



Reconstruction des électrons et recherche de couplages anormaux dans le canal $pp \rightarrow ZZ \rightarrow 4l$ avec le détecteur CMS auprès du LHC

David Sabes

► To cite this version:

David Sabes. Reconstruction des électrons et recherche de couplages anormaux dans le canal $pp \rightarrow ZZ \rightarrow 4l$ avec le détecteur CMS auprès du LHC. Physique des Hautes Energies - Expérience [hep-ex]. Ecole Polytechnique X, 2012. Français. NNT : . pastel-00764836

HAL Id: pastel-00764836

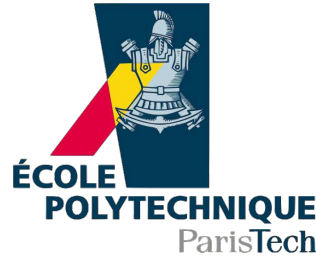
<https://pastel.hal.science/pastel-00764836>

Submitted on 13 Dec 2012

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

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

ÉCOLE POLYTECHNIQUE



THÈSE

présentée pour obtenir le grade de

DOCTEUR EN SCIENCES

par

David Sabes

**Reconstruction des électrons et recherche de
couplages anormaux dans le canal
 $pp \rightarrow ZZ \rightarrow 4l$ avec le détecteur CMS auprès
du LHC.**

Soutenue le 19 Octobre 2012 devant le jury composé de :

M	Jean-Claude	Brient	Président
Mme	Lydia	Iconomidou-Fayard	Rapporteur
M	Gautier	Hamel de Monchenault	Rapporteur
M	Florian	Beaudette	Examineur
M	Grégory	Moreau	Examineur
M	Claude	Charlot	Directeur de thèse

À mes parents

Le vrai bonheur coûte peu ; s'il est cher, il n'est pas d'une bonne espèce.

François-René de Chateaubriand

Table des matières

Remerciements	11
Introduction	13
1 Contexte théorique, Modèle Standard et couplages anormaux	15
1.1 Modèle Standard	16
1.1.1 Fermions et bosons	17
1.1.2 Symétries et interactions fondamentales	18
1.1.2.1 L'interaction électromagnétique	18
1.1.2.2 L'interaction faible	19
1.1.2.3 L'unification électrofaible	20
1.1.2.4 Le mécanisme de Higgs	21
1.1.2.5 L'interaction forte	24
1.1.2.6 Le lagrangien du Modèle Standard	24
1.2 Couplages trilinéaires de bosons neutres	25
1.2.0.7 Notations	25
1.2.0.8 Cas des couplages trilinéaires neutres	26
1.2.0.9 Aspects expérimentaux	28
1.2.0.10 Limites expérimentales actuelles	28
2 Le LHC et le détecteur CMS	31
2.1 Le LHC : grand collisionneur de hadrons	32
2.2 L'expérience et le détecteur CMS	35
2.2.1 Le trajectrographe	37
2.2.1.1 Le détecteur à pixel	37
2.2.1.2 Le détecteur à micro-pistes	39
2.2.1.3 Présence de matériaux non sensibles et cracks.	40
2.2.2 Le calorimètre électromagnétique	41
2.2.2.1 Géométrie	42
2.2.2.2 Les photodétecteurs	43
2.2.2.3 Perte de transparence et correction laser	45
2.2.2.4 Résolution en énergie et position	46
2.2.3 Le calorimètre hadronique	46
2.2.4 L'aimant solénoïdal supraconducteur	47
2.2.5 Les chambres à muons ou spectrographe à muons	48
2.2.6 Le système de déclenchement	50

2.2.6.1	Le déclenchement de premier niveau	51
2.2.6.2	Le déclenchement de haut niveau	53
2.3	Grille de calcul et outils informatiques CMS	54
2.3.1	La grille de calcul	54
2.3.2	Les outils CMS	55
2.3.2.1	CRAB :Cms Remote Analysis Builder	55
2.3.2.2	ROOT et RooFit	56
3	La reconstruction des électrons dans CMS	57
3.1	Introduction	58
3.2	L'algorithme <i>ecaldriven</i>	59
3.2.1	Le super-agrégat	59
3.2.1.1	Partie tonneau : l'algorithme hybrid	59
3.2.1.2	Parties bouchons : l'algorithme multi5×5	60
3.2.2	La reconstruction des pré-traces et traces	61
3.3	L'algorithme <i>trackerdriven</i>	65
3.3.1	Reconstruction des pré-traces électron	65
3.3.2	Les agrégats "particle-flow" et l'équivalent du super-agrégat	66
3.4	Combinaison des deux algorithmes	67
3.4.1	L'algorithme	67
3.4.2	Les attributs de l'électron reconstruit.	68
3.4.3	Performances	72
4	Validation des algorithmes et mesure de l'efficacité de reconstruction des électrons dans CMS	75
4.1	Démarrage, certification des données et mesure des distributions des électrons	76
4.1.1	DQMOffline : certification des données	76
4.1.2	Mesure des distributions des électrons et comparaison avec le Monte-Carlo, données 2011	77
4.2	Sélection des événements et effet sur la mesure d'efficacité	82
4.3	Mesure de l'efficacité de reconstruction des électrons	86
4.3.1	La méthode "tag-and-probe" et mesure d'efficacité par simple comptage	86
4.3.1.1	Efficacité et calcul des incertitudes statistiques	87
4.3.1.2	Résultats	87
4.3.2	Mesure de l'efficacité par une méthode d'ajustement	88
4.3.2.1	Fonctions de modélisation signal et bruit de fond	89
4.3.2.2	Résultats obtenus avec les données 2010	91
4.3.2.3	Résultats obtenus avec les données 2011	94
4.3.2.4	Estimation des incertitudes systématiques	95
4.4	Résultats finaux et conclusion	100
5	Développement d'un algorithme de récupération des radiations dans l'état final	103
5.1	Modélisation de l'émission de FSR, comparaison et choix du générateur . .	105
5.2	Caractéristiques des photons de FSR	106

5.3	Récupération des photons FSR lointains	107
5.3.1	Étude au niveau générateur	107
5.3.1.1	Sélection	107
5.3.1.2	Résultats	108
5.3.2	Étude sur les événements reconstruits	109
5.3.2.1	Sélection	109
5.3.2.2	Performances	111
5.4	Récupération de la composante FSR dans un cône étroit autour du lepton	113
5.4.1	Sélection	113
5.4.2	Performances	114
5.5	Validation sur les données.	115
6	Analyse ZZ et couplages anormaux trilinéaires neutres	117
6.1	Mesure de di-bosons ZZ	118
6.1.1	États finaux $ZZ \rightarrow 4l$	118
6.1.2	Données expérimentales et échantillons Monte-Carlo	120
6.1.2.1	Données expérimentales CMS	120
6.1.2.2	Échantillons Monte-Carlo	121
6.1.3	Sélection et performances	122
6.1.3.1	Sélection des événements	122
6.1.3.2	Performances de sélection	124
6.1.4	Contrôle des bruits de fond	126
6.1.5	Distributions différentielles pour les candidats ZZ.	127
6.2	Recherche de couplages trilinéaires neutres	134
6.2.1	Méthodologie	134
6.2.1.1	Choix et validation du générateur	134
6.2.1.2	Méthode de repondération des distributions	135
6.2.1.3	Choix de la variable discriminante	137
6.2.2	Interprétation statistique, modèles simulés et résultats	138
6.2.2.1	Méthode CL_s	139
6.2.2.2	Limites d'exclusion	141
	Conclusion	147

Remerciements

Je remercie l'École Doctorale de l'École Polytechnique qui m'a permis d'effectuer cette thèse.

Je remercie Madame et Messieurs les membres du jury d'avoir accepté d'évaluer mon travail de thèse. Je remercie particulièrement Madame Lydia Iconomidou-Fayard, Directrice de recherche au Laboratoire de l'Accélérateur Linéaire dans l'expérience ATLAS ainsi que Monsieur Gautier Hamel de Montchenault, Directeur de recherche au CEA dans l'expérience CMS, d'avoir accepté d'être rapporteurs de ce manuscrit ainsi que pour leurs observations constructives. Je remercie, Monsieur Le Directeur, Monsieur Jean-Claude Brient pour m'avoir accueilli au sein du Laboratoire Leprince-Ringuet et d'avoir présider ce jury. Je remercie également Monsieur Florian Beaudette, Docteur au Laboratoire Leprince-Ringuet, et Monsieur Grégory Moreau, Maître de conférence à l'université Paris XI, d'avoir accepté d'examiner le travail réalisé au cours de cette thèse.

Je tiens à exprimer ma gratitude et toute ma reconnaissance à Monsieur Claude Charlot, mon directeur de thèse, pour son enseignement, ses conseils inestimables, sa disponibilité et ses encouragements durant ces trois années.

Je tiens aussi à remercier Monsieur Philippe Busson, responsable du groupe CMS du Laboratoire Leprince-Ringuet au commencement de ma thèse, et Monsieur Yves Sirois, son successeur, pour la confiance qu'ils m'ont témoignée.

J'adresse mes sincères remerciements à l'ensemble des membres de l'équipe CMS, Stéphanie Baffioni-Barillon, Christophe Ochando, Pascal Paganini, Michal Bluj, Roberto Salerno et Alexandre Zabi, pour leur collaboration et les échanges conviviaux que nous avons eus. Un grand merci à David Chamont et Andrea Sartirana du service informatique.

Je souhaite également remercier mes camarades docteurs, Andrea, Arabella, Clémentine, Lorenzo, les doctorants qui nous suivent, Ivo, Marco, Mykhailo, Nadir, et Roko pour leur amitié et les petits moments de détente partagés autour d'un café ou lors de pique-niques, ainsi que les membres "ions lourds" de CMS : Alice, Camelia, Lamia, Matt, Raphaël et Torsten. Merci à Hélène, Jao-Pedro, Johnathan, Marc... pour les matchs de basket, rares efforts sportifs durant ma thèse. Merci à l'ensemble des membres du laboratoire Leprince-Ringuet.

Je remercie tendrement mes parents pour leur affection et leur soutien bienveillant.

Introduction

L'Homme confronté aux beautés et dangers de la Nature s'est sans doute très tôt demandé, lorsqu'il levait les yeux au ciel, d'où venait-il, de quoi était constitué ce qui l'entourait, quelles étaient les forces qui pouvaient bien faire se mouvoir les astres. Plus tard des philosophes grecs tentèrent de répondre à certaines de ces questions en proposant des concepts qui, permettant d'interpréter la complexité du monde, pourraient faciliter la relation de l'Homme à ces mystères. C'est ainsi que naît dans la Grèce Antique, la théorie atomiste selon laquelle l'Univers est formé à partir d'atomes, particules élémentaires, et de vide. L'étonnement et la curiosité pour l'inconnu ont conduit l'Homme à parfaire ses connaissances, abattre les barrières de l'obscurantisme et de l'ignorance. L'évolution de la pensée scientifique et les progrès techniques aidant, il développe des trésors d'ingéniosité pour mettre en place des expériences et développer des outils toujours plus extraordinaires : la lunette astronomique et le microscope hier, le grand collisionneur de hadrons (LHC) aujourd'hui. Ainsi au XX^{ème} siècle, des théories visant à expliquer de nouvelles observations et les questions qu'elles engendrent, sont élaborées : la théorie quantique des champs et le Modèle Standard. Ce Modèle Standard permet de décrire et de comprendre la structure fondamentale de la matière, à savoir les particules élémentaires qui la constituent et les forces qui régissent ces dernières. Cependant de nombreux mécanismes qu'il postule ne sont pas complètement vérifiés et il ne répond malheureusement pas à toutes les questions.

Dans cette quête de savoir, le *Large Hadron Collider* a été construit au CERN afin de prospecter dans des domaines d'énergie jamais atteints, tester le Modèle Standard, découvrir le boson de Higgs ou encore détecter les signes d'une nouvelle physique. Démarré en fin d'année 2009, le LHC, en moins de trois ans, est déjà riche de nombreux résultats dont la découverte d'un nouveau baryon Ξ_b^{*0} ainsi que celle d'un nouveau boson ayant une masse voisine de 125 GeV.c^{-2} en passe de confirmer le mécanisme de brisure spontanée de la symétrie électrofaible. Parmi les différentes études que rendent possibles les énergies mises en jeu au LHC, celle des couplages anormaux entre bosons de jauge, constituant un test naturel du Modèle Standard, sonde des valeurs de couplages jamais envisagées.

La thèse que je vais vous présenter ici a pour sujet la reconstruction des électrons et la recherche de couplages anormaux dans le canal $pp \rightarrow ZZ \rightarrow 4l$ avec le détecteur CMS. Elle débute en fin 2009 avec le démarrage du LHC, et par conséquent avec tous les défis qu'engendre le commencement d'une expérience telle que CMS, de la mise en place d'outils d'analyse à la certification des données acquises, et plus particulièrement des électrons, pour finir par une analyse de près de 5 fb^{-1} de données permettant d'extraire de nouvelles limites sur les valeurs des couplages anormaux.

Dans un premier chapitre, j'introduirai le contexte théorique de cette thèse avec d'abord un détail du Modèle Standard, des particules et interactions qu'il met en jeu, puis j'aborderai les couplages trilinéaires de bosons neutres interdits dans le Modèle Standard et dont la recherche constitue un moyen de découvrir la présence d'une nouvelle physique.

Dans un deuxième chapitre, je présenterai le grand collisionneur de hadrons puis l'expérience et le détecteur CMS avec une revue des différents sous-détecteurs qui le constitue ainsi que des infrastructures mises en place telles que le système de déclenchement ou encore la grille de calcul.

Dans un troisième chapitre, j'exposerai les algorithmes de reconstruction des électrons dans l'expérience CMS et l'étude de leurs performances que j'ai réalisée au début de ma thèse, juste avant le démarrage du LHC. En lien avec ce chapitre, je présenterai dans le quatrième tout d'abord les outils que j'ai développés et utilisés pour la validation des données, et plus particulièrement la certification des électrons, puis je détaillerai la mesure des efficacités de reconstruction des électrons que j'ai effectuée pour les données 2010 et 2011.

Dans un cinquième chapitre, j'introduirai la problématique du rayonnement dans l'état final (FSR) qui altère le spectre de masse 4 leptons et la correction que j'ai réalisée grâce à l'étude et la mise en place d'algorithmes de récupération des photons FSR.

Enfin dans le dernier chapitre j'aborderai mon canal d'étude ainsi que les caractéristiques de l'analyse ZZ et l'exposition des distributions différentielles pour les candidats qu'elle sélectionne. Ensuite ce sera la recherche de couplages anormaux et les nouvelles limites que j'ai extraites pour finir par une discussion comparative avec les résultats existants.

Je conclurai par une synthèse des résultats obtenus.

Chapitre 1

Contexte théorique, Modèle Standard et couplages anormaux

Contenu du chapitre

1.1	Modèle Standard	16
1.1.1	Fermions et bosons	17
1.1.2	Symétries et interactions fondamentales	18
1.1.2.1	L'interaction électromagnétique	18
1.1.2.2	L'interaction faible	19
1.1.2.3	L'unification électrofaible	20
1.1.2.4	Le mécanisme de Higgs	21
1.1.2.5	L'interaction forte	24
1.1.2.6	Le lagrangien du Modèle Standard	24
1.2	Couplages trilineaires de bosons neutres	25
1.2.0.7	Notations	25
1.2.0.8	Cas des couplages trilineaires neutres	26
1.2.0.9	Aspects expérimentaux	28
1.2.0.10	Limites expérimentales actuelles	28

1.1 Modèle Standard

Le Modèle Standard, permettant de décrire et de comprendre la structure fondamentale de la matière, à savoir les particules élémentaires qui la constituent et les forces qui régissent ces dernières, est une théorie de jauge non-abélienne élaborée à partir des résultats expérimentaux du siècle dernier et communément admise par la communauté scientifique. Cette théorie, qui intègre la mécanique quantique et la relativité restreinte, repose sur la notion de symétrie, c'est-à-dire l'invariance d'une grandeur physique par une transformation ; la correspondance entre symétrie et nombre quantique est démontrée par le théorème de Noether [1].

Ainsi toute la matière contenue dans l'univers est composée de particules élémentaires, les fermions, divisées en deux groupes : les quarks et les leptons. Chaque groupe compte trois paires de particules (+ leurs antiparticules), classées des plus légères et plus stables (première famille), aux plus lourdes et plus instables (troisième famille). Notons que les plus instables se désintègrent rapidement en des particules plus stables de la première famille, constituants principaux de la matière stable de l'univers.

Par ailleurs, la matière et donc nécessairement les éléments qui la constituent, sont gouvernés par quatre interactions : l'interaction forte (groupe $SU(3)_C$), l'interaction faible (groupe $SU(2)_L$), l'interaction électromagnétique (groupe $U(1)_{EM}$) et l'interaction gravitationnelle. Les interactions, forte et faible, sont à portées très limitées tandis que les forces électromagnétique et gravitationnelle sont à portées infinies. Nous pouvons parfois trouver dans la littérature la force gravitationnelle comme interaction fondamentale du Modèle Standard. Cependant cette dernière restant négligeable au niveau des particules élémentaires, elle n'est considérée que d'un point de vue conceptuel car n'ayant pu être incluse dans le groupe de symétrie de jauge, comme nous le rappelle celui à la base du Modèle Standard : $SU(3)_C \times SU(2)_L \times U(1)_Y$ avec

- C , la couleur ou charge de couleur, nombre quantique associé aux quarks et propre à l'interaction forte
- L , pour "Left" signifiant "gauche" et traduisant le fait que l'interaction faible n'agit que sur des particules de chiralité gauche
- Y , L'hypercharge, combinaison entre la charge électromagnétique et l'isospin faible.

Ces interactions résultent de l'échange d'autres particules élémentaires, les bosons de jauge, traduisant un transfert d'énergie. Chacune des forces fondamentales sont véhiculées par un ou plusieurs bosons : l'interaction forte par les gluons, l'interaction électromagnétique par le photon, l'interaction faible par les bosons $W^\pm$ et Z^0 , et la force gravitationnelle par le toujours hypothétique graviton. Le fait que les bosons vecteurs $W^\pm$ et Z^0 sont de masses non nulles est intégré dans la théorie par la brisure spontanée de la symétrie électrofaible ou mécanisme de Higgs.

1.1.1 Fermions et bosons

Nous recensons 24 particules élémentaires du Modèle Standard. Nous allons dans cette partie les présenter brièvement. Il y a d'une part, douze bosons de spin entier égal à un, autrement dénommés bosons de jauge :

- Le photon, vecteur de l'interaction électromagnétique. Rappelons que cette interaction agit sur les particules électriquement chargées et est la source de tous les phénomènes électromagnétiques.
- Les bosons W^+ et W^- , vecteurs de l'interaction faible ; interaction pouvant affecter toutes les particules, et qui est responsable notamment de la radioactivité β .
- Le boson Z^0 qui transmet une forme de l'interaction faible sans échange de charges.
- Huit gluons portant l'interaction forte, interaction affectant les quarks et les antiquarks et responsable de leur liaison à l'intérieur des baryons tels que les protons ou les neutrons, ou encore des mésons comme les pions, kaons... Par extension, l'interaction forte affecte tous les éléments composites constitués de quarks : les hadrons. Elle est ainsi responsable notamment de la cohésion du noyau atomique (cohésion entre protons et neutrons).

Nom	Masse [GeV.c ⁻²]	Charge électrique (e)	Spin
photon	0	0	1
$W^\pm$	80.4	± 1	1
Z^0	91.2	0	1
gluons	0	0	1

Table 1.1 – *Tableau des bosons vecteurs des interactions du Modèle Standard.*

D'autre part douze fermions de spin $\frac{1}{2}$, constituants de la matière et que nous pouvons distinguer en deux catégories :

- Six quarks (et leurs antiparticules) classés par saveurs, et constituant les hadrons (cf. table 1.2).

	Saveur	Masse [GeV.c ⁻²]	Charge électrique (e)	Génération
u	Up (haut)	$2.5 \cdot 10^{-3}$	$+2/3$	1 ^{ère}
d	Down (bas)	$5 \cdot 10^{-3}$	$-1/3$	1 ^{ère}
c	Charmed (charmé)	1.27	$-1/3$	2 ^{ème}
s	Strange (étrange)	$104 \cdot 10^{-3}$	$-1/3$	2 ^{ème}
t	Top (top)	172.4	$-1/3$	3 ^{ème}
b	Beauty (beauté)	4.2	$-1/3$	3 ^{ème}

Table 1.2 – *Tableau des quarks.*

- Six leptons et leurs antiparticules (cf. table 1.3).

	Lepton	Masse [GeV.c ⁻²]	Charge électrique (e)	Génération
e ⁻	électron	511.10 ⁻⁶	-1	1 ^{ère}
ν_e	neutrino électronique	≈ 0	0	1 ^{ère}
μ^-	muon	106.10 ⁻³	-1	2 ^{ème}
ν_μ	neutrino muonique	≈ 0	0	2 ^{ème}
τ^-	tau	1.777	-1	3 ^{ème}
ν_τ	neutrino taunique	≈ 0	0	3 ^{ème}

Table 1.3 – *Tableau des leptons. Notons que les neutrinos sont considérés de masses nulles dans le Modèle Standard même si le phénomène d’oscillation constaté expérimentalement démontre le contraire.*

1.1.2 Symétries et interactions fondamentales

Le Modèle Standard est régi par plusieurs symétries : d’une part les symétries continues de jauge (interaction électromagnétique, faible et forte...) que nous détaillerons plus loin et d’autre part les symétries discrètes désignées par les lettres C , P et T :

- la symétrie de conjugaison de charge, notée C , qui appliquée à une particule la transforme en son antiparticule.
- la parité P qui inverse les coordonnées spatiales.
- la symétrie d’inversion temporelle, T , qui permet d’inverser la coordonnée temporelle.

Il est à noter que des combinaisons de ces différentes symétries peuvent être réalisées pour expliquer certaines observations. Ainsi les symétries sont parfois violées : c’est le cas de l’interaction faible qui viole la parité P mais également la symétrie CP , cependant l’ensemble des phénomènes physiques décrits par le Modèle Standard restent invariants par symétrie CPT .

1.1.2.1 L’interaction électromagnétique

L’interaction électromagnétique appliquée aux particules élémentaires dans le cadre relativiste est décrite par la théorie d’électrodynamique quantique, ou QED, c’est-à-dire par la théorie quantique des champs de l’électromagnétisme. Cette dernière fait en effet le lien entre les équations de Maxwell régissant l’électromagnétisme et la mécanique quantique via un formalisme relativiste. Physiquement elle traduit l’interaction agissant sur les leptons chargés par échange de photons. Mathématiquement elle repose sur un groupe de jauge abélien $U(1)$ et noté ici $U(1)_{EM}$.

Soit A_μ , le quadrivecteur potentiel du champ électromagnétique, représentant physiquement un boson de spin 1, vecteur de l’interaction, i.e. le photon. Soumis à une transformation de jauge $U(1)$ de paramètre α il devient :

$$A'_\mu = A_\mu - \frac{1}{e} \partial_\mu \alpha \quad (1.1)$$

avec e , la charge élémentaire, constante de couplage de l'interaction électromagnétique. À ce boson est associé un terme cinétique $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ avec $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, le tenseur de champ électromagnétique nécessaire à l'invariance de jauge par la transformation U(1) décrite au-dessus.

