] DØ Collaboration. Measurement of direct photon pair production cross sections in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV . *Phys. Lett. B* *690*, 2010.
- [39] CDF Collaboration. Measurement of the Cross Section for Prompt Isolated Diphoton Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV . *Phys. Rev. D* *84*, 2011.
- [40] ATLAS Collaboration. Measurement of the isolated di-photon cross-section in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector . *Phys. Rev. D* *85*, 2012.
- [41] CMS Collaboration. Measurement of the Production Cross Section for Pairs of Isolated Photons in  $pp$  collisions at at  $\sqrt{s} = 7$  TeV . *JHEP* *1201*, 2012.
- [42] C. Balazsd, E.L. Berger, P.M. Nadolsky, and C.P. Yuan. Calculation of prompt diphoton production cross-sections at Tevatron and LHC energies. *Phys. Rev. D* *76 (2007) 013009*, 2007.
- [43] T. Sjöstrand. Monte Carlo Generators for the LHC, Lecture 2. *Academic Training Lectures, CERN*, 2005.
- [44] S. Catani, F. Krauss, R. Kuhn, and B.R. Webber. QCD matrix elements + parton showers. *JHEP* *0111*, 063, 2001.
- [45] D.E. Soper. Next-to-leading order QCD calculations with parton showers. II : Soft singularities. *Phys. Rev. D* *69*, 2004.
- [46] K. Nagano. Quantum Chromodynamics Lectures. *France-Asia Particle Physics School*, 2011.
-

- 
- [47] Z. Nagy and D.E. Soper. Matching parton showers to NLO computations. *JHEP* 0510, 024, 2005.
- [48] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO Higgs boson production via gluon fusion matched with shower in POWHEG. *JHEP*, 0904 :002, 2009.
- [49] P. Nason and C. Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 1002 :037, 2010.
- [50] A. Djouadi, M. Spira, and P.M. Zerwas. Production of Higgs bosons in proton colliders : QCD corrections. *Phys. Lett.*, B264 :440, 1991.
- [51] S. Dawson. Radiative corrections to Higgs boson production. *Nucl. Phys.*, B359 :283, 1991.
- [52] M. Spira, A. Djouadi, D. Graudenz, and P.M. Zerwas. Higgs boson production at the LHC. *Nucl. Phys.*, B453 :17, 1995.
- [53] R. Harlander and W. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys. Rev. Lett.*, 88 :201801, 2002.
- [54] C. Anastasiou and K. Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nucl. Phys.*, B646 :220, 2002.
- [55] V. Ravindran, J. Smith, and W.L. van Neerven. NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions. *Nucl. Phys.*, B665 :325, 2003.
- [56] U. Aglietti, R. Bonciani, G. Degrossi, and A. Vicini. Two loop light fermion contribution to Higgs production and decays. *Phys. Lett.*, B595 :432, 2004.
- [57] S. Actis, G. Passarino, C. Sturm, and S. Uccirati. NLO electroweak corrections to Higgs boson production at hadron colliders. *Phys. Lett.*, B670 :12, 2008.
- [58] D. de Florian and M. Grazzini. Higgs production at the LHC : updated cross sections at  $\sqrt{s} = 8$  TeV. 2012.
- [59] C. Anastasiou, S. Buehler, F. Herzog, and A. Lazopoulos. Inclusive Higgs boson cross-section for the LHC at 8 TeV. *JHEP*, 1204 :004, 2012.
- [60] J. Baglio and A. Djouadi. Higgs production at the LHC. *JHEP*, 1103 :055, 2011.
- [61] L. Dixon, M.S. Siu. Resonance-continuum interference in the di-photon Higgs signal at the LHC. *Phys. Rev. Lett.*, 90 :252001, 2003.
- [62] M. Ciccolini, A. Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2-jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99 :161803, 2007.
- [63] M. Ciccolini, A. Denner, and S. Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Phys. Rev.*, D77 :013002, 2008.
- [64] K. Arnold et al. VBFNLO : A Parton level Monte Carlo for processes with electroweak bosons. *Comput. Phys. Commun.*, 180 :1661, 2009.
- [65] P. Bolzoni, F. Maltoni, S.O. Moch, and M. Zaro. Higgs production via vector-boson fusion at NNLO in QCD. *Phys. Rev. Lett.*, 105 :011801, 2010.
- [66] T. Han and S. Willenbrock. QCD correction to the  $pp \rightarrow WH$  and  $ZH$  total cross-sections. *Phys. Lett.*, B273 :167, 1991.
- [67] O. Brein, A. Djouadi, and R. Harlander. NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders. *Phys. Lett.*, B579 :149, 2004.
- [68] M.L. Ciccolini, S. Dittmaier, and M. Krämer. Electroweak radiative corrections to associated  $WH$  and  $ZH$  production at hadron colliders. *Phys. Rev.*, D68 :073003, 2003.
- [69] W. Beenakker et al. Higgs radiation off top quarks at the Tevatron and the LHC. *Phys. Rev. Lett.*, 87 :201805, 2001.
-