Soit ψ le spineur de Dirac représentant la fonction d'onde $\psi(\vec{x}, t)$ associée aux fermions de masse m . L'évolution cinétique de ces fermions est décrite par le lagrangien :

$$\begin{aligned}\mathcal{L}_{Dirac} &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \\ &= \bar{\psi}(i\cancel{D} - m)\psi\end{aligned}\tag{1.2}$$

Cependant la nécessité d'invariance par transformation de jauge locale U(1) décrite dans l'équation 1.1, impose l'utilisation de la dérivée covariante $D_\mu = \partial_\mu - ieA_\mu$ à la place de ∂_μ .

Nous obtenons finalement le lagrangien de l'électrodynamique quantique, invariant par transformation de jauge U(1)_{EM}, et décrit par :

$$\begin{aligned}\mathcal{L}_{QED} &= \bar{\psi}(i\cancel{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \bar{\psi}(i\cancel{D} - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - e\bar{\psi}\gamma^\mu\psi A_\mu\end{aligned}\tag{1.3}$$

Le terme $\bar{\psi}(i\cancel{D}-m)\psi$ représente la propagation et le terme de masse des fermions. Le terme $-e\bar{\psi}\gamma^\mu\psi A_\mu$ traduit l'interaction entre les fermions et le boson vecteur de l'interaction. Nous remarquons qu'aucun terme de masse n'est associé à ce-dernier car cela briserait l'invariance de jauge. Enfin, le terme $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, induit par le caractère abélien du groupe U(1) auquel il doit rester invariant, représente la propagation libre du boson, le photon ne pouvant interagir avec lui-même, expliquant, avec sa masse nulle, la portée infinie de l'interaction électromagnétique.

1.1.2.2 L'interaction faible

L'interaction faible se manifeste physiquement par la désintégration de certaines particules, que ce soit dans la radioactivité β ou la désintégration du $\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e$ induisant l'existence de bosons de jauge chargés massifs nécessaires au couplage. Mathématiquement elle se construit à partir des générateurs du groupe de jauge SU(2), à savoir les trois matrices de Pauli :

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\tag{1.4}$$

La structure non abélienne de ce groupe explique, avec la masse non nulle des bosons vecteurs, la faible portée de cette interaction.

Par ailleurs l'interaction faible n'agit que sur des particules de chiralité gauche. Suivant cette particularité, nous pouvons représenter le fermion par la somme d'une composante gauche, plus précisément un doublet de chiralité gauche, et d'une composante droite, un singulet de chiralité droite, insensible à l'interaction faible. La symétrie SU(2) n'étant valable que pour la composante gauche, le groupe de symétrie associé prend la notation SU(2)_L, avec L pour *Left*. Ainsi chacune des trois générations de fermions peut être

représentée par deux doublets et deux singulets avec l'exemple de la première génération :

$$\begin{aligned} leptons &= \begin{pmatrix} \nu \\ e \end{pmatrix}_L, e_R \\ quarks &= \begin{pmatrix} u \\ d \end{pmatrix}_L, u_R, d_R \end{aligned}$$

Notons que le Modèle Standard ne décrit pas le neutrino d'hélicité droite.

De plus l'interaction faible viole les symétries de parité P , de charge C ainsi que CP et ne conserve pas les saveurs des quarks. Les bosons vecteurs chargés W^+ et W^- , associés aux générateurs du groupe $SU(2)$ τ_1 et τ_2 , ont une masse de $80.398 \pm 0.025 \text{ GeV.c}^{-2}$. La découverte des courants neutres par l'expérience Gargamelle en 1973, atteste l'existence d'un troisième boson vecteur neutre dont la masse, déterminée par la suite et en même temps que celle des $W^\pm$ avec les expériences UA(1) et UA(2), est de $91.1876 \pm 0.0021 \text{ GeV.c}^{-2}$. Son existence n'étant pas clairement expliquée par le formalisme $SU(2)$, de la même manière que le photon n'y est pas clairement introduit, conduit à unifier les secteurs électromagnétique et faible au sein de l'interaction électrofaible.

1.1.2.3 L'unification électrofaible

La théorie électrofaible, également connue sous le nom de théorie Glashow-Weinberg-Salam ([2], [3], [4]) unifie l'interaction électromagnétique et l'interaction faible. L'unification électrofaible rend compte des champs physiques d'interaction neutres et chargés décrivant les deux interactions, en les exprimant comme des combinaisons linéaires des champs de jauge. Elle prédit l'unification des couplages, à savoir que bien que l'interaction faible soit d'intensité beaucoup plus faible que la force électromagnétique à basse énergie, elle parvient à une intensité comparable à haute énergie.

Mathématiquement cette théorie est construite sur le groupe de symétrie $SU(2)_L \times U(1)_Y$ permettant de vérifier la violation de parité par l'interaction faible tout en assurant la conservation de la charge électromagnétique. En effet l'écriture sous forme de doublets de chiralité gauche et singulets de chiralité droite pour les objets soumis à l'interaction faible ne permet plus de vérifier l'invariance de jauge avec la symétrie $U(1)_{EM}$. Nous introduisons un nouveau nombre quantique : l'isospin T , demi-entier pour la composante gauche et nulle pour la droite. Ainsi T vaut $+1/2$ pour les neutrinos et les quarks de charge $+2/3$ et $-1/2$ pour les leptons de charges négatives et les quarks de charges $-1/3$. Ce nouveau nombre quantique, ou plus précisément sa troisième composante T_3 , forme avec la charge électromagnétique Q , l'hypercharge $Y = 2(Q - T_3)$ qui nous permet de conserver la charge électromagnétique sous le groupe de symétrie $U(1)_Y$.

Soit B_μ , le champ associé au boson de jauge du groupe $U(1)_Y$. Soient W_μ^α ($\alpha = 1, 2, 3$), les champs de jauge correspondant à $SU(2)_L$. La dérivée covariante correspondant à la symétrie $SU(2)_L \times U(1)_Y$ s'écrit :

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau_\alpha}{2} W_\mu^\alpha \quad (1.5)$$

De manière comparable au cas de l'électrodynamique quantique, nous écrivons le tenseur de champ pour le groupe $U(1)_Y$ sous la forme :

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.6)$$

Le tenseur correspondant à $SU(2)_L$ s'écrit :

$$W_{\mu\nu} = W_{\mu\nu}^\alpha \frac{\tau_\alpha}{2}, \quad \text{avec } W_{\mu\nu}^\alpha = \partial_\mu W_\nu^\alpha - \partial_\nu W_\mu^\alpha + g_2 \epsilon^{\alpha\beta\gamma} W_{\mu\beta} W_{\nu\gamma} \quad (1.7)$$

Le lagrangien électrofaible sans masse s'écrit donc :

$$\mathcal{L}_{\text{électrofaible}} = \bar{\psi}(i\not{D})\psi - \frac{1}{4}(B_{\mu\nu}B^{\mu\nu} + W_{\mu\nu}^\alpha W^{\mu\nu}_\alpha) \quad (1.8)$$

Nous remarquons que si nous introduisons le terme de masse des fermions $-m\psi\bar{\psi}$ le lagrangien n'est plus invariant sous $SU(2)_L$. De même ce lagrangien ne peut expliquer le caractère massif des $W^\pm$ et Z^0 car si nous introduisons un terme de masse propre aux bosons vecteurs cela brise également l'invariance de jauge $SU(2)_L$. Pour résoudre ce problème il nous faut faire appel à la *brisure spontanée de symétrie électrofaible* ou *mécanisme de Higgs* ([5], [6]).

1.1.2.4 Le mécanisme de Higgs

Champ de Higgs et brisure spontanée de symétrie électrofaible.

La solution au problème des masses soulevé dans la partie précédente 1.1.2.3 réside en l'ajout d'un champ de Higgs, noté ϕ , qui est plus précisément un doublet de champs scalaires complexes et d'isospin faible et d'hypercharge $Y=1$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.9)$$

Nous ajoutons alors au lagrangien électrofaible une composante propre au mécanisme de Higgs, somme d'un terme de propagation et de masse faisant intervenir la dérivée covariante électrofaible définie en 1.5 et d'un potentiel du champ scalaire de Higgs $V(\phi)$:

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \quad (1.10)$$

avec

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad \text{et } \mu^2 < 0, \lambda > 0 \quad (1.11)$$

Le choix de $\mu^2 > 0$ nous donne qu'un seul état fondamental qui ne permet pas de briser la symétrie. Le choix particulier de $\mu^2 < 0$ nous donne une infinité de minima et par conséquent tout champ ϕ tel que $|\phi|^2 = -\frac{\mu^2}{2\lambda} = \frac{v}{2} = 0$ est solution, comme cela est illustré par la figure 1.1, et vérifie de plus l'invariance de jauge de la symétrie $U(1)$. Par contre dès qu'un choix arbitraire est fait pour la valeur de l'état fondamental, ou encore VEV (vaccum expectation value), cet état du vide ne vérifie plus l'invariance sous $U(1)$: une transformation de jauge de symétrie $U(1)$ le transformerait en un autre état fondamental du continuum des minima. Ce phénomène où l'état fondamental possède moins de symétries que le lagrangien originel est appelé *brisure spontanée de symétrie*. Ainsi un état du vide est déterminé par $\phi_1 = \phi_2 = \phi_4 = 0$ et $\phi_3 = v/\sqrt{2}$:

$$\phi_{\text{vev}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad \text{avec } v^2 = -\frac{\mu^2}{\lambda} \quad (1.12)$$

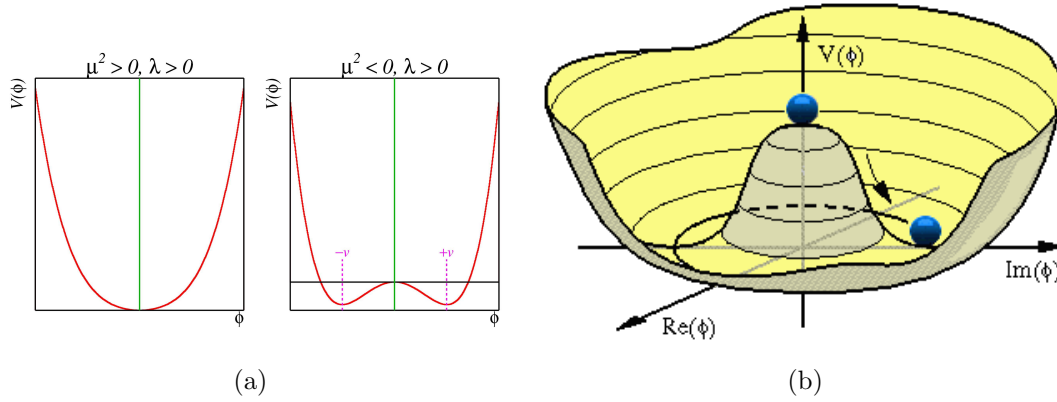


Figure 1.1 – Comparaison du potentiel $V(\phi)$ en fonction du signe du paramètre μ^2 (a). Vue du potentiel de Higgs dans le cas $\mu^2 < 0$ illustrant la dégénérescence des états fondamentaux et le phénomène de brisure spontanée (b).

Le théorème de Goldstone démontre que dès qu'il y a phénomène de brisure spontanée de symétrie il y a également apparition de particules sans masse, les bosons de Goldstone. En effet l'ensemble des états du vide possibles sont donnés par :

$$\phi'_{\text{vev}} = \phi_{\text{vev}} e^{i\frac{\tau_\alpha}{2}\theta_\alpha + iY\mathbb{1}} \quad \text{avec } \alpha = 1, 2, 3 \quad (1.13)$$

où

- τ_α sont les matrices de Pauli, générateurs de $\text{SU}(2)_L$.
- θ_α les trois champs correspondants aux bosons de Goldstone.
- Y , l'hypercharge définie par $Y = 2(Q - T_3)$

Ainsi l'état fondamental du vide n'est pas "vide" car ayant une vev non nulle. Nous devons alors décrire les excitations de ce vide comme étant dues à des quanta de champ, ce qui nous conduit à réinterpréter le champ de Higgs ϕ comme une variation autour de ϕ_{vev} :

$$\phi(x) = \frac{e^{i\tau_\alpha\theta^\alpha(x)}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.14)$$

Nous pouvons par ailleurs appliquer une transformation de jauge locale sur le champ scalaire ϕ en vue d'éliminer les bosons de Goldstone. Par exemple nous pouvons faire la transformation suivante :

$$\phi(x) \rightarrow \phi'(x) = e^{-i\alpha^\alpha(x)\frac{\tau_\alpha}{2}} \phi(x) = e^{i(\theta^a(x) - \frac{\alpha^a(x)}{2})\tau_a} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + h(x)) \end{pmatrix} \quad (1.15)$$

et choisir le cas $\alpha^a(x) = 2\theta^a(x)$. Dès lors le champ ϕ s'exprime simplement par :

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.16)$$

où $h(x)$ est le vrai champ physique de Higgs.

Les masses des bosons de jauge.

Dans le lagrangien de la brisure spontanée de symétrie électrofaible (cf éq. 1.10), le terme $(D_\mu \phi)^\dagger (D^\mu \phi)$ génère les masses de bosons de jauge par l'intermédiaire de la valeur moyenne non nulle de l'état du vide :

$$\begin{aligned} |D_\mu \phi|^2 &= \left| \left(\partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau_a}{2} W_\mu^a \right) \phi \right|^2 \\ &= \frac{1}{2} \partial_\mu h \partial^\mu h + \frac{1}{8} \left| \begin{pmatrix} g_2 W_\mu^3 + g_1 B_\mu & g_2 (W_\mu^1 - iW_\mu^2) \\ g_2 (W_\mu^1 + iW_\mu^2) & -g_2 W_\mu^3 + g_1 B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \right|^2 \end{aligned} \quad (1.17)$$

Les champs des quatre bosons de jauge sont alors définis par :

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^1 \pm iW_\mu^2) \quad (1.18)$$

$$Z_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} (g_2 W_\mu^3 - g_1 B_\mu) \quad (1.19)$$

$$A_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} (g_1 W_\mu^3 + g_2 B_\mu) \quad (1.20)$$

correspondant physiquement aux bosons $W^\pm$, Z^0 et au photon respectivement. Les expressions de Z_μ et A_μ peuvent également être interprétées comme une rotation d'angle θ_w des champs W_μ^3 et B_μ due à la brisure spontanée de symétrie :

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_w & \sin \theta_w \\ -\sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (1.21)$$

avec θ_w , l'angle de Weinberg défini par :

$$\cos \theta_w = \frac{g_2}{\sqrt{g_1^2 + g_2^2}} \quad \text{et} \quad \sin \theta_w = \frac{g_1}{\sqrt{g_1^2 + g_2^2}} \quad (1.22)$$

Notons également qu'il vérifie $\cos \theta_w = \frac{M_W}{M_Z}$.

Les masses des fermions

Comme nous l'avons précédemment souligné, l'ajout d'un terme de masse pour les fermions de la forme $-m\psi\bar{\psi}$ ne permet pas de garder l'invariance de symétrie électrofaible. Le champ de Higgs solutionne également le problème des masses des fermions en introduisant les couplages de Yukawa G_ψ [7] entre le doublet de Higgs et les fermions :

$$\mathcal{L}_{Yukawa} = -G_\psi (\bar{\psi}_R \phi^\dagger \psi_L + \psi_L \phi \psi_R) \quad (1.23)$$

Ceci permet d'introduire les termes de masse des fermions $m_\psi = \frac{G_\psi v}{\sqrt{2}}$ ainsi que le couplage

$g_{H\bar{\psi}\psi} = \frac{iG_\psi}{v}$ entre fermions et champ de Higgs, proportionnel à la masse du fermion.

Si le cas des leptons est simple, étant donné que leurs composantes droites, à savoir les neutrinos, n'ont pas de masse, il n'en est pas de même avec les quarks pour lesquels nous devons introduire une matrice de mélange, également appelée matrice CKM (Cabibbo-Kobayashi-Maskawa [8], chapitre 11).

D'un point de vue expérimental, l'objectif de découverte du boson de Higgs a motivé en grande partie la construction du *Large Hadron Collider* et les expériences qui y sont menées. La découverte d'une nouvelle particule à une masse $\simeq 125 \text{ GeV.c}^{-2}$ en juillet 2012 [9] est en passe de confirmer le principe de brisure spontanée de symétrie.

1.1.2.5 L'interaction forte

L'interaction forte est décrite par la théorie de la Chromodynamique quantique, ou QCD. De manière générale elle agit sur toute particule portant une charge de couleur C , nouveau nombre quantique ($C=R,G,B$) introduit par la QCD et nécessaire pour respecter le principe de Pauli dans le cas du hadron Δ^{++} (u_R, u_G, u_B). Elle est véhiculée par des gluons et traduit l'interaction entre gluons eux-mêmes et entre quarks, expliquant notamment la cohésion des protons et neutrons.

Cette théorie utilise le groupe de jauge non abélien $SU(3)_C$ dont les générateurs sont les matrices de Gell-Mann λ_a ($a=1\dots 8$). Soient G_μ^a les champs de jauge des bosons vecteurs, i.e. des huit gluons, associés aux générateurs de $SU(3)$. Le terme de Dirac est :

$$\mathcal{L}_{Dirac} = \bar{\psi}(i\not{D} - m)\psi \quad (1.24)$$

avec la dérivée covariante $D_\mu = \partial_\mu + ig_s \frac{\lambda_a}{2} G_\mu^a$, $a = 1\dots 8$ et g_s , la constante de couplage de l'interaction forte.

Le terme de propagation des gluons est défini par $-\frac{1}{4}G_{\mu\nu}^\alpha G_{\mu\nu}^\alpha$. A la différence du cas QED, nous devons ajouter un terme supplémentaire dans l'expression de $G_{\mu\nu}^\alpha$ pour rester invariant par transformation de jauge ; ce terme additionnel introduit le couplage entre gluons :

$$G_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha - g_s f_{abc} G_\mu^b G_\nu^c \quad (1.25)$$

f_{abc} désignant les constantes de structure du groupe $SU(3)$. Notons que l'ajout d'un terme de masse $m^2 G_{\mu\nu}^\alpha G_{\mu\nu}^\alpha$ violerait l'invariance de jauge. Finalement le lagrangien complet de QCD est :

$$\mathcal{L}_{QCD} = \bar{\psi}(i\not{D} - m)\psi - \frac{1}{4}G_{\mu\nu}^\alpha G_{\mu\nu}^\alpha \quad (1.26)$$

1.1.2.6 Le lagrangien du Modèle Standard

Le Modèle Standard repose sur le groupe de symétrie $SU(3)_C \times SU(2)_L \times U(1)_Y$ dont nous avons détaillé dans les parties précédentes les différentes composantes.

La dérivée covariante du lagrangien du Modèle Standard est donnée par la formule :

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau_a}{2} W_\mu^a - ig_s \frac{\lambda_a}{2} G_\mu^a \quad (1.27)$$

Le lagrangien du Modèle Standard s'exprime finalement par :

$$\mathcal{L}_{MS} = \mathcal{L}_{fermion} + \mathcal{L}_{Yukawa} + \mathcal{L}_{jauge} + \mathcal{L}_{Higgs} \quad (1.28)$$

avec

$\mathcal{L}_{fermion} = \bar{\psi} i \gamma^\mu D_\mu \psi$, le terme de propagation des fermions :

$\mathcal{L}_{Yukawa} = -G_\psi (\bar{\psi}_R \phi^\dagger \psi_L + \psi_L \phi \psi_R)$, le terme de Yukawa donnant la masse aux fermions

$\mathcal{L}_{jauge} = -\frac{1}{4} (B_{\mu\nu} B^{\mu\nu} + W_{\mu\nu}^a W_a^{\mu\nu} + G_{\mu\nu}^a G_a^{\mu\nu})$ le terme de propagation des champs de jauge

$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) - \lambda ((\phi^\dagger \phi)^2 - v^2 \phi^\dagger \phi)$, le terme de Higgs expliquant la masse des bosons de jauge faible

1.2 Couplages trilinéaires de bosons neutres

En reprenant le lagrangien du Modèle Standard défini dans le paragraphe 1.1.2.6, et plus particulièrement le terme de propagation des champs de jauge, ce dernier comprend notamment des termes de couplage à trois ou quatre champs. Si nous détaillons le terme de couplage à trois champs, nous obtenons :

$$\begin{aligned} \mathcal{L}_3 = & ig_2 \cos \theta_w [(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+) W^{-\mu} Z^\nu - (\partial_\mu W_\nu^- - \partial_\nu W_\mu^-) W^{-\mu} Z^\nu] \\ & + ie [(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+) W^{-\mu} A^\nu - (\partial_\mu W_\nu^- - \partial_\nu W_\mu^-) W^{-\mu} A^\nu] \\ & - g_2 \cos \theta_w (\partial_\mu Z_\nu - \partial_\nu Z_\mu) W^{-\mu} W^{+\nu} \\ & - ie (\partial_\mu A_\nu - \partial_\nu A_\mu) W^{-\mu} W^{+\nu} \end{aligned} \quad (1.29)$$

Nous remarquons alors que seuls les couplages triples entre deux bosons chargés et un boson neutre (WWZ ou WW γ) sont autorisés et que les couplages neutres ZZZ ou ZZ γ sont interdits par la théorie. Nous pouvons immédiatement comprendre que l'étude des couplages anormaux neutres représente un test naturel du Modèle Standard et que toute déviation serait signe d'une nouvelle physique.