- 
- [70] W. Beenakker et al. NLO QCD corrections to  $t\bar{t}H$  production in hadron collisions. *Nucl. Phys.*, B653 :151, 2003.
- [71] S. Dawson, L.H. Orr, L. Reina, and D. Wackerth. Next-to-leading order QCD corrections to  $pp \rightarrow t\bar{t}h$  at the CERN Large Hadron Collider. *Phys. Rev.*, D67 :071503, 2003.
- [72] S. Dawson, C. Jackson, L.H. Orr, L. Reina, and D. Wackerth. Associated Higgs production with top quarks at the large hadron collider : NLO QCD corrections. *Phys. Rev.*, D68 :034022, 2003.
- [73] M. Botje et al. The PDF4LHC working group interim recommendations. 2011.
- [74] H.L. Lai, M. Guzzi, J. Huston, Z. Li, P. Nadolsky, et al. New parton distributions for collider physics. *Phys. Rev.*, D82 :074024, 2010.
- [75] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63 :189, 2009.
- [76] R. Ball R., V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al. Impact of heavy quark masses on parton distributions and LHC phenomenology. *Nucl. Phys.*, B849 :296, 2011.
- [77] A. Djouadi, J. Kalinowski, and M. Spira. HDECAY : A Program for Higgs boson decays in the Standard Model and its supersymmetric extension. *Comput. Phys. Commun.*, 108 :56, 1998.
- [78] A. Bredenstein, A. Denner, S. Dittmaier, and M.M. Weber. Precise predictions for the Higgs-boson decay  $H \rightarrow WW/ZZ \rightarrow 4$  leptons. *Phys. Rev.*, D74 :013004, 2006.
- [79] S. Actis, G. Passarino, C. Sturm, and S. Uccirati. NNLO computational techniques : The cases  $H \rightarrow \gamma\gamma$  and  $H \rightarrow gg$ . *Nucl. Phys.*, B811 :182, 2009.
- [80] A. Denner, S. Heinemeyer, I. Puljak, D. Rebuszi, and M. Spira. Standard Model Higgs-boson branching ratios with uncertainties. *Eur. Phys. J.*, C71 :1753, 2011.
- [81] LHC Higgs Cross Section Working Group and S. Dittmaier and C. Mariotti and G. Passarino and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections : 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.
- [82] K. Schindl. The injector chain for the LHC. *CERN, 1211 Geneva 23, Switzerland*, 1999.
- [83] S. Laplace, G. Unal, and J.B. de Vivie. Proceedings du cours de G. Unal à l'école de GIF 2004 (CERN). pages 8–10, 2006.
- [84] S. Van der Meer. *Internal Report CERN ISR-PRO/68-21*, 1968.
- [85] ATLAS Collaboration. Luminosity Determination in pp Collisions at  $\sqrt{s}=7$  TeV Using the ATLAS Detector at the LHC. *Eur. Phys. J. C71 (2011) 1630*, 2011.
- [86] ATLAS Collaboration. Luminosity Determination in pp Collisions at  $\sqrt{s} = 7$  TeV using the ATLAS Detector in 2011. *ATLAS-CONF-2011-116*, 2011.
- [87] ATLAS Collaboration. Improved Luminosity Determination in pp Collisions at  $\sqrt{s}=7$  TeV using the ATLAS Detector.
- [88] ATLAS Collaboration. Preliminary Luminosity Determination in pp Collisions at  $\sqrt{s}=8$  TeV. *ATL-COM-LUM-2012-013*, 2012.
- [89] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [90] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider . *JINST 3 S08003*, 2008.
- [91] ATLAS Collaboration. Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics. *CERN-OPEN-2008-020*, 2008.
- [92] D. Fournier. Liquid Argon calorimetry, Large Hadron Collider Workshop Aachen. *CERN-90-10, ECFA90-133*, 1990.
- [93] M. Arousseau. Mesure in situ de l'uniformité du calorimètre électromagnétique et recherche des premiers événements di-photons dans ALTAS. 2010.