1.2.0.7 Notations

Nous allons introduire dans ce paragraphe des notations qui seront utilisées par la suite. Soit V^μ le champ des bosons de jauge neutre (i.e. photon A_μ ou Z, Z_μ). Les tenseurs $U_{\mu\nu}$ et $\tilde{U}_{\mu\nu}$ d'un champ U_μ sont définis par :

$$U_{\mu\nu} = \partial_\nu U_\mu - \partial_\mu U_\nu \quad (1.30)$$

$$\tilde{U}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} U^{\rho\sigma} \quad (1.31)$$

L'opérateur d'Alembertien se définit par :

$$\square = \partial^\mu \partial_\mu \quad (1.32)$$

Les symétries discrètes C et P agissent des manières suivantes :

- la symétrie de conjugaison de charge C transforme un champ vectoriel en son anti-symétrique :

$$\begin{aligned} CV_\mu C^{-1} &= -V_\mu \\ \implies CV_{\mu\nu} C^{-1} &= -V_{\mu\nu} = V_{\nu\mu} \\ \implies C\tilde{V}_{\mu\nu} C^{-1} &= -\tilde{V}_{\mu\nu} = \tilde{V}_{\nu\mu} \end{aligned} \quad (1.33)$$

- La parité P transforme un champ vectoriel covariant en un champ vectoriel contra-variant et inversement :

$$\begin{aligned}
PB_\mu P^{-1} &= B^\mu \\
\implies PB_{\mu\nu} P^{-1} &= B^{\mu\nu} \\
\implies P\tilde{B}_{\mu\nu} P^{-1} &= -\tilde{B}^{\mu\nu}
\end{aligned} \tag{1.34}$$

- Nous pouvons aussi démontrer que pour un terme vérifiant l'invariance de Lorentz et formé des champs vectoriels M_ν et N_ν et un tenseur $\tilde{U}_{\mu\nu}$ soumis à la symétrie P , nous obtenons :

$$PM^\mu N^\nu \tilde{U}_{\mu\nu} P^{-1} = -M_\mu N_\nu \tilde{U}^{\mu\nu} = -M^\mu N^\nu \tilde{U}_{\mu\nu} \tag{1.35}$$

1.2.0.8 Cas des couplages trilinéaires neutres

Bien que l'existence de couplages trilinéaires neutres soit interdite dans le Modèle Standard, il est cependant possible d'élaborer un lagrangien effectif au-delà du Modèle Standard incluant de tels termes, sous réserve de quelques contraintes afin d'en limiter le nombre :

- Afin de ne pas remettre en cause le principe fondamental de conservation d'énergie et d'impulsion en introduisant des couplages dépendants de l'espace-temps, nous imposerons l'invariance de Lorentz.
- Les couplages anormaux doivent respecter la symétrie de jauge électromagnétique $U(1)_{EM}$ dont le caractère abélien interdit le couplage $\gamma\gamma$ donc a fortiori le couplage $\gamma\gamma\gamma$.
- La violation de la symétrie de jauge électrofaible par les couplages anormaux sera laissée possible étant donné que cette symétrie est déjà brisée spontanément par le mécanisme de Higgs.
- L'interaction faible ne respecte pas la symétrie discrète CP dans le secteur des quarks. Nous n'appliquerons pas non plus de contraintes relatives à la conservation de C , P ou encore CP concernant les couplages trilinéaires neutres.
- Les composantes scalaires $\partial_\mu V^\mu$ (avec V^μ pour A^μ ou Z^μ) seront négligées
- Le couplage ZZZ , par respect du principe de conservation d'énergie, ne peut vérifier trois bosons sur leur couche de masse. Malgré tout, le fait de se limiter aux bosons sur leur couche de masse entraîne une perte de précision sur les valeurs de couplage négligeable en comparaison de la résolution expérimentale [10].

Ainsi le lagrangien décrivant les couplages trilineaires neutres suivant ces conditions est donné, d'après [11], par la formule suivante :

$$\begin{aligned}
\mathcal{L}_{VV'V''} \frac{M_Z^2}{e} = & - [f_4^\gamma (\partial_\mu F^{\mu\beta}) + f_4^Z (\partial_\mu Z^{\mu\beta})] Z_\alpha (\partial^\alpha Z_\beta) \\
& + [f_5^\gamma (\partial^\sigma F \sigma_\mu) + f_5^Z (\partial^\sigma Z^{\sigma\mu})] \tilde{Z}^{\mu\beta} Z_\beta \\
& - [h_1^\gamma (\partial^\sigma F \sigma_\mu) + h_1^Z (\partial^\sigma Z^{\sigma\mu})] Z_\beta F^{\mu\beta} \\
& - [h_3^\gamma (\partial_\sigma F^{\sigma\rho}) + h_3^Z (\partial_\sigma Z^{\sigma\rho})] Z_\alpha \tilde{F}_{\rho\alpha} \\
& - \left[\frac{h_2^\gamma}{M_Z^2} (\partial_\alpha \partial_\beta \partial^\rho F_{\rho\mu}) + \frac{h_2^Z}{M_Z^2} \partial_\alpha \partial_\beta (\square + M_Z^2) Z_\mu \right] Z^\alpha F^{\mu\beta} \\
& + \left[\frac{h_4^\gamma}{2M_Z^2} (\square \partial^\sigma F^{\rho\alpha}) + \frac{h_4^Z}{2M_Z^2} (\square + M_Z^2) \partial^\sigma Z^{\rho\alpha} \right] Z_\sigma \tilde{F}_{\rho\alpha} \quad (1.36)
\end{aligned}$$

Symétries discrètes

A l'aide des formules 1.33 à 1.35, nous pouvons évaluer les effets des symétries discrètes C , P et de la combinaison CP sur les couplages neutres. Les résultats sont regroupés dans la table 1.4.

Symétrie	h_1^V	h_2^V	h_3^V	h_4^V	f_4^V	f_5^V
C	non	non	non	non	non	non
P	non	non	oui	oui	non	oui
CP	oui	oui	non	non	oui	non

Table 1.4 – *Effet des symétries discrètes dans le cas des couplages trilineaires neutres. "oui", la symétrie est conservée et "non", la symétrie est violée.*

Problème d'unitarité

Les couplages anormaux créent des contributions proportionnelles à $\sqrt{s}$ et s ou engendrent des divergences qui remettent en question l'unitarité de la théorie. Ce problème peut être solutionné en multipliant les constantes de couplage par des facteurs de forme du type :

$$f_i^Z = \frac{f_{i0}^Z}{\left(1 + \frac{\hat{s}}{\Lambda^2}\right)^n} \quad (1.37)$$

avec :

- $\hat{s}$, le carré de la masse invariante du système di-bosons.
- Λ , le facteur d'échelle indiquant la limite au-dessus de laquelle la présence de nouvelle physique au-delà du Modèle Standard devrait apparaître afin de solutionner les divergences à haute énergie.
- f_{i0}^Z , l'approximation à basse énergie des couplages anormaux
- n , la puissance du facteur de forme

Cependant l'utilisation de tels facteurs de forme ne s'avère nécessaire que lorsque $\sqrt{s} \ll \Lambda$. Par ailleurs nous ne connaissons par la forme réelle de la dépendance en énergie, en

conséquence la collaboration CMS a décidé de ne pas appliquer de facteur de forme pour l'étude des coupages anormaux neutres.

1.2.0.9 Aspects expérimentaux

D'un point de vue expérimental, l'étude des couplages anormaux se fait dans les canaux di-bosons ZZ ou $Z\gamma$ où leur présence se traduit généralement par une augmentation de la section efficace de production. Par ailleurs ces couplages peuvent entraîner des déformations importantes dans les distributions de certaines observables physiques. L'étude de cet effet permet d'augmenter la sensibilité des analyses. Les variables discriminantes généralement utilisées sont la masse invariante du système di-bosons ou encore l'impulsion transverse d'un des bosons pour lesquels l'effet se manifeste dans la queue de distribution comme nous pouvons le voir sur la figure 1.2.

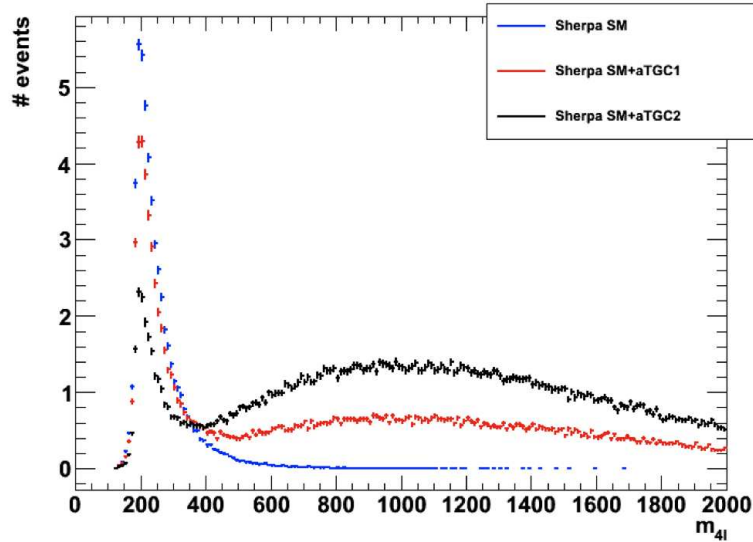


Figure 1.2 – *Effet des couplages anormaux neutres sur le spectre de masse m_{ZZ} . Modèle $aTGC1$ ($f_4^Z=0.1$; $f_5^Z=0.1$) et modèle $aTGC2$ ($f_4^Z=0.1$; $f_5^Z=0$).*

1.2.0.10 Limites expérimentales actuelles

Des mesures de l'auto-couplage entre bosons neutres et en particulier pour les bosons Z ont été effectuées précédemment dans les expériences au LEP et au Tevatron. Les mesures effectuées sont en accord avec les prédictions du Modèle Standard. La table 1.5 présente les limites obtenues par ces expériences pour les couplages impliquant trois bosons Z . À partir des données enregistrées en 2010 et surtout en 2011, les expériences ATLAS et CMS ont recherché des couplages anormaux et publié de nouvelles limites sur les couplages trilineaires et en particulier les couplages à trois bosons Z . Ces résultats seront présentés et discutés au chapitre 6 avec les résultats obtenus dans cette thèse.

Expérience	$\mathbf{f}_4^Z$	$\mathbf{f}_5^Z$	Publication
ALEPH	$[-0.60 ; +0.61]$	$[-1.22 ; +1.10]$	[12]
DELPHI	$[-0.40 ; +0.42]$	$[-0.38 ; +0.62]$	[13]
L3	$[-1.9 ; +1.9]$	$[-5.0 ; +4.5]$	[14]
OPAL	$[-0.45 ; +0.58]$	$[-0.94 ; +0.25]$	[15]
LEP EWK WG	$[-0.30 ; +0.30]$	$[-0.34 ; +0.38]$	[16]
CDF	$[-0.12 ; +0.12]$	$[-0.13 ; +0.12]$	[17]
DØ	$[-0.28 ; +0.28]$	$[-0.31 ; +0.29]$	[18]

Table 1.5 – *Tableau comparatif des limites expérimentales sur les couplages anormaux ZZZ obtenues avant le LHC.*

Chapitre 2

Le LHC et le détecteur CMS

Contenu du chapitre

2.1	Le LHC : grand collisionneur de hadrons	32
2.2	L'expérience et le détecteur CMS	35
2.2.1	Le trajectrographe	37
2.2.1.1	Le détecteur à pixel	37
2.2.1.2	Le détecteur à micro-pistes	39
2.2.1.3	Présence de matériaux non sensibles et cracks.	40
2.2.2	Le calorimètre électromagnétique	41
2.2.2.1	Géométrie	42
2.2.2.2	Les photodétecteurs	43
2.2.2.3	Perte de transparence et correction laser	45
2.2.2.4	Résolution en énergie et position	46
2.2.3	Le calorimètre hadronique	46
2.2.4	L'aimant solénoïdal supraconducteur	47
2.2.5	Les chambres à muons ou spectrographe à muons	48
2.2.6	Le système de déclenchement	50
2.2.6.1	Le déclenchement de premier niveau	51
2.2.6.2	Le déclenchement de haut niveau	53
2.3	Grille de calcul et outils informatiques CMS	54
2.3.1	La grille de calcul	54
2.3.2	Les outils CMS	55
2.3.2.1	CRAB :Cms Remote Analysis Builder	55
2.3.2.2	ROOT et RooFit	56

2.1 Le LHC : grand collisionneur de hadrons

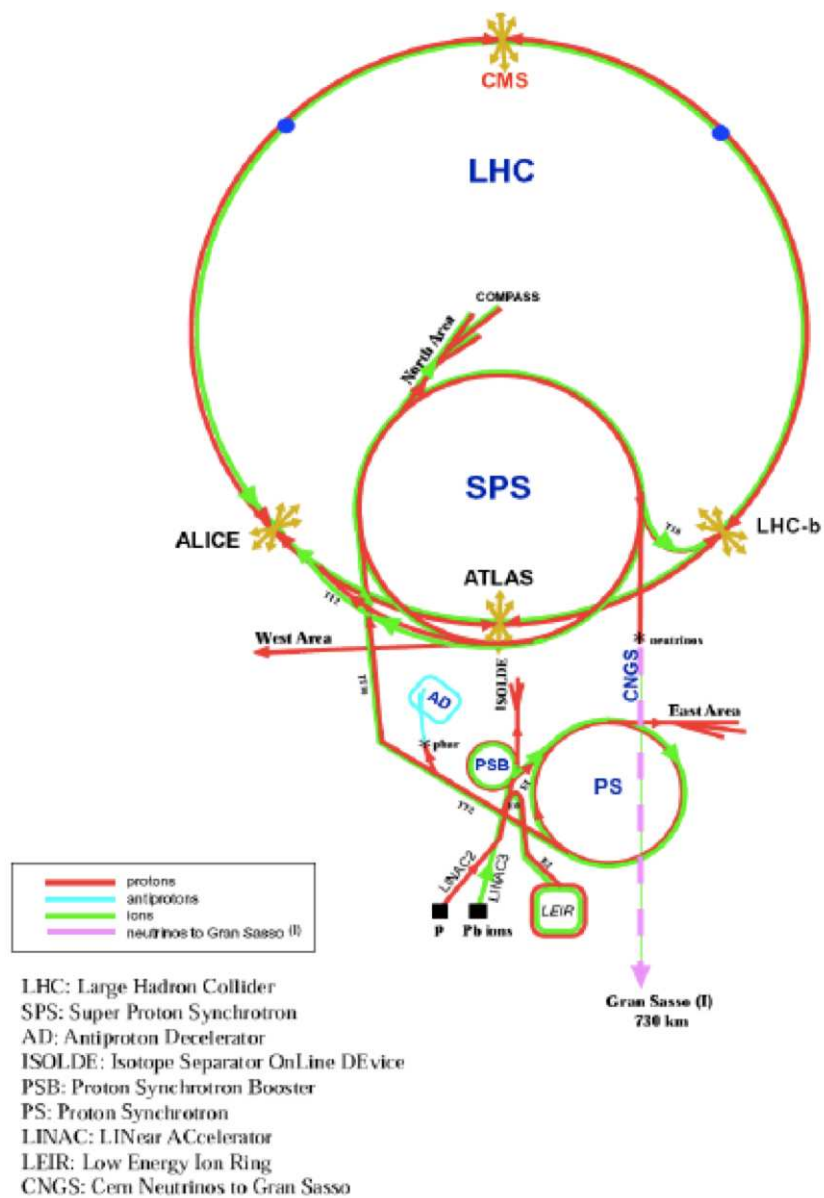


Figure 2.1 – *Les différents accélérateurs du CERN.*

Le Large Hadron Collider (cf. figure 2.1) est le plus puissant accélérateur de particules jamais construit. Un tel instrument est réalisé afin d'étudier des phénomènes rares et encore non découverts, dans le domaine de la physique des particules et nécessitant des énergies de collision élevées. Le LHC est plus précisément un collisionneur protons-protons où chaque faisceau atteindrait dans son mode nominal de fonctionnement, l'énergie de 7 TeV ce qui nous donnerait alors une énergie de collision de 14 TeV. Le LHC présente un deuxième mode : celui d'accélérateur de faisceaux d'ions lourds, plus précisément d'ions plomb, présentant un grand nombre de protons d'où une énergie de collision atteignant les 1150 TeV.

Au sein de ce collisionneur sont installés plusieurs détecteurs correspondant chacun à une expérience particulière :

- ALICE (A Large Ion Collider Experiment) a pour sujet d'étude le plasma de quarks et de gluons, état supposé de la matière quelques millisecondes après le Big Bang, naissance de l'univers. Cette expérience œuvrera avec un LHC en mode ions lourds dont les collisions occasionneront une quantité incroyable de particules, beaucoup plus importante que dans le cas de collisions protons-protons.
- ATLAS (A Toroidal LHC ApparatuS) tire son nom notamment de l'électroaimant de forme toroïdal permettant de refermer le champ magnétique sur lui-même sans utiliser de culasses de retour de champ comme c'est le cas dans l'expérience CMS (cf. partie 2.2.5). Cette expérience a un but de prospection généraliste avec un accent particulier sur la recherche du boson de Higgs ou encore la nouvelle physique, c'est-à-dire les extensions du Modèle Standard, telles que la supersymétrie ou les dimensions supplémentaires.
- CMS (Compact Muon Solenoid) est une expérience similaire à ATLAS quant à ses objectifs : découvrir le boson de Higgs et révéler de nouvelles particules issues des théories supersymétriques ou encore des dimensions supplémentaires. Son nom vient de :
 - la compacité de sa géométrie induite par l'aimant solénoïdal parmi les plus puissants du monde
 - l'accent particulier mis sur la détection des muons
- LHCb (Large Hadron Collider beauty) est une expérience sur le quark beau (b) et de manière générale, la physique de la beauté au sens où elle implique un quark b. Elle étudie notamment les désintégrations des mésons B et la violation CP.

D'un point de vue historique, les premières propositions d'un collisionneur de hadrons datent de 1992, avec la mise en place de collaborations et la réalisation de prototypes, mais ce n'est qu'en 1994 que le conseil du CERN adopte le projet de construction du LHC dans l'optique de remplacer le collisionneur électrons-positrons : le LEP (Large Electron Positron collider), qui devra être démantelé afin de céder les 27 km de tunnel qu'il occupe au LHC. En 1996 les deux premières expériences ATLAS et CMS sont approuvées. En 1998 débutent les travaux d'un projet devenu mondial. En 2000 le LEP cesse définitivement son activité tandis qu'en 2001 le projet de grille européenne est lancé en vue de répondre aux énormes quantités de données qu'engendrera le LHC et ses expériences. Après construction et tests, le LHC lance son premier faisceau le 10 septembre 2008 mais est arrêté quelques jours plus tard à la suite d'un incident au niveau de connexions électriques entraînant des dégâts dans le système de refroidissement. Un arrêt afin d'isoler le problème,

le solutionner et mettre en place de nouveaux contrôles est alors réalisé. En fin d'année 2009 le LHC redémarre avec des collisions à 900 GeV dans le centre de masse. En 2010 le LHC décroche un premier record d'énergie en atteignant 7 TeV (3.5 TeV par faisceau) et accumule pour les années 2010-2011, une luminosité intégrée de 5 fb^{-1} permettant, parmi d'autres résultats, de réduire la fenêtre de recherche du boson de Higgs à 116-126 GeV/ c^2 . En 2012, l'énergie des collisions atteint le nouveau record de 8 TeV.

Abordons maintenant l'aspect technique de cet accélérateur. Les protons sont fabriqués à partir d'hydrogène puis accélérés jusqu'à 50 MeV au sein du LINAC (LInear particle ACcelerator). Ils sont ensuite injectés dans le BOOSTER pour atteindre 1.4 GeV puis transmis successivement à deux synchrotrons à protons : le Proton Synchrotron (PS) et le Super Proton Synchrotron (SPS) où ils seront accélérés à 25 et 450 GeV respectivement. Notons au passage que ces diverses structures étaient déjà existantes du temps du LEP. Enfin ils sont dirigés vers le LHC où ils seront accélérés jusqu'aux 7 TeV dans le cas nominal, 4 TeV pour l'année 2012. (cf. figure 2.1)

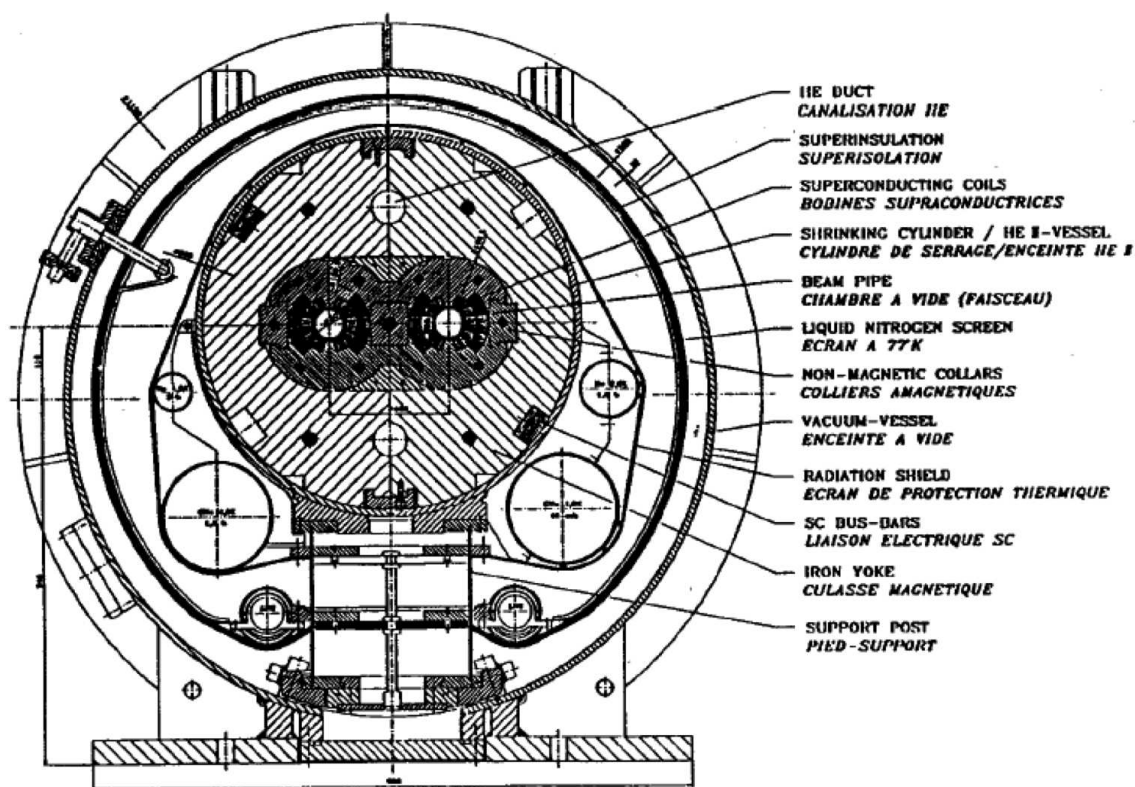


Figure 2.2 – *Vue en coupe d'un dipôle du LHC.*

Le LHC proprement dit est formé de huit arcs contenant des aimants dipolaires. Au total 1232 dipôles (cf. figure 2.2) assurent la courbure des faisceaux tandis que 392 quadrupôles (assemblage de quatre pôles magnétiques de signes alternés) servent à les focaliser. Par ailleurs ces aimants dipolaires, sous forme de câbles en alliage supraconducteur Niobium-Titane, fournissent un champ magnétique de 8.33 T, dans le cas nominal, et

ce, grâce à un système de refroidissement à base d'hélium superfluide sous pression pour une température de 1.9 K, soit -271 °C. Chaque structure dipolaire mesure 15 mètres de long pour un poids de 35 tonnes. L'accélération est quant à elle réalisée par huit cavités radiofréquence par faisceau œuvrant à 400.8 MHz et délivrant chacune une tension de 2 MV, soit un champ accélérateur de $5 \text{ MV} \cdot \text{m}^{-1}$.

Quelques caractéristiques supplémentaires du LHC et d'autres déjà introduites sont regroupées dans le tableau suivant :

Paramètres	Valeur nominale	2010	2011
Circonférence [m]	26659	-	-
# cavités radiofréquence [par faisceau]	8	-	-
# aimants	9593	-	-
# quadripôles principaux	392	-	-
# dipôles principaux	1232	-	-
T[°C] des dipôles	-271.3	-	-
Champ magn. dipolaire max. [T]	8.33	-	-
Energie par faisceau	7 TeV	900 GeV - 3.5 TeV	3.5 TeV
Luminosité instantanée [$\text{cm}^{-2}\text{s}^{-1}$]	10^{34}	$2 \cdot 10^{32}$	$3.6 \cdot 10^{33}$
Luminosité intégrée sur un an	100 fb^{-1}	36 pb^{-1}	4.98 fb^{-1}
# paquets : k_B	2808	368	1380
# protons par paquet : N [10^{11}]	1.15	1.2	1.5
Durée entre 2 paquets [ns]	25	150	50
β^* [m]	0.55	3.5	1

Table 2.1 – *Caractéristiques du LHC.*

2.2 L'expérience et le détecteur CMS

CMS (Compact Muon Solenoid) est l'une des quatre grandes expériences menées au sein du nouveau grand collisionneur de hadrons du CERN. Plus concrètement, CMS est une expérience polyvalente au sens où elle explore plusieurs domaines de la physique. Principalement élaboré pour vérifier l'existence du boson de Higgs, le détecteur permet également la recherche de nouvelles dimensions ou encore celle des particules supersymétriques. D'un point de vue aventure humaine, CMS est le fruit d'une collaboration entre plus de 3000 ingénieurs et scientifiques, dont environ 400 étudiants, issus de 183 laboratoires et universités et provenant de 38 pays différents au travers du monde (Europe, Amérique, Asie...).

Nous allons maintenant nous intéresser plus particulièrement à l'aspect technique et fonctionnel de ce détecteur. Le détecteur CMS en résumé et en quelques chiffres c'est d'après [19] :

- Environ 21 mètres de long pour 15 mètres de large et 15 mètres de haut.
- Un poids de 12 500 tonnes
- Configuration résumée : tonneau et bouchons.
- Situation géographique : Cessy, France, 100 mètres sous terre.

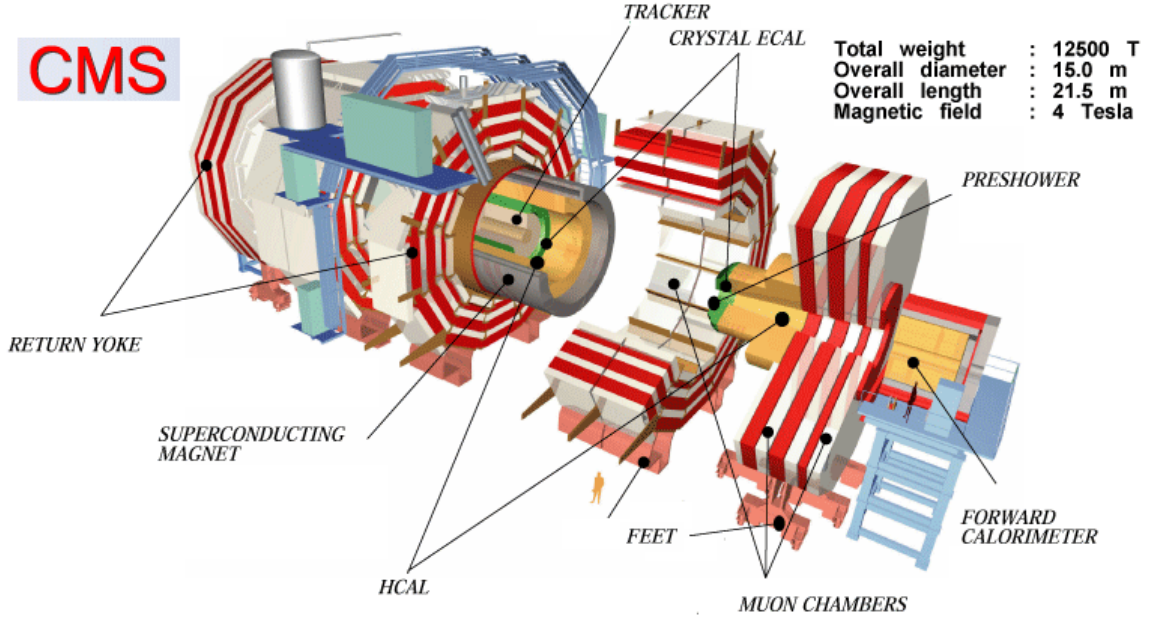


Figure 2.3 – Vue éclatée du détecteur CMS.

Le détecteur se décompose en plusieurs sous-détecteurs (figure 2.3) ayant des fonctions bien spécifiques. Ces derniers sont répartis sous différentes géométries, les deux principales étant le tonneau et les bouchons. Ils sont en partant du plus interne, c'est-à-dire du plus proche du faisceau :

- Le trajectographe
- Le calorimètre électromagnétique
- Le calorimètre hadronique
- Les chambres à muons

Entre le calorimètre hadronique et les chambres à muons se situe l'aimant solénoïdal supraconducteur délivrant un champ magnétique de 3.8 T.

Conventions et système de coordonnées.

Le détecteur CMS est de forme cylindrique allongée selon l'axe z . L'origine du système de coordonnées est le point de collision. L'axe x est l'axe horizontal pointant vers le centre du LHC tandis que l'axe y est vertical pointant vers le haut. L'angle azimutal ϕ mesuré à partir de l'axe x est défini dans le plan transverse (x, y) et l'angle polaire θ est mesuré à partir de l'axe z . De manière générale, le système de coordonnées utilisé pour déterminer les positions ou directions des particules reconstruites, points de mesure, dépôts d'énergie... est le système (η, ϕ) avec η , la pseudo-rapacité définie par :

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (2.1)$$

La pseudo-rapacité est une approximation de la rapidité y

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (2.2)$$

approximation qui n'est valide que dans le cadre de la limite relativiste $\frac{mc^2}{E} \rightarrow 0$. La coordonnée η est préférentiellement utilisée car elle maintient invariante par translation selon l'axe z , une distribution différentielle selon η .

2.2.1 Le trajectographe

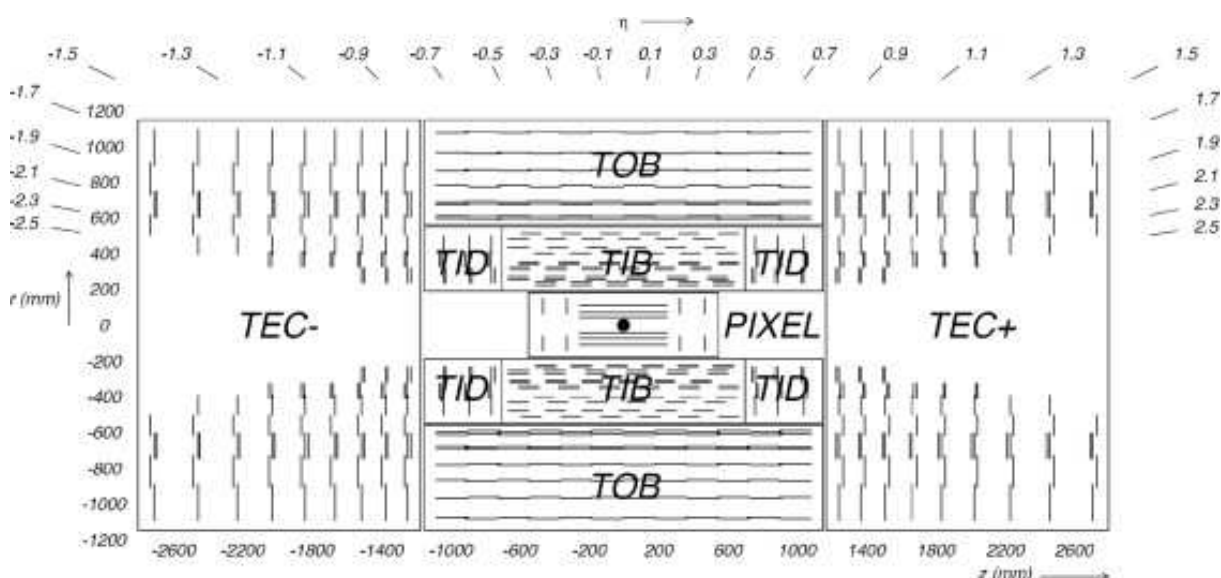


Figure 2.4 – Structure du trajectographe.

Le trajectographe (figure 2.4) ou trajectomètre ou encore détecteur de traces chargées est un composant cylindrique de 5,5 mètres de long pour un rayon de 1,1 mètre. Il permet de reconstruire les traces chargées à partir des points d'impact ("hits") mesurés dans chaque couche du détecteur, d'identifier le vertex d'interaction et les vertex secondaires. L'acceptance du trajectographe, c'est-à-dire sa partie utile est définie en termes de pseudorapidité : $|\eta| < 2,5$.

Le trajectographe est lui-même constitué de deux sous-détecteurs au silicium :

- Le détecteur à pixels, qui forme la partie la plus proche du faisceau de protons.
- Le détecteur à micro-pistes.

2.2.1.1 Le détecteur à pixel

Le détecteur à pixels, formant la partie la plus interne (rayon transverse $r_T < 15$ cm), est conçu de trois couches cylindriques (dont la première ne se situe qu'à 4,4 cm du faisceau, les deux autres à 7.3 et 10.2 cm respectivement) dans sa partie centrale de 53 cm de long. Deux bouchons sont disposés de part et d'autre et sont formés de deux disques en $z = \pm 34.5$ et $z = \pm 46.5$ cm constituant une couronne de rayon compris entre 6 et 15 cm. Il s'agit d'un détecteur de très fine granularité avec 48 millions de pixels pour une surface de 0.78 m^2 dans sa partie tonneau contre 18 millions de pixels pour une surface de 0.28 m^2 dans les parties bouchons. Il fournit des germes de traces, encore qualifiés de graines

ou de pré-traces dans les algorithmes de reconstruction, et en général dotés de deux ou trois points d'impact avec une résolution de $10\ \mu\text{m}$ en (r,φ) et de $15\ \mu\text{m}$ en z .

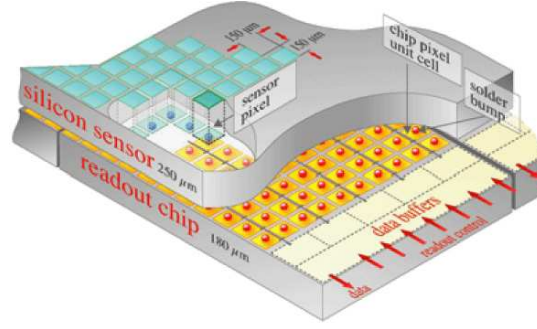


Figure 2.5 – *Détail d'un segment de capteur du trajectographe.*

Concernant la structure des couches, chacune est découpée en segments de capteurs (modules) à pixel en silicium de $150\ \mu\text{m}$ par $150\ \mu\text{m}$ pour $250\ \mu\text{m}$ d'épaisseur (cf. figure 2.5). Sous ces capteurs sont disposées des puces de lecture qui vont récolter l'information délivrée par chaque pixel sous forme de signaux électriques de $50\ \mu\text{W}$ ce qui crée beaucoup de chaleur à l'échelle de l'ensemble du détecteur. De ce fait un système de refroidissement est également intégré au support mécanique des puces. La très grande résolution du détecteur cumulée au fait qu'il se situe très près de la zone d'interaction des faisceaux est indispensable pour reconstruire les vertex de particules ayant un temps de vie long.

Cependant la très grande proximité des points de collisions implique un énorme taux de particules. A titre d'illustration, la couche la plus interne qui se situe à $4\ \text{cm}$ environ du point de collisions protons-protons sera exposée, lorsque le LHC aura atteint sa luminosité maximale, à une fluence de $3 \cdot 10^{14}\ \text{n}_{eq}\ \text{cm}^{-2}\ \text{an}^{-1}$, avec n_{eq} désignant le nombre équivalent de neutrons d'énergie cinétique d'un MeV. Pour les couches suivantes la fluence diminue à $1.2 \cdot 10^{14}\ \text{n}_{eq}\ \text{cm}^{-2}\ \text{an}^{-1}$ et $0.6 \cdot 10^{14}\ \text{n}_{eq}\ \text{cm}^{-2}\ \text{an}^{-1}$. Plus explicitement, nous aurons aux alentours de la deuxième couche près de 10 millions de particules par cm^2 et par seconde. Ceci agit sur le temps de vie des composants : ainsi la première couche devra être changée tous les deux ans tandis que la troisième dépassera les dix années. Par ailleurs, la proximité entre première couche et faisceau implique un très grand contrôle quant à la localisation du faisceau : si ce dernier entre en contact direct avec le détecteur, il le détruira.

2.2.1.2 Le détecteur à micro-pistes

Ensemble comprenant les autres sous-éléments du trajectographe pour un rayon allant de 20 cm à 116 cm :

- TIB (Tracker Inner Barrel) partie tonneau interne du trajectographe ;
- TID (Tracker Inner Disk) disques venant compléter le TIB ;
- TOB (Tracker Outer Barrel) partie tonneau externe du trajectographe ;
- TEC (Tracker End Caps) disques dans les parties bouchons du trajectographe

Le TIB est constitué de 4 couches à 25.5, 33.9, 41.8 et 49.8 cm de l'axe du faisceau. Les deux premières sont double-faces tandis que les deux autres sont simple face. Plus court (longueur selon l'axe des collisions = 140 cm : $-70 < z < +70$ cm, l'origine étant le point d'interaction protons-protons) que la partie externe TOB qui l'entoure, cet élément est complété par les TID, chacun composé de 3 petits disques en $|z|=70, 80$ et 90 cm, de part et d'autre. Le TOB est quant à lui constitué de 6 couches (rayon = 60.8, 69.2, 78, 86.8, 96.5, 108 cm), les deux plus internes étant double-faces, pour un rayon compris entre 55.5 et 116 cm. Les cas "double-faces" présentent des parties actives, autrement dit des capteurs, dos à dos sur les côtés interne et externe des supports mécaniques. Ces-derniers sont désignés dans la littérature sous le nom de "rod" et constituent en fait un assemblage de tiges permettant le maintien de 6 modules, ou 12 dans les cas "double-faces", tout en autorisant le passage des inter-connexions et des sorties de lecture ainsi que du circuit de refroidissement. En incluant le câblage à ses extrémités, cette partie du trajectographe mesure 236 cm.

Concernant les TEC, parties bouchons du trajectographe, leurs disques ont un rayon compris entre 22 et 113.5 cm et sont disposés selon l'axe z entre ± 124 et ± 280 cm. Chaque TEC (TEC- situé en $-z$, TEC+ en $+z$) est formé de neuf disques servant de soutien aux modules détecteurs ainsi que de deux disques formant leurs faces avant et arrière. Ces supports sont en composite à fibre de carbone et structurés en nids d'abeille. Ce matériau et cette forme particulière permettent la dissipation de chaleur ainsi que la possibilité d'insérer un circuit de refroidissement sous forme de canules en titane de diamètre inférieur à 4 mm. Afin de faciliter l'accès aux différents modules, ces derniers ont été montés en pétales sur les disques et arrangés en anneaux autour de l'axe z . Seize pétales sont fixés sur chacun des neuf disques d'un TEC tandis que huit sont montés sur les disques faces avant et arrière de chaque TEC. Cette formation en pétale en autorise le retrait d'un ou plusieurs sans avoir à décâbler ou désassembler la structure entière.

D'un point de vue éléments actifs, le détecteur à micro-pistes contient 15148 modules à haute sensibilité comprenant près de dix millions de pistes lues par 80000 puces. Chaque module est constitué d'un jeu de capteurs maintenu par un support mécanique et de l'électronique servant à la lecture des sorties. Plus concrètement un module possède soit un capteur fin ($320 \mu\text{m}$) (TIB, TID et anneaux 1 à 4 des TEC) soit deux capteurs épais ($500 \mu\text{m}$) (TOB, anneaux 5 à 7 des TEC) ce qui nous donne, suivant la géométrie et le nombre de capteurs, un total de 24244 capteurs pour une surface active comprise entre 62.4 cm^2 (cas du 1er anneau du TEC) et 172 cm^2 (cas du module TOB). Le support mécanique, en graphite ou fibre de carbone suivant les cas, a été étudié pour fournir une structure stable, la possibilité de dissiper la chaleur tout en permettant le passage de l'électronique nécessaire pour lire les sorties des capteurs.

2.2.1.3 Présence de matériaux non sensibles et cracks.

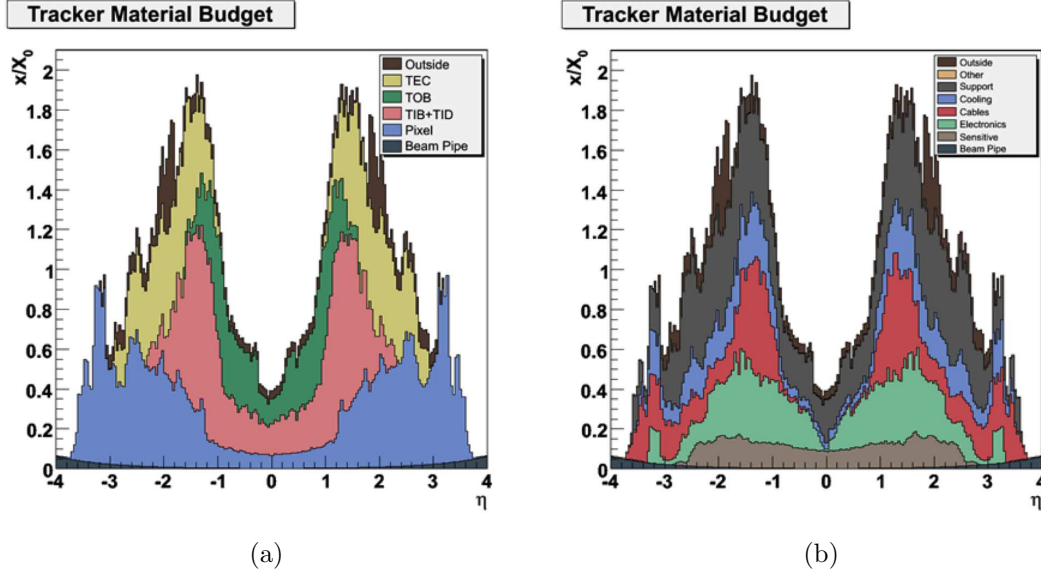


Figure 2.6 – Épaisseur de matière dans le trajectographe suivant η pour différents sous détecteurs (a) et différentes fonctionnalités (b) du trajectographe.

Le trajectographe est donc un ensemble de modules disposés selon des géométries particulières. La partie active du détecteur se limite donc à ces modules qui, d'un point de vue couverture spatiale, ne représentent que des éléments discrets. Ces divers capteurs sont de plus reliés ou connectés à un important dispositif de câbles électriques, de puces électroniques, de structures de maintien mécanique, de circuits de refroidissement... constituant une épaisseur de matériaux non directement utiles pour la détection car non sensibles.

Non seulement ces matériaux sont non sensibles, mais ils perturbent les mesures car les particules que nous cherchons à détecter interagissent avec. Les répercussions sont présentes aussi bien au niveau de la reconstruction des traces dans le trajectographe qu'au niveau de la mesure de l'énergie dans les calorimètres.

Ainsi l'épaisseur de ces matériaux affectant la mesure des électrons et photons est illustrée par la figure 2.6 où nous pouvons notamment voir leur influence en termes d'unité de radiation (X_0 , désigne la longueur moyenne au bout de laquelle un électron voit son énergie réduite d'un facteur $1/e$ sous forme de rayonnement de freinage, encore appelé effet de bremsstrahlung). Ceci est relativement important dans les zones aux environs de $\eta = \pm 1.5$.

2.2.2 Le calorimètre électromagnétique

Le calorimètre électromagnétique : "ECAL" (Electromagnetic CALorimeter, figure 2.7) permet une mesure précise de l'énergie des photons et des électrons. Il est constitué de cristaux de tungstate de plomb (PbWO_4).

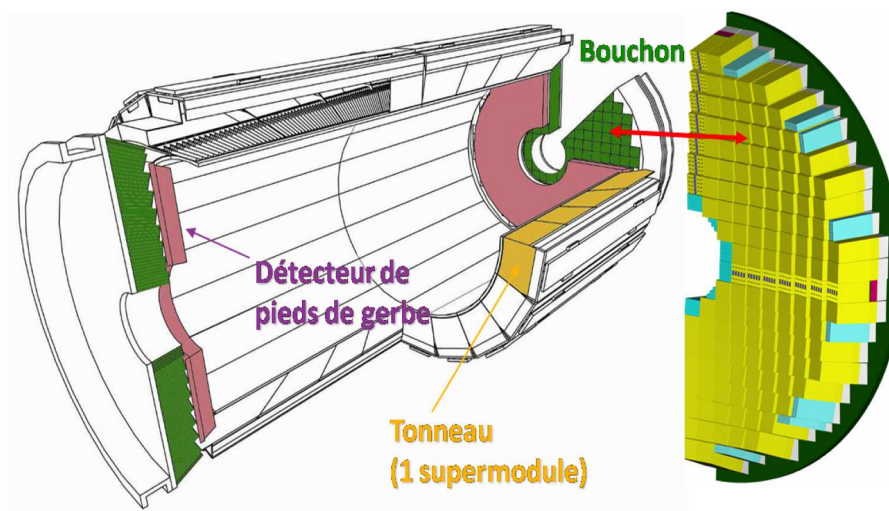


Figure 2.7 – Vue éclatée du calorimètre électromagnétique.

Le tungstate de plomb est un scintillateur qui présente quatre critères de choix pour réaliser un calorimètre électromagnétique :

- Sa densité élevée (8.28 g.cm^{-3}).
- Sa faible longueur de radiation ($X_0 = 0.89 \text{ cm}$) assure la compacité de l'ECAL.
- Son faible rayon de Molière ($R_M = 2.19 \text{ cm}$) contribue à la fine granularité du détecteur et donc à sa précision. En effet ce rayon caractérise l'extension transverse des gerbes électromagnétiques : l'interaction entre e/γ très énergétiques et le PbWO_4 provoque une production importante de particules secondaires sous forme de gerbes. Il est à noter qu'en moyenne 90% de l'énergie de la gerbe est contenu dans un cylindre de rayon R_M . Par conséquent le tungstate de plomb permet de contenir l'essentiel du développement d'une gerbe au sein d'un seul cristal, dans le cas idéal d'une particule impactant le centre du cristal.
- Le tungstate de plomb est un scintillateur très rapide à même de répondre au fort flux de particules. En effet 85% de la lumière est émise en 20 ns.

Par ailleurs le spectre d'émission de ce matériau présente deux bandes à 440 nm et 530 nm, fixant le domaine d'efficacité des photodétecteurs, éléments discutés un peu plus loin dans cette partie. Le tungstate de plomb a cependant l'inconvénient d'être peu lumineux avec un gain dépendant de surcroît de la température : le taux de lumière est compris, à température ambiante, entre 60 et 100 photons/MeV (contre 9000 pour le germanate de bismuth, scintillateur utilisé dans le calorimètre électromagnétique du LEP), défauts qui

devront être compensés par le choix de photodétecteurs adéquats.

Le design de ce détecteur remplit les caractéristiques essentielles requises par le LHC avec un temps de réponse court pouvant faire face à la très haute fréquence de collisions et au très fort flux de particules en résultant, une résistance aux radiations assurant la longévité et la fiabilité du détecteur, une fine granularité et une excellente résolution due au choix d'une colorimétrie homogène. Cependant les cristaux sont soumis, durant les collisions, à une perte de transparence qu'il faut corriger comme nous le verrons plus loin.