- [94] ATLAS Collaboration. ATLAS Calorimeter Performance. *CERN/LHCC/96-40*, 2008.
- [95] ATLAS Collaboration. ATLAS liquid-argon calorimeter : Technical Design Report. *CERN-LHCC-96-041*, 1996.
- [96] ATLAS Collaboration. The monitoring and data quality assessment of the ATLAS liquid argon calorimeter. *ATL-COM-LARG-2013-004*, 2013.
- [97] ATLAS Collaboration. ATLAS Combined Testbeam : Computation and Validation of the Electronic Calibration Constants for the Electromagnetic Calorimeter. *ATL-LARG-PUB-2006-003*, 2006.
- [98] J. Colas and al. Electronics calibration board for the ATLAS liquid argon calorimeters. *Nuclear Instruments and Methods in Physics Research A*, 593, 2008.
- [99] ATLAS Collaboration. Measurement of isolated-photon pair production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector . *JHEP01(2013)086*, 2013.
- [100] ATLAS Collaboration. Calorimeter clustering algorithms : description and performance. *ATL-LARG-PUB-2008-002,ATL-COM-LARG-2008-003*, 2008.
- [101] ATLAS Collaboration. Expected photon performance in the ATLAS experiment. *ATL-PHYS-PUB-2011-007*, 2011.
- [102] ATLAS Collaboration. <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20110512/CalibratedZee/ATL-COM-PHYS-2011-1637/index.html>.
- [103] ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. *Eur. Phys. J. C72 (2012) 1909*, 2012.
- [104] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally invariant Kt clustering algorithms for hadron hadron collisions. *Nucl. Phys. B 406 (1993) 187-22*, 1993.
- [105] T. Barillari and al. Local Hadronic Calibration.
- [106] ATLAS Collaboration. <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20120404/TopoIsolation/ATL-COM-PHYS-2012-362/ATL-COM-PHYS-2012-362.pdf>.
- [107] ATLAS Collaboration. Calorimeter isolation and pile-up. *ATL-COM-PHYS-2012-467*, 2012.
- [108] ATLAS Collaboration. Measurements of the photon identification efficiency with the ATLAS detector using  $4.9 \text{ fb}^{-1}$  of  $pp$  collision data collected in 2011. *ATLAS-CONF-2012-123*, 2012.
- [109] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. TMVA 4, toolkit for multivariate data analysis with root, users guide. *CERN-OPEN-2007-007*, 2009.
- [110] ATLAS Collaboration. Signal studies in H to gamma gamma search with  $4.9 \text{ fb}^{-1}$  of 7 TeV data. *ATLAS-COM-PHYS-2012-501*, 2012.
- [111] J.C. Collins and D.E. Soper. Angular distribution of dileptons in high-energy hadron collisions. *Phys. Rev. D*, page 2219, 1977.
- [112] T. Sjöstrand et al. High-Energy-Physics Event Generation with PYTHIA 6.1. *Comput. Phys. Commun.*, 135 :238, 2001.
- [113] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP*, 2009.
- [114] A. Sherstnev and R.S. Thorne. Parton distributions for LO generators. *Eur. Phys. J. C*, 55 :553, 2008.
- [115] M.L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307 :001, 2003.
- [116] ATLAS Collaboration. Performance of the ATLAS Electron and Photon Trigger in  $pp$  Collisions at  $\sqrt{s} = 7$  TeV in 2011 . *ATLAS-CONF-2012-048*, 2012.