2.2.2.1 Géométrie

Examinons la géométrie du calorimètre (cf. figure 2.8). Comme la plupart des autres sous-détecteurs de CMS, l'ECAL possède la structure tonneau-bouchons :

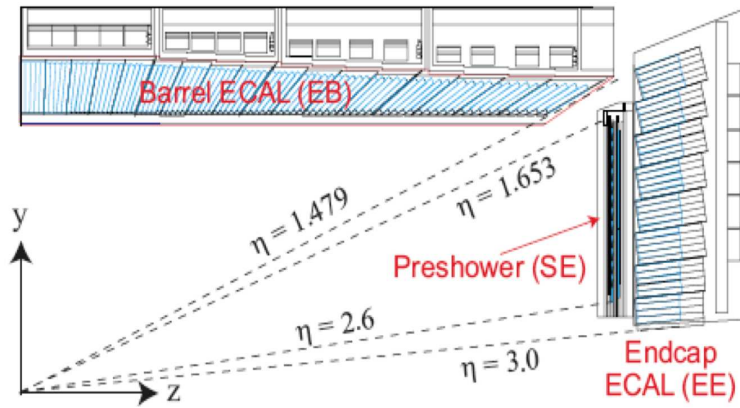


Figure 2.8 – Coupe transverse du ECAL, vue 1/4.

- La partie centrale ou tonneau couvre le domaine en $|\eta| < 1.479$. 61200 cristaux trapézoïdaux de largeur avoisinant un R_M et de longueur représentant environ $25.8 X_0$ (23 cm) sont regroupés au sein d'alvéoles en fibre de verre constituant des matrices 5×2 (les sous-modules) elles-mêmes regroupées par 40 ou 50 pour former un module. Au final ce sont 36 supermodules, association de 4 modules maintenus par une structure en aluminium, qui constituent la partie tonneau de l'ECAL pour un volume de plus de 8 m^3 et un poids de 67.4 tonnes. Notons également que la face avant des cristaux se situe à 1.29 m de l'axe du faisceau.
- Les bouchons ferment la partie tonneau et couvrent le domaine en $1.479 < |\eta| < 3.0$. Chaque bouchon est partagé en deux moitiés, qualifiées de "D" (dees) : demi-disques de 3662 cristaux regroupés en matrice 5×5 . Les cristaux ne possèdent pas les mêmes dimensions que dans la partie centrale avec cette fois-ci une largeur de $1.3 R_M$ pour une longueur de $24.7 X_0$ (22 cm). Le volume d'un bouchon est de 2.9 m^3 pour un poids de 24 tonnes.
- De plus un détecteur de pieds de gerbes, encore nommé initiateur de gerbes, est positionné devant les cristaux des bouchons pour le domaine en $1.653 < |\eta| < 2.6$. Le détecteur se décompose sous forme d'un radiateur en plomb initiant la

gerbe électromagnétique et d'une couche de capteurs en silicium chargée de mesurer l'énergie déposée.

Ainsi lorsqu'un photon ou un électron parvient au niveau du calorimètre, il y dépose toute son énergie sous forme d'une gerbe électromagnétique qui sera transformée en lumière dans les scintillateurs (cristaux), que des photodétecteurs, placés à l'arrière de ces derniers, convertiront en signaux électriques qu'ils amplifieront ensuite afin de les rendre exploitables.

Il est à noter que le calorimètre électromagnétique présente, comme le trajectographe, des zones "difficiles" pour la reconstruction des électrons. Ces zones sont notamment dues à l'absence de matériaux actifs en $\eta=0$ et $\eta=\pm 1.5$. En $\eta=0$, il s'agit de la jonction entre deux supermodules tandis qu'en $\eta=\pm 1.5$, c'est la zone de recouvrement entre partie tonneau et parties bouchons qui induit ces difficultés.

2.2.2.2 Les photodétecteurs

Je vais maintenant détailler le principe des photodétecteurs avec l'examen particulier de ceux utilisés dans le calorimètre électromagnétique et introduits précédemment. La structure caractéristique de ces derniers peut se résumer, au sein d'un tube sous vide, à :

- une cathode photosensible
- un dispositif de collecte d'électrons
- un dispositif de multiplication des électrons (succession de dynodes par exemple)
- une anode donnant en sortie un signal électrique

Ainsi, lorsqu'un photon issu du scintillateur parvient à la cathode, il y est converti par effet photoélectrique en un électron. Ce dernier est, sous l'effet d'un champ électrique, accéléré vers une première dynode du dispositif photomultiplicateur, où par choc il va donner naissance à des électrons secondaires qui vont être, à leur tour, accélérés jusqu'à la dynode suivante... Dans certains autres systèmes, c'est un processus d'avalanche qui va multiplier la quantité d'électrons. Nous aurons alors une cascade d'électrons que la cathode va collecter et restituer sous forme d'un signal électrique.

Les pré-requis de ces détecteurs au sein de CMS sont : un faible temps de réponse, une résistance aux radiations afin de réduire l'importance du courant de fuite dû aux défauts introduits par les neutrons, un gain conséquent pour pallier un scintillateur PbWO_4 peu lumineux et le devoir de fonctionner sous un champ magnétique important. Ainsi dans la partie tonneau ce sont des photodiodes à avalanches (APD, Avalanche PhotoDiode) qui sont utilisées tandis que dans les parties bouchons ce sont des phototriodes à vide (VPT, Vacuum PhotoTriode), mieux à même de résister aux radiations beaucoup plus importantes dans ces parties du détecteur.

La photodiode à avalanche (APD)

Les photodiodes à avalanche utilisées dans CMS, dont le principe de fonctionnement

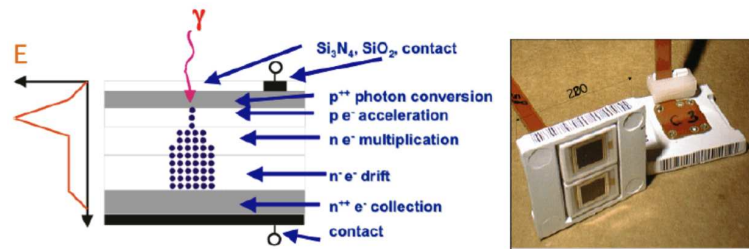


Figure 2.9 – Photodiode à avalanche (APD), schéma et photo.

est illustré par la figure 2.9, présentent un gain égal à 50, une efficacité quantique de 80% pour des longueurs d'onde du domaine visible et une sensibilité à la température de -2.4% par °C. Elles sont disposées par deux à l'arrière de chaque cristal, chacune représentant une surface active de $5 \times 5 \text{ mm}^2$. Il est à noter qu'afin d'éviter un dépôt direct d'énergie au sein des APD, il faut en réduire l'épaisseur ce qui a pour inconvénient d'en augmenter la capacité. Un compromis entre les deux critères a été trouvé en réalisant des APD d'une épaisseur de $300 \text{ }\mu\text{m}$.

La phototriode à vide (VPT)

Le photoélectron généré au niveau de la cathode, passe au travers de l'anode grillagée

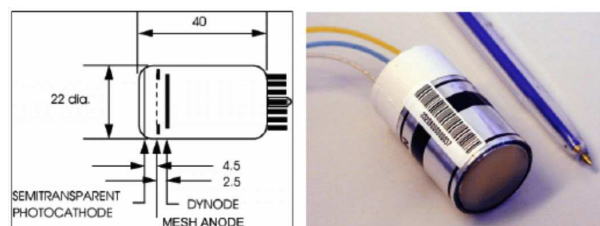


Figure 2.10 – Phototriode à vide (VPT), schéma et photo.

et atteint la dynode au niveau de laquelle sont créés les électrons secondaires. Ces derniers sont attirés par l'anode qui va alors les collecter et les restituer en signal électrique exploitable. La VPT (figure 2.10) présente un gain de 10, contre 50 pour les APD, mais est moins sensible à la température avec une perte en efficacité inférieure à 1% par °C. Son efficacité quantique est de 15-20% pour les pics d'émission du PbWO_4 , faiblesse compensée par environ 280 mm^2 de surface active.

Elles sont, par ailleurs, à même de résister aux fortes radiations et à l'important champ magnétique qui sont présents dans les parties bouchons du détecteur.

2.2.2.3 Perte de transparence et correction laser

Les cristaux de tungstate de plomb formant le ECAL sont soumis à de très fortes radiations faisant apparaître au sein du scintillateur des centres de couleurs pouvant absorber une fraction de la lumière qui y est créée. Ainsi les cristaux perdent en transparence au fur et à mesure des collisions. Cet effet est cependant limité, entraînant une réduction de quelques pour cents de la transmission optique au cours d'un *run*. De plus le cristal tend à retrouver naturellement sa transparence par phénomène de recuit : soumis aux fortes radiations, le cristal se trouve à une température légèrement supérieure à la température de recristallisation puis refroidit lors de l'arrêt des collisions ; ce processus change alors certaine propriété du matériaux, sa transparence dans notre cas.

Ainsi la transparence des cristaux du ECAL suit un schéma cyclique avec perte de transparence lors des collisions et recouvrement de sa transparence lors des phases d'arrêt, comme nous pouvons le voir sur la figure 2.11. Cependant cet effet n'est pas uniforme : il est plus important pour des hautes luminosités et dans les parties bouchons. Ainsi les pertes sont de l'ordre de 1 à 2% dans le tonneau et à basse luminosité contre quelques dizaines de pour cents pour les région de hautes pseudo-rapidité et à haute luminosité.

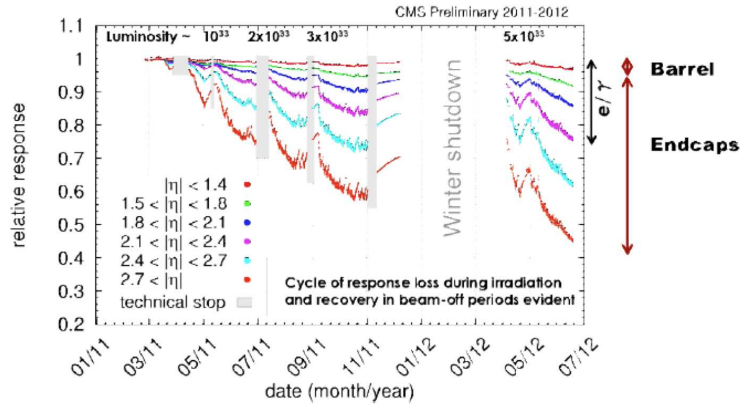


Figure 2.11 – Cycle de perte de transparence et récupération en fonction du temps et de région en η .

Afin de corriger ce phénomène et d'obtenir une estimation correcte de l'énergie déposée dans le ECAL, un système de lasers injecte par fibre optique des impulsions lumineuses d'amplitude connue dans les cristaux. Le rapport entre la réponse de ces derniers et l'amplitude initiale des impulsions laser évalue la transparence. Ce système intervient au moment des intervalles creux lors du cycle des faisceaux. Les performances d'une telle correction sont présentées dans la figure 2.12.

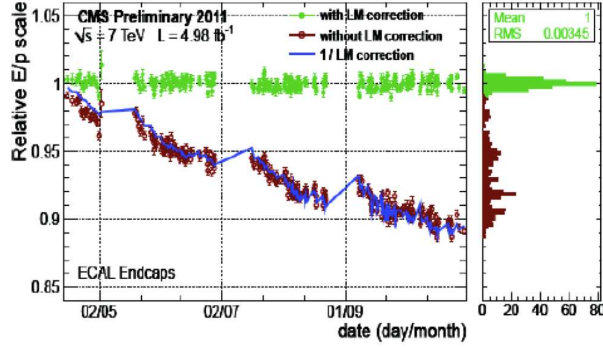


Figure 2.12 – *Effet de la correction laser.*

2.2.2.4 Résolution en énergie et position

La résolution en énergie du calorimètre s'exprime par la formule :

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.3)$$

Avec :

- E, l'énergie de la particule laissant un dépôt dans le ECAL
- S, terme stochastique traduisant les fluctuations énergétiques de la gerbe
- N, contribution due au bruit électronique
- C, constante

La résolution a été mesurée à l'aide de tests basés sur des faisceaux d'électrons :

$$S = 2.8\% \quad N = 0.12 GeV \quad C = 0.3\% \quad (2.4)$$

La résolution en position du ECAL est liée à sa résolution en énergie. En effet les fluctuations des mesures des dépôts d'énergie affectent la position de ce dernier. Cette résolution en position a été étudiée par simulation, vérifiée par test faisceau et est d'environ 10^{-3} unité en η et 1.6 mrad en ϕ , correspondant à ± 2 mm dans le plan (x,y).

2.2.3 Le calorimètre hadronique

Le calorimètre hadronique : "HCAL" (Hadron CALorimeter) permet de mesurer l'énergie des hadrons neutres et chargés. Il est constitué, dans sa partie tonneau, d'un empilement de plaques de laiton (70% Cu + 30% Zn, pour une densité de 8.53 g.cm^{-3}) de 5 cm d'épaisseur, présentant le double avantage de ne pas être sensible au champ magnétique et de ne pas trop perturber les muons qui vont traverser le calorimètre. Cet empilement comprend également des feuilles de scintillateurs de 4 mm d'épaisseur dont la lumière est collectée par fibre optique. Les photodétecteurs utilisés cette fois-ci sont des photodiodes hybrides. L'ensemble forme la partie tonneau de 1,8 m à 2,95 m de rayon sur une longueur de 9 m et couvre le domaine en $|\eta| < 1.48$. Il est à noter que cette partie tonneau se distingue en une partie (HB) située à l'intérieur de l'aimant supraconducteur et d'une

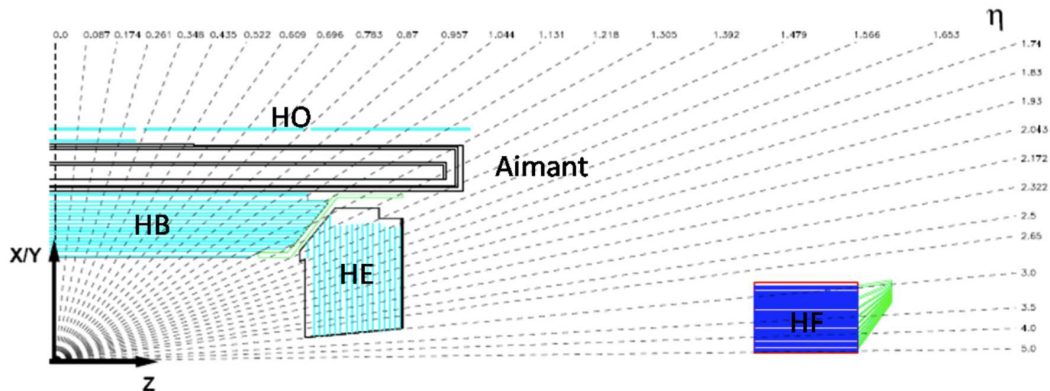


Figure 2.13 – Coupe transverse des différents éléments constituant le HCAL, vue 1/4.

partie externe (HO) entre l’aimant et les chambres à muons, la partie (HB) n’étant pas suffisamment épaisse pour contenir la totalité du développement de la gerbe hadronique. En effet (HO) se sert de l’aimant comme absorbeur et peut ainsi mesurer les gerbes tardives ou longues. Cf. figure 2.13.

Dans la partie bouchon (HE), la structure reste inchangée à l’exception de la largeur des plaques de laiton qui passe de 5 cm à 8 cm. Les bouchons couvrent un domaine en $1.48 < |\eta| < 3.0$ et ont une épaisseur de 1.8 m permettant de contenir à elle seule la totalité des gerbes hadroniques.

Afin d’augmenter le domaine d’acceptance jusqu’à $|\eta| < 5$, deux calorimètres (HF) sont placés très en avant (calorimètres à petit angle sur la figure 2.13) de part et d’autre en $z = \pm 11.1$ m du point d’interaction, dans la partie la plus externe du détecteur CMS. D’un rayon de 1.4 m pour une longueur de 1.65 m, le HF permet de compléter la reconstruction de particules très en avant ($3.0 < |\eta| < 5.0$).

2.2.4 L’aimant solénoïdal supraconducteur

L’aimant supraconducteur de CMS assure un champ magnétique de 3.8T, soit près de 100000 fois le champ magnétique terrestre moyen. Ce sont ses dimensions, sa géométrie et ses caractéristiques qui ont contraint celles des sous-détecteurs, notamment ceux qui sont situés à l’intérieur (trajectographe, calorimètres électromagnétique et hadronique) : impératif de compacité, d’où des matériaux très denses, et nécessité de fonctionnement sous un champ magnétique important impliquant certains choix particuliers, photodétecteurs par exemple, comme cela a été évoqué dans la partie 2.2.2.2. Rappelons également qu’il donne son nom au détecteur et à l’expérience : CMS = Compact Muon Solenoid. Examinons sa structure. En résumé, il s’agit d’une bobine supraconductrice isolée thermiquement au sein d’une enceinte, le cryostat, le tout entouré de culasses en acier appelées ”culasses de retour de champ”. Plus explicitement :

- Cinq bobines supraconductrices, formées de 4 couches en niobium-titane, juxtaposées et maintenues à 4.15 K par de l’hélium liquide, afin d’ôter toute résistance électrique au matériau supraconducteur et ainsi d’atteindre de très hautes intensités

sans dégagement thermique. Avec 2168 spires parcourues par un courant nominal de 19.14 kA, l'aimant peut produire un champ de 3.8T au niveau du point d'interaction. L'énergie stockée dans la bobine s'élève à 2.6 GJ. Les grandeurs sont : une longueur de 12.5 m, un diamètre de 5.9 m, une épaisseur de 31 cm et un poids de 220 tonnes. La plupart de ces chiffres sont des records et font de l'aimant de CMS, le plus grand solénoïde supraconducteur jamais construit.

- Une enceinte sous vide, le cryostat, chargée d'isoler thermiquement le bobinage supraconducteur. Sa face interne se situe à un rayon de 2.9 m du faisceau de protons, tandis que sa face externe est à 3.7 m. L'épaisseur de l'armature de cet étui est d'une dizaine de centimètres environ.
- Les culasses de retour de champ sont constituées de plaques d'acier entre lesquelles sont intercalées les chambres à muons. Leurs épaisseurs relatives sont de 30, 63 et 63 cm dans la partie tonneau contre 60, 60 et 25 cm pour les bouchons, de la plus interne vers la plus externe respectivement. Le diamètre extérieur de la dernière plaque est de 14 mètres. La partie tonneau recouvre une longueur de 13 m. Ces plaques constituent la majeure partie du poids global de CMS avec 10000 tonnes sur 12500. Cet ensemble de culasses est appelé "culasses de retour de champ" au sens où il permet de refermer le champ magnétique sur lui-même et de le maintenir uniforme.

L'aimant permet la mesure des impulsions des traces chargées en complément du trajectographe. Il effectue de plus un tri entre particules chargées, qui seront alors courbées par le champ magnétique, et particules neutres, magnétiquement insensibles et donc de trajectoire rectiligne. Le fait d'utiliser un champ magnétique fort (3.8T) permet également de distinguer les particules de forte impulsion de celles d'impulsion plus faible. En effet les particules de faible impulsion sont extrêmement courbées et restent sous forme de spirales dans le trajectographe. Ce filtrage est surtout avantageux pour les calorimètres, et particulièrement dans la région tonneau : nous aurons ainsi éliminé les énergies de ces particules qui constituent un bruit de fond dans la mesure où les particules intéressantes, notamment du point de vue du canal $ZZ \rightarrow 4l$, ont de plus grandes impulsions. En revanche leur confinement au sein du trajectographe augmente le nombre de points d'impact ; phénomène gênant pour la reconstruction des trajectoires.

2.2.5 Les chambres à muons ou spectrographe à muons

Les chambres à muons, constituant le spectrographe à muons, reconstruisent les traces des muons, seules particules à pouvoir atteindre cette partie la plus externe du détecteur CMS. L'intérêt particulier des muons, surtout dans le canal $ZZ \rightarrow 4\text{leptons}$, vient du fait que ces particules sont faciles à isoler et sont moins affectées, que les électrons notamment, par les pertes dues au rayonnement de freinage lors d'interactions avec les matériaux non sensibles du trajectographe. Ainsi la mesure de leur impulsion est plus précise, ce qui aurait pour conséquence, dans le cas idéal d'un boson de Higgs qui se désintégrerait en 4 muons, d'avoir la meilleure résolution sur la masse reconstruite du boson. La reconstruction des muons est donc au cœur de cette expérience puisqu'elle contribue à son nom : Compact Muon Solenoid.

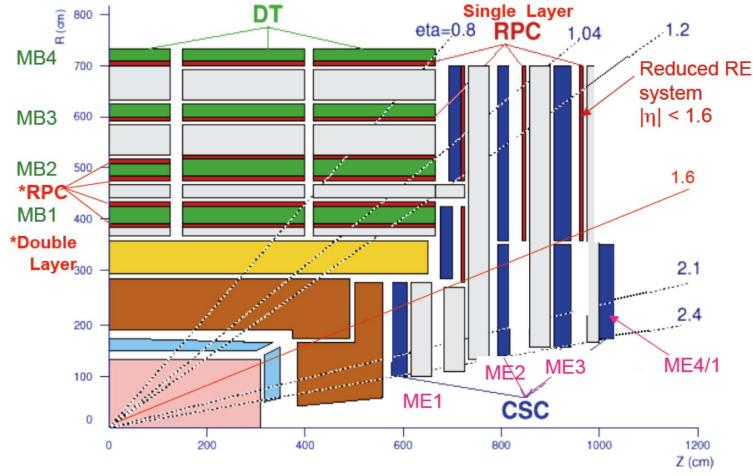


Figure 2.14 – Coupe transverse des chambres à muons, vue 1/4.

Ce détecteur (figure 2.14) comprend une partie tonneau pour le domaine $|\eta| < 1.3$ et deux bouchons couvrant le domaine $0.9 < |\eta| < 2.4$. Toutes les parties sont constituées de quatre stations à muons entre lesquelles sont intercalées les plaques d'acier des culasses de retour de champ. Chacune de ces chambres contribue à la reconstruction des traces par l'apport de segments de trace formés par quelques points d'impact. Notons également que parce que le spectrographe à muons a une surface de 25000 m² d'éléments actifs, il doit remplir les impératifs de robustesse et de fiabilité à un moindre coût.

Dans la partie centrale, où le taux de muons est faible ainsi que le champ magnétique résiduel (1.8 T), des chambres de tubes à dérive (DT= "Drift Tube") constituent la part active du détecteur. Les trois premières stations (ou couches) contiennent chacune 8 chambres, en deux jeux de quatre, fournissant les coordonnées r et φ du segment de trajectoire, tandis que quatre autres chambres donnent une mesure selon l'axe z du faisceau de protons. La quatrième station ne contient quant à elle que les huit chambres du premier type.

Dans les deux bouchons, le taux de muons ainsi que le bruit de fond dû aux neutrons sont élevés d'où l'utilisation de chambres à pistes cathodiques (CSC = "Cathod Strip Chamber"). Les pistes cathodiques fonctionnent radialement afin de fournir les coordonnées r et φ du segment de trajectoire. L'anode mesure la pseudorapacité η . Les quatre stations formant chaque bouchon sont réparties selon six niveaux de coordonnées z : 5.7, 6.6 et 7 mètres pour les trois étages de la première station, la deuxième station se situe en z égal à 7.9 mètres, et z vaut 9.1 et 10 mètres pour les troisième et quatrième stations.