- 
- [117] ATLAS Collaboration. Measurement of the isolated di-photon cross section in  $4.9 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 7 \text{ TeV}$  with the ATLAS detector. *ATL-COM-PHYS-2012-592*, 2012.
- [118] ATLAS Collaboration. Purity Estimates for the Inclusive Isolated Photons. *ATL-PHYS-INT-2011-015*, 2011.
- [119] ATLAS Collaboration. Performance of the ATLAS trigger system in 2010. *Eur. Phys. J. C72 (2012) 1849*, 2012.
- [120] G. D’Agostini. A multidimensional unfolding method based on Bayes’ theorem. *Nucl. Instrum. Methods A*, 362 :487, 1995.
- [121] S. Schmitt. TUnfold : an algorithm for correcting migration effects in high energy physics. 2012.
- [122] T. Binoth, J.Ph. Guillet, E. Pilon, and M. Werlen. Beyond leading order effects in photon pair production at the Fermilab Tevatron. *Phys. Rev. D*, 63, 2001.
- [123] J. Pumplin et al. New Generation of Parton Distributions with Uncertainties from Global QCD Analysis. *JHEP*, 07 :012, 2002.
- [124] R.M. Whalley, D. Bourilkov, and R.C. Group. The Les Houches accord PDFs (LHAPDF) and LHAGLUE.
- [125] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini. *arXiv :1110.2375*.
- [126] S. Frixione. Isolated photons in perturbative QCD. *Phys. Lett. B429*, pages 369–374, 1998.
- [127] JA. Maestre, S. Alioli, JR Andersen, RD. Ball, A. Buckley, M. Cacciari, F. Campanario, N. Chanon, G. Chachamis, V. Ciulli, F. Cossutti, and al. The SM and NLO Multileg and SM MC Working Groups : Summary Report. *arXiv :1203.6803*, 2012.
- [128] A. Buckley, J. Butterworth, L. Lonnblad, H. Hoeth, J. Monk, et al. Rivet user manual. 2010.
- [129] S. Agostinelli et al. Geant4 - a simulation toolkit. *Nucl. Instrum. Methods A*, page 250, 2003.
- [130] H. Stefan, S. Steffen, and F. Siegert. Hard photon production and matrix-element parton-shower merging. 2010.
- [131] CDF Collaboration. Measurement of the Cross Section for Prompt Isolated Diphoton Production with  $9.5 \text{ fb}^{-1}$  of CDF Run II Data. *CDF note 10160*, 2012.
- [132] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B.716*.
- [133] ATLAS Collaboration. Observation of an excess of events in the search for the Standard Model Higgs boson in the  $\gamma\gamma$  channel with the ATLAS detector. *ATLAS-CONF-2012-091*, 2012.
- [134] N. Andari. Observation of a BEH-like boson decaying into two photons with the ATLAS detector at the LHC. *CERN-THESIS-2012-144*, 2012.
- [135] ATLAS Collaboration. Search for the Higgs Boson in the Diphoton Decay Channel with  $1.08 \text{ fb}^{-1}$ . *ATLAS-COM-PHYS-2011-786*, 2011.
- [136] ATLAS Collaboration. Search for the Standard Model Higgs boson in the diphoton decay channel with  $4.9 \text{ fb}^{-1}$  of ATLAS data at  $\sqrt{s} = 7 \text{ TeV}$ . *ATLAS-CONF-2011-161*, 2011.
- [137] OPAL Collaboration. Search for Anomalous Production of Di-lepton Events with Missing Transverse Momentum in  $e^+e^-$  Collisions at  $\sqrt{s} = 161$  and  $172 \text{ GeV}$ . *Eur. Phys. J. C4 (1998)*, 1998.
- [138] M. Vesterinen and T.R. Wyatt. A Novel Technique for Studying the Z Boson Transverse Momentum Distribution at Hadron Colliders. *Nucl. Instrum. Meth. A 602 (2009)*, 2009.
-

- [139] M. Cacciari, G.P. Salam, and G. Soyez. The anti-k(t) jet clustering algorithm. *JHEP 0804 (2008) 063*, 2008.
- [140] <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013Pileup1>.
- [141] ATLAS Collaboration. Background Studies for the Search of Higgs Boson Decaying to Two Photons with  $4.83 \text{ fb}^{-1}$  with the ATLAS Experiment. *ATLAS-COM-PHYS-2012-515*, 2012.
- [142] ATLAS Collaboration. BG study for  $H \rightarrow \gamma\gamma$  with 8 TeV data. *ATLAS-COM-PHYS-2012-754*, 2012.
- [143] J. Gaiser. *Charmonium spectroscopy from radiative decays of the  $J/\psi$  and  $\psi'$* . PhD thesis, 1982. SLAC-0255.
- [144] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605 :026, 2006.
- [145] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. *Eur. Phys. J. C 71*.
- [146] N. Berger. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/Hfitter>.
- [147] L. Lyons. *Statistics for nuclear particle physicists*. 1986.
- [148] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Using the Profile Likelihood in Searches for New Physics. 2010.
- [149] S. Wilks. The Large-sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Ann. Math. Stat.*, 9 :60–62, 1938.
- [150] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B.716*.
- [151] ATLAS Collaboration. Measurements of the properties of the Higgs-like bosons in the two photon decay channel with the ATLAS detector using  $25 \text{ fb}^{-1}$  of proton-proton collision data. *ATLAS-CONF-2013-012*, 2013.
- [152] ATLAS Collaboration. Observation and study of the Higgs boson candidate in the two photon decay channel with the ATLAS detector at the LHC. *ATLAS-CONF-2012-168*, 2012.
- [153] ATLAS Collaboration. Selection for  $H \rightarrow \gamma\gamma$  analysis supporting note, for Moriond 2013. *ATL-COM-PHYS-2013-093*, 2013.
- [154] ATLAS Collaboration. Supporting Note for Coupling Measurements in Diphoton Channel, for Moriond 2013. *ATL-COM-PHYS-2013-094*, 2013.
- [155] ATLAS Collaboration. Performance of missing transverse momentum reconstruction in ATLAS with 2011 proton-proton collisions at  $\sqrt{s} = 7 \text{ TeV}$ . *ATLAS-CONF-2012-101*, 2012.
- [156] ATLAS Collaboration. Multivariate analysis for the search of VBF Higgs to gamma gamma. *ATL-COM-PHYS-2013-076*, 2013.
- [157] F. Urs Bernlochner and D. Gillberg. <https://indico.cern.ch/conferenceDisplay.py?confId=231636>.
- [158] D. Rainwater, R. Szalapski, and D. Zeppenfeld. Probing color-singlet exchange in  $Z+2$ -jet events at the CERN LHC. *Phys. Rev. D*, 54 :6680, Dec 1996.
- [159] N. Chanon. [http://www.phys.ethz.ch/~pheno/Lectures2012\\_StatisticalTools/slides/Chanon1.pdf](http://www.phys.ethz.ch/~pheno/Lectures2012_StatisticalTools/slides/Chanon1.pdf).
- [160] J. Campbell and R.K. Ellis. MCFM for the Tevatron and the LHC. *Nucl.Phys.Proc.Suppl.*, 205-206 :10–15, 2010.
- [161] ATLAS Collaboration. Summary of ATLAS Pythia 8 tunes. *ATL-PHYS-PUB-2012-003*, Aug 2012.
- [162] N. Lorenzo-Martinez. <https://indico.cern.ch/getFile.py/access?contribId=4&resId=0&materialId=slides&confId=247480>.

- 
- [163] ATLAS Collaboration. Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to  $25\text{fb}^{-1}$  of proton-proton collision data. *ATL-CONF-2013-034*, 2013.
- [164] ATLAS Collaboration. Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC. *CERN-PH-EP-2013-103*, 2013.
- [165] J.Ph. Guillet. Photon production at hadronic colliders. Winter Workshop. 2011.
- [166] L. Bourhis, M. Fontannaz, and J.Ph. Guillet. Quark and gluon fragmentation functions into photons. *Eur. Phys. J, C2* :529–537, 1998.
- [167] Z. Bern, L. J. Dixon, and C. Schmidt. *Phys. Rev. D66, 074018*, 2002.
- [168] J. M. Campbell, R. K. Ellis, and C. Williams. *JHEP 1107, 018*, 2011.
- [169] The RooUnfold package and documentation.
- [170] ATLAS Collaboration. Measurement of the inclusive isolated prompt photon cross section in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector. *Phys. Rev. D83 (2011) 052005*, 2011.
- [171] S. Catani, D. de Florian, M. Grazzini, and P. Nason. Soft gluon resummation for Higgs boson production at hadron colliders. *JHEP*, 0307 :028, 2003.
- [172] L. Lyons. Open statistical issues in particle physics. *The annual of applied statistics*, 2 :887–915.
- [173] E. Bagnaschi, G. Degrossi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 1202 :088, 2012.
- [174] S. Gangal and F. Tackmann. NLO Uncertainties in Higgs + 2 jets from Gluon Fusion. 2013.
- [175] G. Belanger, F. Fawzi, and J.Ph. Guillet. Physics at TeV colliders. Proceedings, 6th Workshop, dedicated to Thomas Binoth, Les Houches, France. 2009.
- [176] ATLAS Collaboration. Signal studies for  $H \rightarrow \gamma\gamma$ . *ATLAS-COM-PHYS-2012-501*, 2012.
- [177] S. Elitzur. Impossibility of Spontaneously Breaking Local Symmetries. *Phys. Rev. D12*, pages 3978–3983, 1975.