Par ailleurs des chambres à plaques résistives (RPC = "Resistive Plate Chamber") sont implantées dans les deux parties, tonneau et bouchons, du spectrographe. Dans la partie tonneau, il y a six couches coaxiales pour un total de 480 RPC double-gap tandis que chaque bouchon comporte trois couches de 72 RPC de forme trapézoïdale. Elles ont pour but d'ajouter un déclenchement rapide indépendant de ceux des systèmes DT ou CSC. Les chambres à plaques résistives produisent une réponse rapide, de l'ordre de la nano-seconde, avec une très bonne résolution en temps mais une résolution en position plus grossière que celle des tubes à dérive ou des pistes cathodiques. Les RPC aident aussi à résoudre les cas ambigus de muons partageant leur trace. Six couches de plaques résistives

sont insérées dans la partie tonneau : deux couches, une au recto et l'autre au verso, pour les deux premières stations, tandis que les stations suivantes n'en possèdent qu'une sur leur face interne. Les doubles couches dans les premières stations permettent de procéder à une reconstruction pour les traces de basse impulsion transverse qui n'atteindraient pas les stations plus externes. Dans chaque bouchon, les RPC sont présents au niveau des trois premières stations.

En termes de performance, le spectrographe à muons possède une efficacité de reconstruction proche de 100% hors des cracks. De plus la résolution sur l'impulsion transverse des muons est excellente lorsque nous ajoutons, lors de la reconstruction, des informations issues du trajectographe.

Les tubes à dérive (DT = Drift Tube)

Ils sont constitués d'enceintes en aluminium, renfermant un mélange gazeux Argon-Dioxyde de carbone (85%-15%), avec en leur centre un fil d'acier inoxydable plaqué or de $50\mu\text{m}$ de diamètre : l'anode. Lorsqu'un muon traverse l'enceinte, il ionise le mélange gazeux et les charges électriques ainsi produites sont collectées par phénomène de dérive selon une durée maximale de 400 ns.

Les chambres à pistes à cathodes (CSC= Cathod Strip Chamber)

De forme trapézoïdale, elles sont composées de 7 rangées de pistes cathodiques de différentes longueurs comprises entre 3 et 16 mm et, intercalées orthogonalement aux pistes d'anodes, constituées d'un enroulement de fils de tungstène plaqués or de $50\mu\text{m}$ de diamètre. Dans les espacements entre rangées, baigne un mélange gazeux Ar-CO₂-CF₄ (40-50-10 %).

Les chambres à plaques résistives (RPC = Resitive Plate Chamber)

Elles sont formées de deux plaques parallèles en bakélite de 2 mm d'épaisseur ($\rho = 1010$ à $1011 \Omega\cdot\text{m}$), espacées de 2mm de mélange gazeux. Les deux plaques sont cuites avec un recouvrement peu résistif en graphite constituant l'anode.

2.2.6 Le système de déclenchement

Afin de faire face à l'énorme quantité de collisions, dont une proportion importante inutile du point de vue des analyses physiques, une sélection des événements intéressants doit être mise en place et ce grâce à un système de déclenchement complexe et performant à même de répondre aux immenses flux de données. En effet pour le cas du LHC fonctionnant de manière nominale, le détecteur CMS sera amené à gérer une collision proton-proton toutes les 25 ns, soit une fréquence de collision de 40 MHz correspondant à un montant d'informations d'environ $40 \text{ To}\cdot\text{s}^{-1}$ auquel ne peuvent répondre ni les fermes de calcul ni les fermes de stockage de l'expérience : le système d'acquisition ne pouvant prendre en entrée qu'environ $100 \text{ Go}\cdot\text{s}^{-1}$ et la capacité de stockage se limitant à $100 \text{ Mo}\cdot\text{s}^{-1}$. La réduction nécessaire est faite en deux étapes :

- système de déclenchement de premier niveau (ou Level 1 Trigger : LV1) réduisant le taux à un maximum de 100 kHz grâce aux informations parcellaires du détecteur

fournies par les systèmes électroniques d'acquisition.

- système de déclenchement de haut niveau (ou High Level Trigger : HLT) réduisant le taux aux 150 Hz que peuvent emmagasiner les fermes de stockage à l'aide d'algorithmes complexes comparables à ceux utilisés dans les analyses.

Une illustration de l'architecture globale du système d'acquisition pour les expériences ATLAS et CMS est fournie par la figure 2.15. Détaillons plus précisément ces deux étapes de réduction.

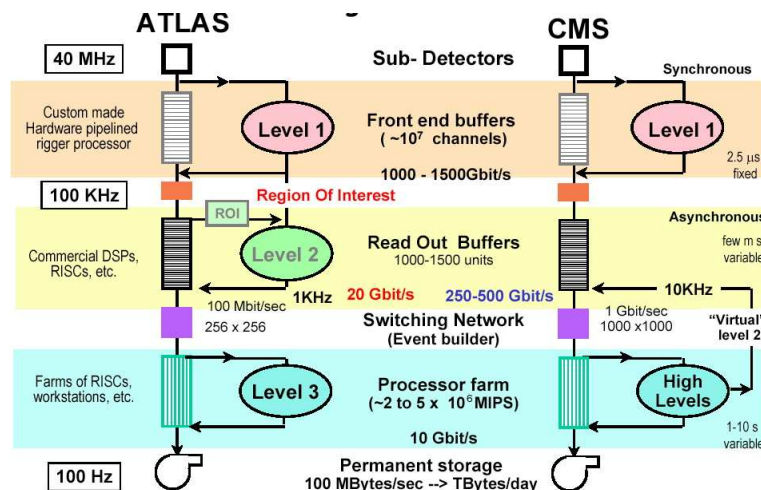


Figure 2.15 – Comparaison des architectures du système de déclenchement pour les expériences ATLAS et CMS.

2.2.6.1 Le déclenchement de premier niveau

Le système de déclenchement de premier niveau se sert d'informations issues des systèmes électroniques d'acquisition, souvent constitués de circuits logiques programmables (Field-Programmable Gate Array : FPGA par exemple) ou de circuits intégrés spécialisés tels que les ASICs, afin de choisir de garder ou non l'événement pour un traitement postérieur par le déclenchement de haut niveau. Par conséquent l'intégralité des informations détecteurs de l'événement est stockée en attendant la décision de niveau 1 dans les différents pipelines de l'électronique frontale (également dénommée "électronique front-end"). Cependant ce stockage ne peut se faire que pendant un temps réduit de 3.2 μs, comprenant le temps de traitement par les algorithmes de premier niveau et le temps de propagation des informations.

La décision de transmettre un événement pour traitement par le HLT doit se faire toutes les 25 ns, forçant une organisation des informations en pipeline pour le déclenchement de premier niveau et de ne pas utiliser de processus itératifs comme ceux pouvant être employés pour la recherche de traces. Par conséquent seuls les chambres à muons et les calorimètres électromagnétique et hadronique sont considérés pour la décision de niveau 1 comme cela est illustré dans la figure 2.16. Nous remarquons également avec cette figure, l'organisation en trois composantes du système de déclenchement de niveau 1 :

- la composante locale (Trigger Primitive Generator pour la partie calorimétrique, Local DT et CSC Trigger pour les muons) collecte les dépôts d'énergie calorimétrique

- ou des fragments de traces dans les chambres à muons, i.e. ce qui constitue les informations dites "primitives" et forme en conséquence les *primitives de déclenchement*.
- la composante régionale (Regional Calorimeter Trigger pour la partie calorimétrique, DT et CSC track finder ainsi que RPC Trigger pour la partie muons) forme les candidats L1, détaillés un peu plus loin, à partir des primitives de déclenchement.
 - la composante globale (Global Calorimeter Trigger et Global Muon Trigger) reçoit de la part de la composante régionale les quatre meilleurs candidats L1. Elle calcule également d'autres grandeurs, la "somme totale des énergies transverses" et l'"énergie transverse manquante" notamment.

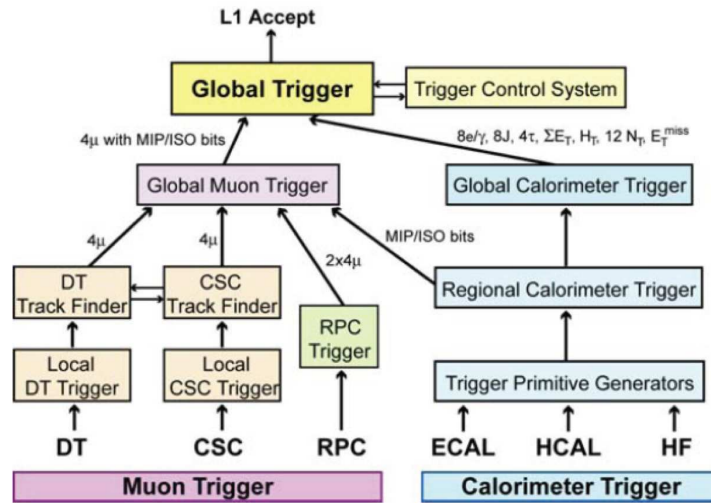


Figure 2.16 – Arbre de décision du système de déclenchement de premier niveau.

Ainsi des informations fragmentaires du détecteur sont regroupées pour former les candidats L1 suivants :

- les candidats L1 muons formés par le système de déclenchement propre aux muons, reposant sur les données issues des RPC, CSC et DT (cf. 2.2.5)
- les candidats L1 électrons/photons (isolés et non-isolés), les candidats L1 taus et les candidats L1 jets sont reconstruits par la composante régionale du système, c'est-à-dire dans le RCT (Regional Calorimeter Trigger)
- les candidats L1 "somme totale des énergies transverses", et son pendant "énergie transverse manquante" sont notamment reconstruits par la composante globale du système de déclenchement de premier niveau : le Global Calorimeter Trigger.

Notons que les meilleurs candidats L1 de chaque catégorie reconstruits ou sélectionnés par la composante globale sont ensuite transmis au déclenchement global (Global Trigger : GT) qui prend la décision finale de transmettre, ou non, l'événement au traitement de haut niveau conformément aux critères de sélection regroupées au sein du "menu de déclenchement". Ce menu regroupe 128 algorithmes implémentés dans des circuits logiques programmables qui utilisent les caractéristiques des candidats L1 (p_T , position,...) afin de déterminer les bits de décisions : ainsi un algorithme "L1.SingleIsoEG8" requiert un

candidat L1 électron/photon isolé d'énergie transverse supérieure à 8 GeV. Notons que ces bits peuvent être combinés à l'aide de OU logiques et que certains de ces algorithmes peuvent être échantillonnés pour satisfaire les 100 kHz de limite en sortie.

2.2.6.2 Le déclenchement de haut niveau

Le déclenchement de haut niveau (HLT) a pour but de réduire le taux d'événements à un niveau compatible avec les capacités de calcul et de stockage. Sachant que la taille moyenne d'un événement est de 1.5 Mo environ et que les capacités optimales de traitement permettraient d'atteindre 500 événements par seconde, le système de déclenchement réduit finalement le taux d'événements à environ 100-150 Hz afin de garantir une certaine marge.

Le HLT est organisé en une série de filtres utilisant chaque fois plus de données. Ainsi les premiers filtres dits de niveau 2 doivent être validés afin que le reste des données soit transmis aux filtres de niveau 2.5 et 3 impliquant des algorithmes plus complexes. Par ailleurs, la reconstruction des objets physiques HLT ne se fait qu'après association avec un candidat L1. Cette organisation en étage permet d'éviter d'encombrer le système, avec des données du trajectographe par exemple, si les données des calorimètres sont suffisantes pour rejeter l'événement.

Illustrons dans le cas des électrons le rôle des différents filtres :

- au niveau 2, des agrégats d'énergie sont constitués à l'aide des informations du ECAL et du détecteur de pieds de gerbe puis associés au candidats L1. La différence avec le candidat L1 est qu'à ce stade, la formation de l'agrégat est plus complexe car elle permet de récupérer l'énergie perdue par rayonnement de freinage en constituant le *super-agrégat*, plus de détails sur cette notion dans le chapitre 3. Des critères de sélection sur ce super-agrégat sont également appliqués tels que son énergie transverse, son isolation au vu de l'activité ECAL et HCAL environnante...
- au niveau 2.5 des points d'impact dans les couches les plus internes du trajectographe sont recherchés dans une fenêtre obtenue par extrapolation à partir du super-agrégat, la courbure étant approximée par mesure de l'énergie transverse de ce dernier. Deux points d'impact forment une *pré-trace*.
- au niveau 3, si nous avons obtenu des pré-traces au niveau 2.5, la reconstruction complète de la trace à partir des pré-traces et de l'ensemble des données de l'événement est alors effectuée.

Dans le cas HLT, le temps de latence n'est pas limité à $3.2 \mu\text{s}$ comme dans le cas du déclenchement de niveau 1 mais doit rester de l'ordre de la centaine de ms. De la même manière que dans le cas de premier niveau, le système de haut niveau est organisé selon un menu de déclenchement : nous obtenons finalement une collection d'événements identifiés par les algorithmes de déclenchement HLT qu'ils ont vérifiés.

2.3 Grille de calcul et outils informatiques CMS

2.3.1 La grille de calcul

Initialement introduit à la fin des années 90 par l'idée de partager la puissance de calcul non utilisée des ordinateurs au sein d'un réseau, le concept de grille de calcul, gardant parfois la dénomination anglaise de "computing grid" ou plus simplement "grid", reste très novateur et en voie de développement. Bien que nous ne soyons qu'aux commencements, le partage de la puissance de calcul mais aussi des ressources de stockage se répand dans plusieurs domaines : quelques entreprises commencent à le développer au sein de leurs propres réseaux, le concept portant alors plutôt le nom de "cloud computing", tandis que des laboratoires l'utilisent à l'heure actuelle déjà très largement. Notons au passage que la grille n'est viable que parce que les performances des réseaux ont énormément progressé avec des communications très rapides.

L'idée initiale a évolué selon deux voies distinctes : la première, s'appliquant aux entreprises notamment, suit le concept original : mettre à disposition des autres collaborateurs de la société la puissance de calcul et les ressources de stockage non utilisées par un poste personnel et ainsi optimiser au mieux les ressources du parc informatique de l'entreprise, en particulier lorsqu'il s'agit de multinationales. La deuxième, en revanche, se distingue quelque peu : en effet le concept s'est changé, notamment au niveau des laboratoires, et en particulier dans le domaine des hautes énergies, en celui de fermes de calcul. Il s'agit alors de développer des ensembles d'ordinateurs, les fermes incorporées dans la grille, uniquement alloués aux calculs et/ou au stockage de données.

Les avantages du concept de grille de calcul n'ont pas tardé à se révéler. Ainsi dans le but d'accélérer le développement de nouveaux traitements médicaux, IBM, l'Association Française contre la Myopathie (AFM) et Génomining, société de bioinformatique, ont construit le plus gros ordinateur virtuel : 100000 ordinateurs particuliers ont contribué en quelques mois, chacun par 10 heures de calcul, à ce qu'un seul ordinateur aurait obtenu en 1140 années par traitement séquentiel. Le même type d'exemple est celui de la recherche sur la grippe aviaire H5N1 : 2000 ordinateurs ont permis d'isoler en 4 semaines les médicaments efficaces, contre cent ans pour un seul ordinateur. Le deuxième avantage est plutôt une alternative. En effet depuis les débuts de la technologie informatique à base de transistors, le nombre de ces derniers dans les microprocesseurs suit à peu près la loi de Moore (doublé tous les 18 mois). Cependant nous allons atteindre, d'ici une dizaine d'années, les limites de cette technologie avec notamment l'apparition d'effets quantiques dus aux dimensions infimes des futurs transistors. En attendant le développement de transistors quantiques, le partage via le système de grille apparaît donc comme une alternative de choix pour augmenter la puissance de calcul.

La grille de calcul trouve son utilité dans trois fonctions principales : le calcul intensif, la visualisation et les grandes bases de données. Elle est déjà utilisée dans trois domaines scientifiques : la physique des hautes énergies, les sciences du vivant et les géosciences. Cependant, notons que l'idée de grille de calcul a été essentiellement développée et mise en œuvre par et pour la physique des particules, et notamment afin de répondre aux impératifs du LHC, à savoir les immenses flux de données qui seraient récoltés par les divers détecteurs. Rappelons également que c'est au CERN, qu'est née l'idée de partage d'informations avec le premier Web. Nous voyons ici, ainsi qu'avec l'aspect détecteur détaillé

en partie 2.2, l'importance de l'innovation et du développement de nouvelles technologies pour répondre aux besoins de la physique des hautes énergies ; technologies pouvant être exportées vers d'autres domaines d'application ou dans le monde de l'entreprise.

Dans le cadre de ma thèse, j'ai été amené à produire des échantillons sources pour mes programmes d'analyse grâce à la grille de calcul et au travers de l'interface CRAB propre à CMS. Je vais donc aborder maintenant plus spécifiquement la grille de calcul du CERN. En effet la collaboration LHC, toute expérience confondue, a dû faire face à environ 15 petabytes de données par an. La grille destinée à répondre à ces enjeux, la WLCG (Worldwide LHC Computing Grid), regroupe 100000 processeurs, valeur a minima car en constante évolution, répartis dans 32 pays selon divers échelons d'importance :

- Tier 0 : niveau zéro de WLCG établi au sein du centre de calcul du CERN, sur le site de l'accélérateur dont sont issues les données des détecteurs. Il les stockera et les redistribuera aux niveaux 1, mais il fournit moins de 20% de la capacité totale de calcul.
- Tier 1 : niveau 1 de WLCG réparti sur onze sites : Allemagne, Canada, Espagne, France (CCIN2P3 de Lyon), Italie, Pays-Bas, Pays Scandinaves (Danemark, Finlande, Norvège, Suède), Royaume-Unis, Taïwan, et deux sites aux USA. Ce niveau recevra les données Tier 0, les archivera pour le long terme et fournira l'essentiel des ressources pour la reconstruction.
- Tier 2 : niveau 2 de WLCG réparti sur plus de 140 sites dans le monde. Il sera chargé d'assurer le lien entre données reconstruites des Tier 1 et les physiciens, pour analyse. Par ailleurs ce niveau 2 de la grille fournira environ 50% de la puissance de calcul et sera également amené à stocker provisoirement, pour analyse, quelques pour cents des données Tier 1.
- Tier 3 : niveau 3 de WLCG réparti à l'échelle des laboratoires impliqués dans une expérience LHC

Comme cela a été évoqué en début de partie, ce système de répartition des ressources n'est possible que grâce à un réseau fibre-optique dédié très performant avec 10 Gbit/s entre niveau zéro et niveau un.

2.3.2 Les outils CMS

2.3.2.1 CRAB :Cms Remote Analysis Builder

CRAB est un programme écrit en python dont la fonction est de simplifier les procédés de création et de soumission de jobs d'analyse ou de production d'événements dans l'environnement de la grille. En regroupant tous les paramètres fixés par l'utilisateur au sein d'un fichier de configuration python ".cfg", CRAB génère, pour chaque job, des scripts et des fichiers de données qu'il soumet directement à la grille. Un programmeur, ou scheduler selon la dénomination anglaise, se charge de la distribution des jobs selon les sites, suivant la répartition des données et les paramètres fixés par l'utilisateur.

Un fichier type de configuration contient les groupes suivants :

[CRAB] rassemble les informations sur le type du job ou plus précisément les programmes qu'il utilise, le type du programmeur de la grille ou encore le nom du serveur.

[CMSSW] : contient diverses informations parmi lesquelles nous relevons :

- le chemin et le nom des données sources, en particulier pour des jobs d’analyse ; dans le cas de jobs de production ce champ n’est pas forcément renseigné.
- le nom du fichier python d’analyse ou de production comprenant les paramètres spécifiques de l’analyse ou de la production.
- le nombre total d’événements, à lire ou à générer.
- le nombre de jobs.
- le nom du fichier de sortie.

[USER] : regroupe des paramètres personnels de l’utilisateur comme notamment :

- l’option de retourner directement les données produites ainsi que les noms et chemins des répertoires d’accueil.
- ou l’option de les copier sur des sites de stockages de la grille, le code d’un site particulier de stockage (T2_ FR_ GRIF_ LLR pour les éléments de stockage du LLR par exemple) ainsi que le nom et chemin des répertoires d’accueil.
- la possibilité de publier les données, ainsi que le nom et l’URL de publication.

[GRID] : renseigne les paramètres grille :

- liste des sites de calcul préférentiels (liste blanche).
- liste noire des sites de calcul.
- liste blanche et liste noire des sites de stockage.

L’intérêt de la publication réside dans le fait que la localisation des données, via le chemin du répertoire les contenant et l’URL de publication, est disponible sur le site internet du système de référencement de CMS : DAS (Data Aggregation System). Ceci facilite leur accès pour l’utilisateur et leur partage auprès d’autres membres de la collaboration grâce à un classement de ces données en fonction de leur type et de la version du logiciel qui les a produites notamment.

2.3.2.2 ROOT et RooFit

ROOT [20] est un programme orienté objet qui permet l’analyse de diverses données en associant à la fois un environnement C++, permettant l’exécution directe de codes en C ou C++, et un environnement graphique. Ce dernier autorise l’affichage de spectres, sous forme d’histogrammes de plusieurs dimensions notamment, mais aussi leur manipulation, décoration. *RooFit* est un module orienté statistique qui peut être intégré à ROOT.

Chapitre 3

La reconstruction des électrons dans CMS

Contenu du chapitre

3.1	Introduction	58
3.2	L'algorithme <i>ecaldriven</i>.	59
3.2.1	Le super-agrégat	59
3.2.1.1	Partie tonneau : l'algorithme hybrid	59
3.2.1.2	Parties bouchons : l'algorithme multi5×5	60
3.2.2	La reconstruction des pré-traces et traces	61
3.3	L'algorithme <i>trackerdriven</i>.	65
3.3.1	Reconstruction des pré-traces électron	65
3.3.2	Les agrégats "particle-flow" et l'équivalent du super-agrégat	66
3.4	Combinaison des deux algorithmes	67
3.4.1	L'algorithme	67
3.4.2	Les attributs de l'électron reconstruit.	68
3.4.3	Performances	72

3.1 Introduction

Les performances d'une analyse, quelle qu'elle soit, repose sur les performances de reconstruction et d'identification des objets physiques (électrons, muons...) qu'elle met en jeu. De manière schématique, un électron est constitué d'une trace révélée dans le trajectographe associée à un dépôt d'énergie dans le calorimètre électromagnétique. Concernant les

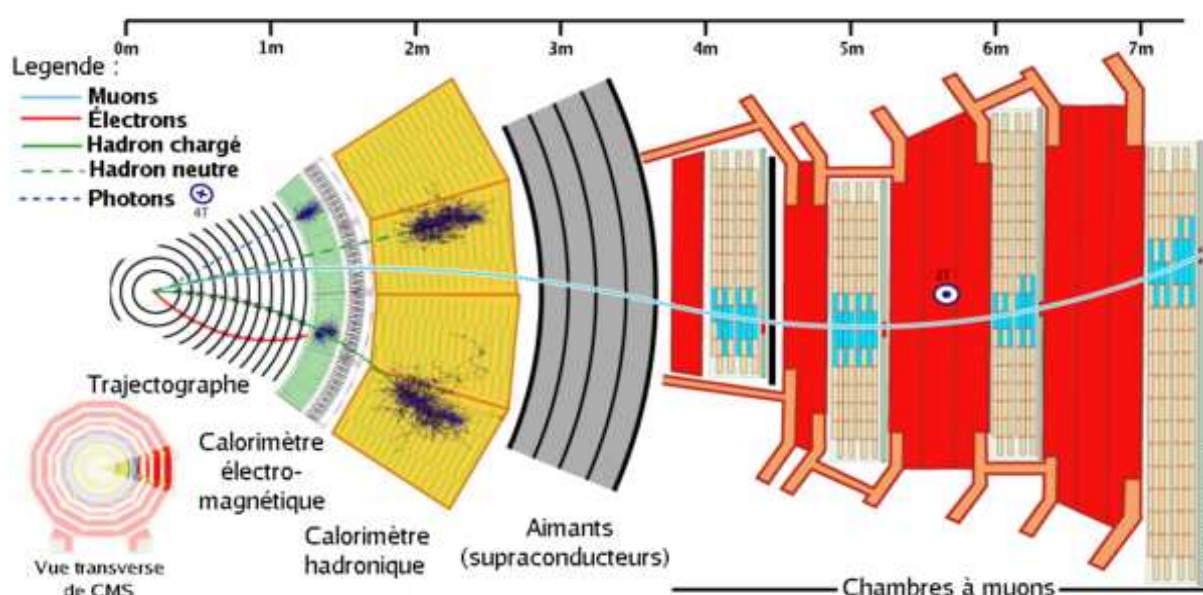


Figure 3.1 – *Trajet des différentes particules dans les divers couches du détecteur.*

électrons, la collaboration CMS utilise deux algorithmes différents. Le premier, dénommé "*ecaldriven*" est initié par la reconstruction d'un super-agrégat de cristaux du calorimètre électromagnétique. Il a été optimisé pour la détection d'électrons isolés, caractéristiques des canaux de désintégration des di-bosons notamment, et qualifiés ainsi car non présents dans des faisceaux de particules provenant de l'hadronisation de quarks et de gluons : les jets. Le deuxième algorithme, qualifié de "*trackerdriven*" ou "*pflow*", ne suit pas la même approche de reconstruction : il est initié par la reconstruction des traces dites "générales", i.e. l'ensemble des traces formées à partir des points d'impact collectés dans le trajectographe en utilisant des paramètres adaptés aux particules subissant une perte d'énergie essentiellement non radiative. A l'inverse du précédent, cet algorithme a été optimisé pour les électrons contenus dans les jets. Nous allons maintenant détailler ces algorithmes de reconstruction ainsi que la combinaison qui en a été faite en vue de cumuler leurs avantages respectifs et que nous illustrerons au travers d'une étude de performances en fin de chapitre.

3.2 L'algorithme *ecaldriven*.

L'électron *ecaldriven* [21] est constitué d'un dépôt dans le calorimètre électromagnétique, représenté au niveau détecteur sous la forme d'un super-agrégat de cristaux activés, associé à une pré-trace à partir de laquelle la trace proprement dite est reconstruite. Un agrégat graine construit autour d'un cristal graine, et défini comme tel car dépassant un certain seuil en énergie, initie la construction du super-agrégat ayant pour but de collecter en plus les agrégats propres aux photons issus du rayonnement de freinage de l'électron. La différence de géométrie du calorimètre électromagnétique entre ses parties bouchons et sa partie tonneau donne lieu à deux procédés de reconstruction de super-agrégats : l'algorithme *hybrid* pour la partie tonneau et l'algorithme *multi5x5* pour les parties bouchons. La pré-trace est quant à elle reconstruite à l'aide des points d'impact collectés dans les couches du trajectographe les plus proches du vertex nominal, et caractérise donc la trajectoire initiale de l'électron.

3.2.1 Le super-agrégat

3.2.1.1 Partie tonneau : l'algorithme hybrid

L'algorithme hybrid, utilisé dans la partie tonneau, regroupe au sein d'un même super-agrégat, l'agrégat propre au dépôt de l'électron et ceux des photons de rayonnement de freinage collectés selon une fenêtre fixe et optimisée en ϕ . Cette stratégie de sommer au sein d'un contour fixe les énergies des cristaux issues de gerbes électromagnétiques se révèle plus performante, dans le cas des électrons et photons isolés, qu'une recherche et collecte dynamique de maxima locaux de dépôts d'énergie. En effet pour des topologies complexes où l'étalement des dépôts peut donner lieu à la fausse identification d'agrégats secondaires, cette tendance est amplifiée dans le cas des algorithmes de recherche dynamique (cf [22]).

Partant d'un cristal graine, ainsi défini car il dépasse le seuil d'énergie $E_T^{hypseed}$, des dominos de 1x3 cristaux étalés en η , chacun ayant son cristal central aligné en η avec le cristal graine, sont formés. Si l'énergie du cristal central de ces dominos dépasse E_{wing} leur dimension est étendue à 1x5 cristaux. Ce procédé est réitéré selon N_{step} pas, de part et d'autre en ϕ du cristal graine. Les dominos ayant une énergie inférieure à E_{thresh} ne sont plus considérés. Enfin tous les agrégats distincts de dominos ayant une énergie totale supérieure E_{seed} ainsi collectés formeront les sous-agrégats du super-agrégat tonneau. La taille du super-agrégat à grande énergie est donc 5 cristaux de large pour ± 17 de long correspondant à ± 0.05 en η et ± 0.3 radian en ϕ . La structure plus allongée en ϕ permet de récupérer de manière optimale les photons issus du rayonnement de freinage et correspond techniquement à une coupure sur la part d'impulsion transverse enlevée à l'électron par le phénomène de bremsstrahlung modélisée par la fraction $\frac{p_T^\gamma}{p_T^e}$. La valeur de cette coupure est de 0.5 et a été établie pour des photons émis très tôt ($r_\gamma \simeq 0$) et de manière à optimiser la reconstruction des électrons jusqu'à $p_T = 5 \text{ GeV.c}^{-1}$. Les différents paramètres sont détaillés dans la table 3.1.

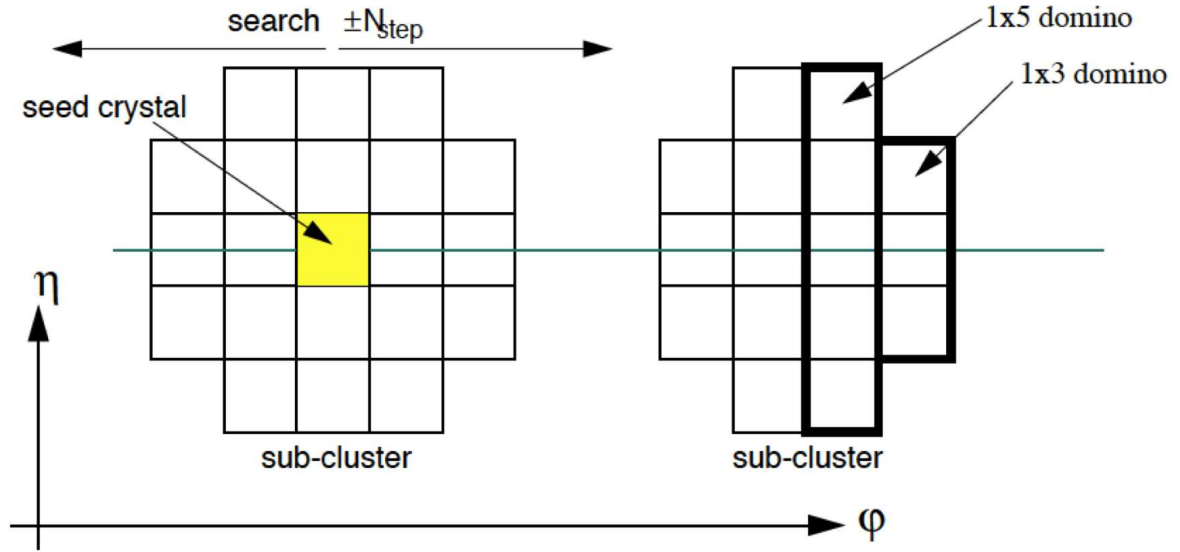


Figure 3.2 – Illustration de la formation des agrégats de dominos de cristaux ECAL selon l'algorithme hybrid [22].

Paramètres	Signification	Valeur
$E_T^{hybseed}$	Seuil pour initier un agrégat	1 GeV
N_{step}	Nombre de pas en ϕ	± 17
E_{wing}	Seuil pour passer à un domino 1×5	1 GeV
E_{thresh}	Seuil pour conserver un domino	0.1 GeV
E_{seed}	Seuil pour conserver un agrégat de dominos	0.35 GeV

Table 3.1 – Paramètres utilisés pour la formation des agrégats dans la partie tonneau du ECAL (algorithme hybrid).

3.2.1.2 Parties bouchons : l'algorithme multi5×5

Contrairement à l'algorithme *hybrid*, qui est par construction un procédé de formation de super-agrégat, l'algorithme *multi5×5* ne forme le super-agrégat qu'a posteriori. Premièrement tout cristal activé représentant un maximum local, au regard de ses quatre cristaux voisins selon une forme de croix, d'énergie transverse supérieure à 0.18 GeV constitue un cristal graine initiant un agrégat de cristaux de dimension 5×5 autour de lui. Un cristal graine, ainsi que ses voisins immédiats, ne peuvent faire partie que d'un seul agrégat, cependant les 16 cristaux constituant la bordure du domino 5×5 peuvent initier, s'ils en remplissent les conditions, d'autres agrégats (cf figure 3.3). Cette particularité autorise la reconstruction d'agrégats malgré le recouvrement des gerbes qui peut advenir notamment entre un électron et ses photons de bremsstrahlung. Une fois ces agrégats primaires formés, la construction du super-agrégat a lieu. Pour cela, une liste d'agrégats primaires graines, c'est-à-dire ayant une énergie transverse supérieure à 1 GeV, est constituée. Une fenêtre étroite en η et allongée en ϕ de part et d'autre de l'agrégat graine est alors utilisée afin d'amalgamer l'ensemble des agrégats qui y sont présents en

un seul et même super-agrégat. Comme dans le cas tonneau, la taille de cette fenêtre est de ± 0.05 en η et ± 0.3 radian en ϕ . Comme nous l'avons vu dans le chapitre 2 un détecteur

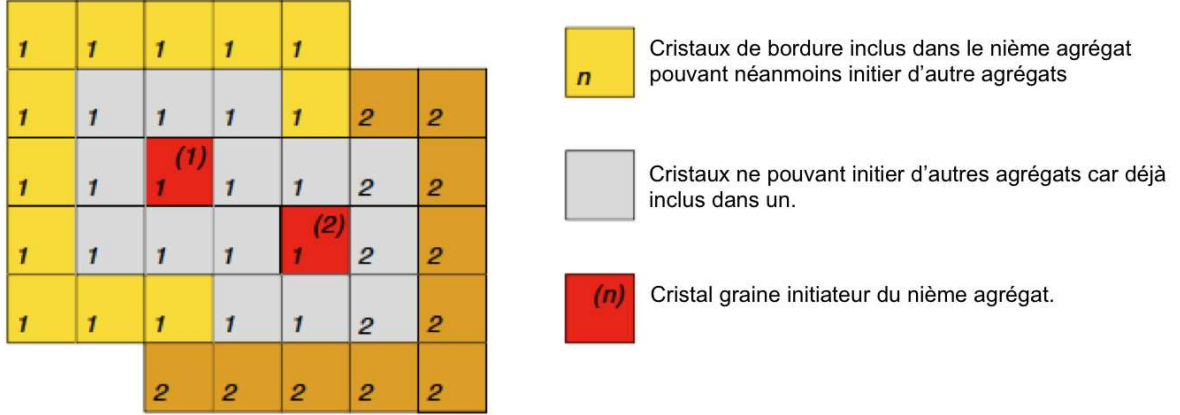


Figure 3.3 – Illustration de la formation des super-agrégats dans les parties bouchons selon l'algorithme multi5x5, cas du recouvrement de deux agrégats primaires [23].

de pieds de gerbes est positionné devant les cristaux des bouchons pour le domaine en $1.653 < |\eta| < 2.6$. Afin de récupérer la fraction d'énergie perdue ($\sim 5\%$) dans ce sous détecteur, la position de chaque cristal activé du calorimètre électromagnétique ainsi que celle du vertex nominal (0,0,0), permettent d'extrapoler au niveau de la couche de capteurs en silicium, partie sensible du détecteur de pieds de gerbes, une zone de recherche. Les capteurs activés servent à estimer le nombre de particules qui laisseront une contribution dans le calorimètre en plomb du détecteur de pieds de gerbe. Cette contribution est prise en compte afin de corriger l'énergie du super-agrégat dans les parties bouchons.

3.2.2 La reconstruction des pré-traces et traces

La reconstruction des traces des particules chargées dans le trajectographe est initiée par la présence de deux ou trois points d'impact dans les couches les plus internes du détecteur : les pré-traces. Parmi ces dernières, celles pouvant initier la reconstruction d'une trace électron sont identifiées grâce à une association avec un super-agrégat d'énergie transverse supérieure à 4 GeV. En effet via la détermination du barycentre du dépôt énergétique de l'électron, la position des premiers points d'impact peut être déterminée et nous pouvons ainsi isoler les candidats pré-traces électron. Cette propriété repose sur les performances de la reconstruction des super-agrégats à collecter correctement les dépôts énergétiques issus du rayonnement de freinage afin de ne pas biaiser la position du barycentre.

L'algorithme de recherche et reconstruction des pré-traces.

Cet algorithme repose sur la propriété suivante : la position du barycentre du super-agrégat peut être déterminée par une simple hélice représentant la trajectoire initiale de l'électron et de ce fait ne dépend pas de l'effet de bremsstrahlung. Ainsi la portion interne de la trajectoire de l'électron, constituée de deux ou trois points d'impact, peut être

prédite en propageant les paramètres de l'hélice à travers le champ magnétique de 3.8T, de la partie externe vers la partie interne du trajectographe où il est très peu probable que soit survenu un rayonnement de freinage.

Tout d'abord l'ensemble des pré-traces, c'est-à-dire des paires ou des triplets de points d'impact dans les couches les plus internes, est constitué. L'algorithme de reconstruction des pré-traces électron va chercher parmi cet ensemble, les pré-traces pouvant correspondre à une trace électron grâce à la propriété énoncée précédemment. Ainsi l'algorithme va tout d'abord isoler le point d'impact le plus interne : une hélice est construite à travers le champ magnétique, à partir du barycentre du super-agrégat localisé à la surface du ECAL, jusqu'au vertex d'interaction nominal (0,0,0), en utilisant l'énergie transverse mesurée pour en déterminer la courbure. Notons qu'à présent l'électron n'est pas encore défini et que par conséquent cette démarche est réalisée pour les deux hypothèses de signe (+,-). Une fois la couche la plus interne atteinte, des fenêtres de recherche en ϕ et z sont définies. La fenêtre en ϕ est construite de manière asymétrique selon l'hypothèse de signe afin de tenir compte de l'incertitude sur la courbure due au rayonnement de freinage tandis que la fenêtre en z (ou r_T pour les parties avant des détecteurs) est définie de façon peu contraignante dans le but de considérer le σ_z caractérisant l'étalement du point d'interaction (cf. table 3.2).

Une fois ce premier point d'impact déterminé, ses propriétés combinées à celles du super-agrégat servent à mettre à jour la coordonnée z du vertex d'interaction. Ensuite une seconde hélice est propagée à partir de ce premier point d'impact jusqu'au ECAL, en conservant l'hypothèse de charge qui a mené à la détermination de ce premier point, afin de définir de nouvelles fenêtres de recherche selon ϕ et z aux paramètres plus contraignants (cf. table 3.3). Si cette deuxième étape trouve un point d'impact compatible, nous définissons par les deux points associés, une pré-trace électron.

E_T^{SC}	δz (Pix. tonneau)	δr_T (Pix. bouchons et TEC)	$\Delta\phi$ [mrad] (e^-)	$\Delta\phi$ [mrad] (e^+)
5 GeV	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-155 ; +75]	[-75 ; +155]
10 GeV	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-96 ; +46]	[-46 ; +96]
35 GeV	$\pm 5\sigma_z$	$\pm 5\sigma_z$	[-54 ; +26]	[-26 ; +54]

Table 3.2 – Paramètres des fenêtres de recherche du premier point d'impact. Pour $E_T^{SC} < 5$ GeV, les paramètres sont ceux de $E_T^{SC} = 5$ GeV. Pour $E_T^{SC} > 35$ GeV les paramètres sont ceux de $E_T^{SC} = 35$ GeV.

δz (Pix. tonneau)	δr_T (Pix. bouchons)	δr_T (TEC)	$\Delta\phi$ (Pix. tonneau)	$\Delta\phi$ (Pix. bouch., TEC)
$\pm 0.9\text{mm}$	$\pm 1.5\text{mm}$	$\pm 2\text{mm}$	$\pm 4\text{mrad}$	$\pm 6\text{mrad}$

Table 3.3 – Paramètres des fenêtres de recherche du second point d'impact.

L'algorithme CKF (Combinatorial Kalman Filter) : application de filtres de Kalman.

Ces pré-traces électrons servent alors de germes à la procédure de reconstruction des traces via la technique dite "Filtre de Kalman". Cette technique permet de déterminer une zone de recherche de points d'impact dans la couche suivante (n+1) à partir des points de mesure des couches précédentes et des paramètres de l'embryon de trace constitué par ces derniers. Le point d'impact trouvé, les paramètres du début de trace sont alors réajustés. La procédure reprend pour la recherche du prochain point de la trace(n+2). Ainsi pour chaque pré-trace, un ensemble de candidats traces est reconstruit. Ensuite un ajustement est opéré sur chacun d'eux : en partant du dernier point de mesure constituant la trace jusqu'au point de mesure (n), une mise à jour de l'estimation de position et de la direction associées au point de mesure de la couche précédente (n-1) est effectuée et sert à renforcer l'identification des points d'impact appartenant réellement à la trace. Cette technique est équivalente à un ajustement par moindre carré pour les cas où les diffusions coulombiennes multiples dues à la présence de matériel constituent l'effet dominant et peuvent être modélisées avec des fluctuations gaussiennes, comme dans le cas des muons par exemple.

Appliqué aux électrons, cette méthode présente l'inconvénient de conduire à des traces courtes, la reconstruction de la trace par l'algorithme CKM étant stoppée lorsque surviennent d'importants changements de courbure dus au rayonnement de freinage.

L'ajustement GSF(Gaussian Sum Filter).

Dans le cas des électrons soumis au rayonnement de freinage, l'approximation gaussienne n'est plus valable. L'algorithme GSF modélise, par une superposition de distributions gaussiennes, la densité de probabilité $f(z)$ de perte d'énergie par bremsstrahlung définie dans le modèle de Bethe-Heitler [24]

$$f(z) = \frac{[-\ln(z)]^{c-1}}{\Gamma(c)} \quad (3.1)$$

avec

- $c = \frac{t}{\ln(2)}$ et t , l'épaisseur de matériel traversée, exprimée en unités de longueur de radiation
- z , fraction d'énergie restant à l'électron après la traversée du matériel.

L'algorithme GSF recherche donc les points d'impact compatibles de la couche suivante après une propagation rendant compte de la présence de matériel entre les deux couches et des pertes radiatives associées. La compatibilité entre la zone de recherche prédite et les points d'impact réellement mesurés est évaluée par un test de χ^2 . Si plusieurs points d'impact compatibles sont trouvés, les trajectoires correspondantes sont reconstruites en parallèle, en ne gardant par électron que les deux meilleures traces, i.e. celles minimisant le χ^2 . Cette procédure s'arrête une fois la dernière couche atteinte ou si aucun point d'impact compatible n'a été trouvé dans deux couches successives.

Ainsi grâce à la méthode GSF, et au contraire de la technique CKF, nous pouvons reconstruire des traces jusqu'au ECAL malgré la perte par effet de Bremsstrahlung. Nous obtenons alors des traces plus longues comme le montre la figure 3.4 représentant la distribution du nombre de points d'impact collectés par trace pour chacun des algorithmes.

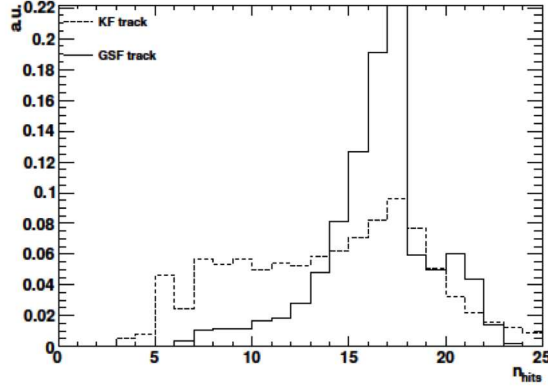


Figure 3.4 – Nombre de points d’impact collectés par trace pour l’algorithme GSF (ligne pleine) et pour l’algorithme CKF (ligne pointillée) sur un échantillon d’électrons simulés provenant de $Z \rightarrow e^+ e^-$.

L’algorithme complet de reconstruction des traces électrons

Les pré-traces électrons initient une procédure de recherche de traces électrons, suivant l’algorithme CKF décrit plus haut, puis les traces trouvées subissent un traitement GSF pour tenir compte des pertes par rayonnement de freinage. Ainsi lors du premier traitement CKF, le test de χ^2 à chaque étape est maintenu à une valeur très peu contraignante de 2000. Les combinaisons possibles sont cependant limitées à cinq candidats traces pour chaque couche du trajectographe avec au maximum un seul point d’impact manquant. De plus, afin de réduire la probabilité de combiner la trace d’un électron primaire avec celle d’un électron issu de la conversion d’un photon, un χ^2 limité à 90 est requis dans les cas présentant un point d’impact manquant.

Les points de trace ainsi collectés sont ensuite soumis à l’algorithme GSF pour l’estimation finale des paramètres de la trace où la perte d’énergie dans chaque couche est modélisée par une somme pondérée de distributions gaussiennes. Les paramètres physiques de la trace peuvent alors être évalués selon deux méthodes. La première, dite ”mean”, renvoie la moyenne pondérée des paramètres obtenus par chacune des gaussiennes et, par conséquent, décrit un comportement moyen en modélisant complètement la perte par bremsstrahlung, ce qui en fait la technique d’estimation la moins biaisante. La seconde, dite ”mode”, ne considère que les gaussiennes de poids maximum, ce qui a pour effet d’être plus précis pour les traces ayant peu rayonnée comme nous pouvons le voir sur la figure 3.5 ; sur cette même figure nous remarquons qu’au contraire les traces ayant subies d’importantes pertes par bremsstrahlung voient leur impulsion reconstruite sous-estimée.

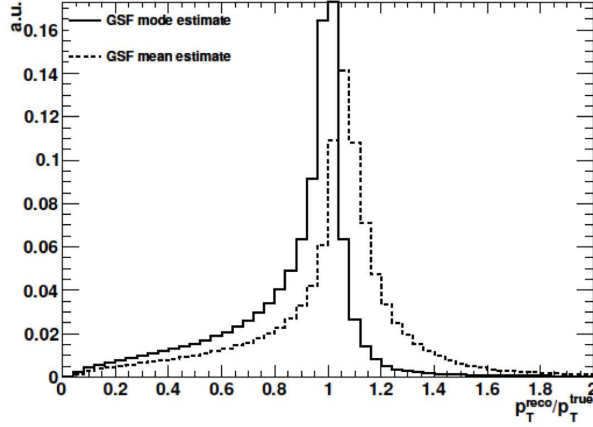


Figure 3.5 – Impulsion transverse de la trace évaluée pour le "mode" (ligne pleine) et pour le "mean" (ligne pointillée) obtenue sur un échantillon d'électrons simulés provenant de $Z \rightarrow e^+ e^-$.