# Résumé

Depuis 2010, les performances du LHC sont remarquables, plus de  $25 \text{ fb}^{-1}$  de données ont été délivrées et enregistrées par ATLAS. Le détecteur ATLAS est une expérience polyvalente, l'un de ses principaux but était la découverte de la dernière particule du Modèle Standard non-observée : le boson de Higgs. Dans la recherche du boson de Higgs dans le canal diphoton, la production de paires de photons non résonnantes constitue le bruit de fond irréductible. La mesure de section efficace différentielle de production de paires de photons fournit donc une solide compréhension du bruit de fond dans le canal  $H \rightarrow \gamma\gamma$  mais c'est également un excellent test pour les prédictions de QCD perturbative.

Avec l'ensemble des données collectées en 2011 à  $\sqrt{s} = 7 \text{ TeV}$ , soit  $5 \text{ fb}^{-1}$ , nous avons effectué la mesure de section efficace différentielle en fonction de quatre variables cinématiques, la masse invariante du système diphoton, l'impulsion transverse, l'angle azimuthal et le cosinus de l'angle polaire entre les deux photons. Le bruit de fond lié à la mauvaise identification des jets et des électrons isolés est estimé à partir des données et soustrait à l'échantillon de données initial. Les résultats, dominés par l'erreur systématique, sont ensuite comparés aux prédictions des générateurs avec parton shower (PYTHIA, SHERPA) et aux prédictions NLO (DIPHOX) et NNLO ( $2\gamma\text{NNLO}$ ) des générateurs partoniques.

Avec  $4.9 \text{ fb}^{-1}$  de données à  $\sqrt{s} = 7 \text{ TeV}$  et  $5.9 \text{ fb}^{-1}$  à  $\sqrt{s} = 8 \text{ TeV}$ , une première observation d'une nouvelle résonance dans le canal diphoton a été faite par ATLAS. Avec l'intégralité des données, soit  $\sim 25 \text{ fb}^{-1}$ , la découverte dans le canal diphoton est établie. Afin d'évaluer d'éventuelles tensions avec les prédictions du Modèle Standard, l'erreur sur la mesure des couplages de cette particule avec les bosons vecteurs doit être réduite autant que possible ; cela a été fait avec l'ensemble de la statistique en isolant les événements caractéristiques de la production par fusion de bosons vecteurs (VBF) et de la production associée aux bosons vecteurs (WH et ZH). Aucune déviation significative par rapport aux prédictions du Modèle Standard n'a été observée à ce jour.

# Abstract

Since 2010, the LHC performances are remarkable, more than  $25 \text{ fb}^{-1}$  have been delivered and recorded by the ATLAS detector. The ATLAS detector is a multi-purpose experiment, one of its main aim was the discovery of the last Standard Model particle predicted : the Higgs boson. In the search for the Higgs boson in the diphoton channel, the non resonant photon pairs production corresponds to the irreducible background. The differential cross-section measurement of photon pair production provides then a deep understanding of the background in  $H \rightarrow \gamma\gamma$  channel and it is also an excellent test for the perturbative QCD predictions.

With the full dataset collected in 2011 at  $\sqrt{s} = 7 \text{ TeV}$ , corresponding to  $5 \text{ fb}^{-1}$ , we performed the differential cross-section measurement of photon pairs production at LHC as a function of four variables, the diphoton mass, transverse momentum, azimuthal separation, and cosine of the polar angle of the leading photon. The background arising from misidentified hadronic jets and isolated electrons, is estimated with data-driven techniques and subtracted from the data samples. The results, dominated by the systematic errors, are then compared to the prediction of parton-shower (PYTHIA, SHERPA) and parton-level NLO (DIPHOX) and NNLO ( $2\gamma\text{NNLO}$ ) generators.

With  $4.9 \text{ fb}^{-1}$  of data at  $\sqrt{s} = 7 \text{ TeV}$  and  $5.9 \text{ fb}^{-1}$  at  $\sqrt{s} = 8 \text{ TeV}$ , a first observation of a new resonance in the diphoton channel has been made by ATLAS. With the full dataset, corresponding to  $\sim 25 \text{ fb}^{-1}$ , the discovery in the  $H \rightarrow \gamma\gamma$  channel alone is established. In order to assess any tensions with the Standard Model predictions, the error on coupling measurement of this particle to vector bosons must be reduced as much as possible ; this was done with the full statistics by isolating events representative of vector boson fusion (VBF) and associated WH and ZH production. So far, no significative deviation from the Standard Model predictions has been observed.