3.3 L'algorithme *trackerdriven*.

3.3.1 Reconstruction des pré-traces électron

Au contraire de l'algorithme *ecaldriven*, l'algorithme *trackerdriven* part des traces dites "générales". Ces traces générales, ou traces CKF, sont initiées par des pré-traces et formées par toutes les combinaisons de points d'impact, selon la technique des filtres de Kalman détaillée dans la partie 3.2.2, sans pour autant être associées à un quelconque type de particule : électrons, muons ou encore hadrons chargés. Il s'agit alors pour l'algorithme *trackerdriven* d'identifier en premier lieu celles qui appartiennent à un électron : c'est l'étape de pré-identification de l'électron, c'est-à-dire une association entre trace et agrégat électron. Lorsque le rayonnement de freinage est peu important, la trajectoire de l'électron peut être approximée par une hélice. L'algorithme CKF fournit dans ce cas particulier une très bonne estimation de l'impulsion de la trace et de ce fait une association entre agrégat et trace reposant sur des critères de position en η et ϕ ainsi que sur un rapport entre énergie déposée dans le ECAL et impulsion de la trace, E/p , proche de l'unité permet d'identifier les traces électrons. Ces traces sont ensuite réajustées selon l'algorithme GSF pour tenir compte d'éventuelles perturbations dues au rayonnement de freinage. Dans les cas où la trace CKF ne peut être pré-identifiée, cela peut s'expliquer par une altération due à l'effet de bremsstrahlung trop important, elle subit alors un réajustement GSF pour après, faire de nouveau l'objet d'une tentative d'association avec un agrégat électron. Nous obtenons à l'issue de ces étapes des candidats traces électron. Ensuite une analyse multivariable reposant sur un arbre de décision mettant en jeu différents paramètres de qualité de la trace tels que χ_{GSF}^2 , χ_{KF}^2/χ_{GSF}^2 , son nombre de points de mesure, la perte d'énergie mesurée lors de sa reconstruction, la qualité de l'association trace-agrégat, détermine la trace électron.

3.3.2 Les agrégats "particle-flow" et l'équivalent du super-agrégat

La formation des agrégats de dépôts d'énergie dans le ECAL pour reconstruire les électrons selon l'algorithme trackerdriven ne suit pas la même approche que celle détaillée dans la partie 3.2.1 mais suit la procédure développée par l'algorithme de reconstruction "particle-flow". Elle procède en trois étapes :

- la première est d'isoler des cellules "graines" définies comme telles car maxima locaux d'énergie dépassant un seuil E_{seed}^{min} défini dans la table 3.4, autour desquelles seront ensuite agrégées d'autres cellules activées .
- lors de la seconde étape, l'agrégat proprement dit est formé selon un algorithme de type topologique, c'est-à-dire en associant à la graine dans un premier temps, à l'agrégat en formation ensuite, les cellules activées ayant au moins un coté en commun avec la ou les cellules déjà agrégées et dépassant également un certain seuil en énergie, E_{min} . Ce seuil, défini également en table 3.4, représente les déviations standards dues au bruit électronique.
- la dernière étape consiste à déterminer de manière itérative, en exploitant la granularité du calorimètre, l'énergie et la position de l'agrégat "particle-flow" obtenu en fin de seconde étape. Pour ce faire :
 - Premièrement, la position de l'agrégat "particle-flow" est simplement celle de la cellule graine tandis que l'énergie de chaque cellule activée est partagée entre tous les agrégats "particle-flow" proportionnellement à $\exp(-d_{ij}^2/R^2)$, avec d_{ij} la distance entre l'agrégat "particle-flow" i , i.e. lors de cette première itération la cellule graine i , et la cellule j .
 - Lors de la deuxième itération, la position de l'agrégat devient le barycentre de neuf de ses cellules centrales, à savoir la cellule graine et ses huit voisines, pondéré par le logarithme du rapport entre l'énergie de ces cellules et le seuil E_{min} . L'énergie de l'agrégat est recalculée de la même manière que lors de la première itération, en prenant cette fois-ci la nouvelle position.
 - Cette dernière étape est réitérée jusqu'à convergence obtenue lorsque la position de l'agrégat "particle-flow" ne varie que d'une simple fraction de la résolution en position.

Notons également que la position de l'agrégat "particle-flow" est corrigée du fait que les cellules ne pointent pas vers le point d'interaction moyen.

Paramètres	ECAL tonneau	ECAL bouchons	détecteur pied de gerbe
E_{min} (GeV)	0.08	0.3	7.10^{-6}
E_{seed}^{min} (GeV)	0.23	0.8	5.10^{-4}
R (cm)	5	5	0.2
# cellule pour calcul position	9	9	toutes

Table 3.4 – Paramètres utilisés pour la formation des agrégats dans l'algorithme particle-flow.

A l'issu des ces trois étapes nous avons formé des agrégats "particle-flow" caractéristiques d'une activité électromagnétique détectée dans le ECAL et propres aux électrons et photons. L'identification de ceux appartenant à des électrons se fait comme décrit dans la

partie 3.3.1.

Les composantes électron déterminées, le but n'est pas de former un super-agrégat mais

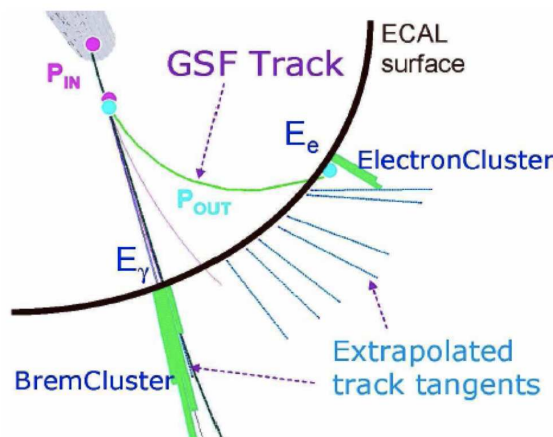


Figure 3.6 – Illustration de la reconstruction d'un électron par l'algorithme trackerdriven.

plutôt d'attacher à la trace électron une liste d'agrégats : le premier étant l'agrégat électron attaché à l'extrémité de l'extrapolation de la trace, les autres sont les agrégats des photons de bremsstrahlung et ainsi identifiés par la technique des tangentes : à chaque point d'impact dans le trajectographe, une tangente à la trace est propagée jusqu'au calorimètre électromagnétique et si un agrégat de cristaux peut lui être associé, il est alors identifié comme provenant d'un photon de bremsstrahlung. Une recherche des photons de bremsstrahlung convertis est également menée, d'une part pour améliorer la collecte d'énergie et d'autre part afin d'identifier les traces correspondantes.

3.4 Combinaison des deux algorithmes

Chacun des deux algorithmes présente ses propres avantages. L'algorithme *ecaldriven*, par sa fenêtre fixe pour former le super-agrégat électron, permet de reconstruire de manière performante les électrons isolés et les électrons de $p_T > 10 \text{ GeV} \cdot c^{-1}$ pour lesquels le rayonnement de freinage reste proche de la trace de l'électron. L'algorithme *trackerdriven* permet, par la technique des tangentes, de récupérer et d'isoler correctement les différents dépôts électromagnétiques de l'électron notamment dans les cas particuliers d'électrons de faible impulsion transverse dont la courbure importante de la trace peut éloigner les agrégats de bremsstrahlung au-delà de la taille limite des super-agrégats *ecaldriven*. La combinaison des deux méthodes permet d'améliorer la reconstruction des électrons isolés en particuliers dans la région des bas p_T . Un des premiers travaux de ma thèse a été d'évaluer les performances de cette combinaison.

3.4.1 L'algorithme

L'algorithme combiné commence par la construction de manière indépendante des pré-traces par chacun des algorithmes telle que présentée dans les parties précédentes. L'étape

suivante réunit les pré-traces issues des deux algorithmes puis identifie celles communes aux deux méthodes en vérifiant qu'elles partagent l'ensemble de leurs points d'impact :

- Dans l'affirmative une référence sur la trace *trackerdriven* est ajoutée à la pré-trace *ecaldriven*. Cet objet est alors identifié comme provenant des deux algorithmes et remplit la liste finale des pré-traces.
- Dans la négative les pré-traces uniquement *trackerdriven* et uniquement *ecaldriven* sont également ajoutées à la liste finale des pré-traces.

A l'issue de cette étape de mélange, les pré-traces servent alors à reconstruire les traces via l'algorithme CKF dédié aux électrons et l'ajustement GSF (cf 3.2.2).

Puis l'association est réalisée entre trace et super-agrégat en tenant cette fois-ci compte des équivalents de super-agrégats fournis par un module dénommé "PFTranslator" qui "traduit" en architecture type super-agrégat, les agrégats électrons et les agrégats de photons de bremsstrahlung reconstruits par la méthode *particle-flow*, tout en ajoutant le résultat "mva" issu de l'analyse multivariable. Nous obtenons finalement une collection de candidats électrons reconstruits par l'un ou l'autre ou les deux algorithmes *ecaldriven* et *trackerdriven*.

3.4.2 Les attributs de l'électron reconstruit.

A l'issue de cet algorithme combiné de reconstruction nous obtenons un candidat électron pourvu de divers variables de qualité, d'identification et d'isolation spécifiques aux deux algorithmes. En plus du quadrivecteur impulsion, de l'impulsion transverse, de références sur la trace et sur le super-agrégat ainsi que leurs attributs propres, de variables booléennes indiquant l'algorithme qui l'a reconstruit, l'électron possède notamment :

- **E** : l'énergie de l'électron n'est pas seulement déterminée par le dépôt électromagnétique mais est le résultat d'une combinaison E-p. Si la direction du vecteur impulsion est toujours mieux déterminée par les paramètres de la trace reconstruite, il n'en est pas de même pour son amplitude. En effet le calorimètre électromagnétique présente une précision qui décroît avec l'énergie incidente. A l'inverse le trajectographe présente une meilleure mesure de l'impulsion pour de faibles énergies (<15 GeV) ainsi que pour les cas d'électrons proches des cracks. Par conséquent l'estimation de l'impulsion de l'électron reconstruit est obtenue par une combinaison des informations issues du ECAL et de celles provenant du trajectographe dont les performances sont illustrées par la figure 3.7. La résolution y est présentée en termes de déviation standard effective (effective RMS), cette dernière représentant la demie-taille de la plus petite fenêtre contenant 68% de la distribution ; fenêtre construite de part et d'autre du maximum de la distribution de l'énergie.

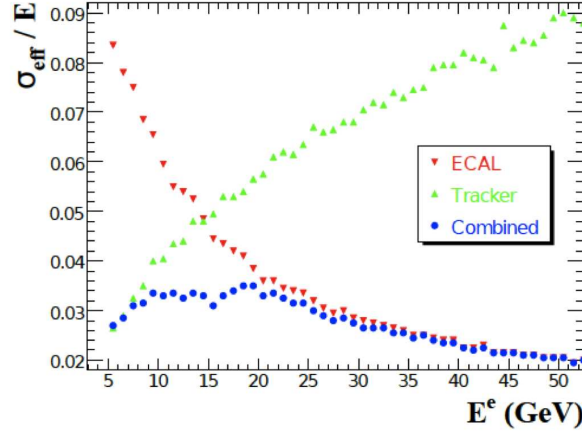


Figure 3.7 – Résolution en termes de déviation standard effective pour l'énergie du dépôt électromagnétique (triangles inversés rouges), pour l'impulsion de la trace (triangles verts) et pour la combinaison (ronds bleus).

- **La charge.** Elle est déterminée par les résultats de trois algorithmes différents via une logique majoritaire : la valeur commune à au moins deux des trois algorithmes définit la charge. Elle est établie par la courbure de la trace pour deux de ces algorithmes (charge trace Gsf et charge trace Ctf), ou par la déviation par rapport au point origine entre la position du premier point d'impact de la trace et celle du super-agrégat. (Cf [25])
- **E/p ,** rapport entre l'énergie du super-agrégat et l'impulsion de la trace. Cette variable rend compte de l'adéquation entre la mesure du dépôt d'énergie de l'électron et la reconstruction de sa trace. Une valeur proche de 1, comme le montre la figure 3.8, caractérise un électron bien reconstruit que ce soit du point de vue de sa trace ou de son dépôt électromagnétique.

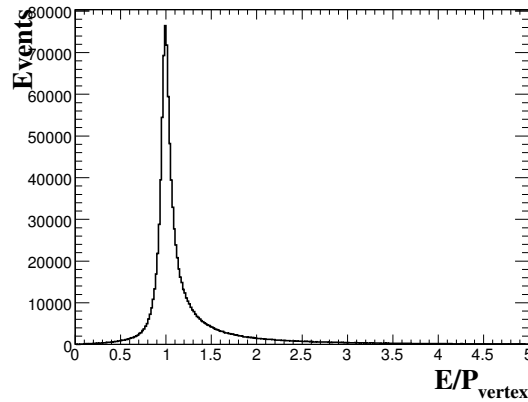


Figure 3.8 – Rapport E/p pour des électrons simulés issus d'un échantillon $Z \rightarrow e^+e^-$

- **f_{brem} ,** variable mesurant l'effet de bremsstrahlung

$$f_{\text{brem}} = \frac{p_{\text{in}} - p_{\text{out}}}{p_{\text{in}}}$$

avec p_{in} , valeur de l'impulsion au point le plus proche du vertex de création de la trace et p_{out} , valeur de l'impulsion en son point le plus éloigné.

Une valeur proche de zéro est caractéristique d'un électron ayant peu rayonné au contraire d'une valeur proche de 1. Cette variable, ingrédient essentiel des algorithmes d'identification, est illustrée par la figure 3.9.

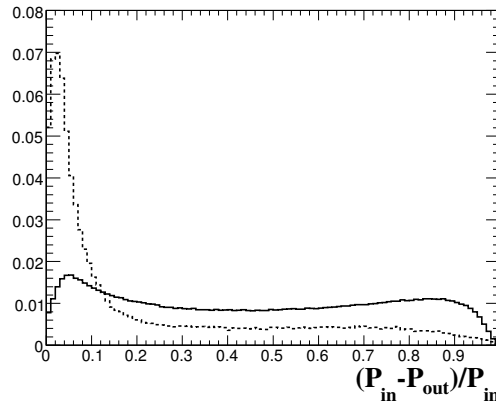


Figure 3.9 – f_{brem} pour des électrons simulés issus du canal $Z \rightarrow e^+ e^-$ (ligne pleine) et pour du bruit de fond QCD d'événements di-jets de $80 < p_T < 120 \text{ GeV}/c$ (ligne pointillée).

- $\sigma_{i\eta i\eta}$ caractérise la largeur de la gerbe électromagnétique :

$$\sigma_{i\eta i\eta} = \frac{\sum_j w_j (i\eta_j - i\bar{\eta}_{SC})}{\sum_j w_j}$$

avec le poids associé au cristal j : $w_j = 4.2 + \ln\left(\frac{E_j}{E_{5 \times 5}}\right)$

où j désigne l'itération sur les cristaux d'énergie E_j et d'index $i\eta_j$ selon η formant le super-agrégat de pseudo-rapacité moyenne $i\bar{\eta}_{SC}$. $E_{5 \times 5}$ est l'énergie du domino de dimension 5×5 cristaux.

Un vrai électron possède un $\sigma_{i\eta i\eta}$ de faible valeur. Notons que la largeur est ici évaluée seulement suivant η afin d'éviter la distorsion en ϕ due au rayonnement de freinage.

- $\Delta\eta$, $\Delta\phi$ caractérisant l'association trace-Super-agrégat :

$$\begin{aligned} \Delta\eta &= \eta_{SC} - \eta_{in}^{extrap} \\ \Delta\phi &= \phi_{SC} - \phi_{in}^{extrap} \end{aligned}$$

avec η_{SC}, ϕ_{SC} , coordonnées du super-agrégat et $\eta_{in}^{extrap}, \phi_{in}^{extrap}$, coordonnées de l'extrapolation du point d'impact de la trace à la surface du ECAL. Plus les valeurs de ces variables sont proches de zéro, plus le candidat reconstruit a de chance d'être un vrai électron comme nous pouvons le voir dans le cas de vrais électrons sur la figure 3.10.

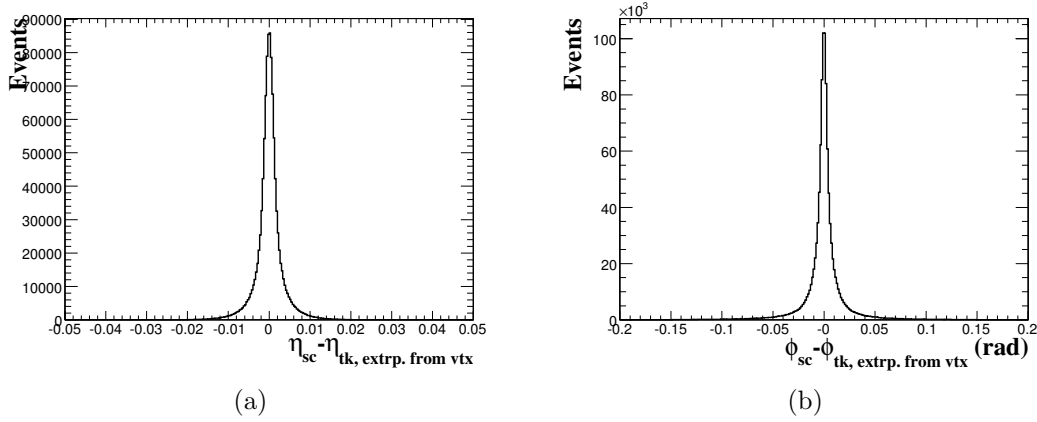


Figure 3.10 – $\Delta\eta$ (a) et $\Delta\phi$ (b) pour des électrons simulés issus d'un signal $Z \rightarrow e^+ e^-$.

- **H/E** rapport entre énergies hadronique et électromagnétique déposées dans la région du super-agrégat. Il doit être faible pour de vrais électrons. E est l'énergie du super-agrégat. H désigne l'énergie déposée dans les tours activées du HCAL se situant dans un cône de rayon égal à 0.15 centré sur la position du super-agrégat.
- **mva**, résultat de l'analyse multivariable accomplie via arbre de décision par l'algorithme *trackerdriven* (cf section 3.3)
- **Classification**. En fonction de certaines variables précédemment définies, des classes d'électrons sont établies
 - "golden" : Super-agrégat formé d'un seul agrégat, $E/p > 0.9$, $f_{brem} < 0.5$. Il s'agit d'électrons peu affectés par le rayonnement de freinage et dont la trace reconstruite est correctement associée au super-agrégat.
 - "big brem" : Super-agrégat formé d'un seul agrégat, $E/p > 0.9$, $f_{brem} > 0.5$. Les électrons sont dans ces cas hautement affectés par l'effet de bremsstrahlung mais ceci ne se traduit pas par une dégradation de la mesure de l'énergie.
 - "showering" : Super-agrégat formé d'un seul agrégat mais dont l'électron n'appartient pas aux deux catégories précédentes ou super-agrégat formé d'une collection d'agrégats. Ces électrons ont subi une perte d'énergie importante par rayonnement de freinage.
 - "crack" : électrons dont le cristal graine se situe proche d'une limite de structure du détecteur telle qu'entre les modules de la partie tonneau du ECAL ou dans les zones de transition entre le tonneau et les parties bouchons.

À ces diverses variables d'identification ou rendant compte de la qualité de la reconstruction de l'électron, s'ajoutent les variables d'isolations. Dans le cas *ecaldriven* il s'agit d'isolations calculées sur la base d'éléments détecteurs :

- **trackIso** ou isolation trace : somme des impulsions transverses des traces présentes dans le trajectographe et regroupées au sein d'un cône $\Delta R < 0.3$ autour de la trace électron tout en prenant soin de ne pas considérer la composante même de ce dernier grâce à l'utilisation d'un cône de veto de dimension $\Delta R < 0.015$.
- **ecalIso** ou isolation ECAL : somme des énergies transverses déposées dans le ECAL,

regroupées au sein d'un cône $\Delta R < 0.3$ sans la composante électron isolée à l'aide d'un veto cône interne ainsi qu'une bande veto interne pour rejeter la composante issue du rayonnement de freinage.

- **hcalIso** ou isolation HCAL : somme des énergies transverses déposées dans le HCAL, regroupées au sein d'un cône $\Delta R < 0.3$.

Plus généralement c'est leur valeur relative, i.e. divisée par l'énergie transverse de l'électron, qui est considérée.

Dans le cas de l'algorithme *particle-flow*, nous nous servons de la connaissance que nous avons de l'événement et de la nature des particules reconstruites par l'algorithme en identifiant trois composantes : "photons", "hadrons chargés" et "hadrons neutres", résultat de la somme, au sein d'un cône $\Delta R < 0.3$, de l'énergie transverse des types de particules respectivement considérées.

3.4.3 Performances

Pour achever la reconstruction, les candidats électrons sont soumis à une pré-sélection afin de réduire le taux de jets étant faussement reconstruits comme électrons. Pour les candidats reconstruits par l'algorithme *ecaldriven*, les coupures suivantes sont utilisées :

- $E_T > 4 \text{ GeV}$
- $H/E < 0.15$
- $|\Delta\eta| < 0.02$
- $|\Delta\phi| < 0.15$

Pour ceux reconstruits par l'algorithme *trackerdriven* seulement, une valeur $mva > -0.4$ est requise. A ce stade, les coupures utilisées ne sont que très peu contraignantes permettant de conserver une efficacité maximale compatible avec un taux de faux électrons admissible pour le stockage.

Au début de ma thèse, j'ai travaillé à l'estimation des performances de la combinaison des deux algorithmes *ecaldriven* et *trackerdriven* dans le cas d'électrons isolés. La mesure de l'efficacité de cette combinaison a été faite à deux niveaux : au niveau de la formation des pré-trace et au niveau de la reconstruction elle-même, c'est-à-dire au niveau de l'association trace-Super-agrégat.

L'étude a été menée sur un échantillon d'électrons générés d'impulsion transverse ayant une distribution plate pour $p_T^e > 2 \text{ GeV.c}^{-1}$, en associant la pré-trace reconstruite à la trace de l'électron généré dans un cône de rayon 0.15 autour de la pré-trace.

En examinant la figure 3.11, nous pouvons constater que bien que l'algorithme *trackerdriven* ait été initialement développé pour reconstruire des électrons non isolés, sa contribution permet d'améliorer l'efficacité de formation des pré-traces, en particulier dans des régions difficiles : régions "cracks" ($\eta \simeq 0$ et $\eta \simeq 1.5$) ainsi qu'à faible impulsion transverse. Ainsi l'efficacité de reconstruction des pré-traces est augmentée de 12.5% à 5 GeV.c^{-1} grâce à la combinaison des deux algorithmes tandis qu'au niveau du plateau atteint à hauts p_T , l'augmentation prend une valeur constante comprise entre 1 et 2%.

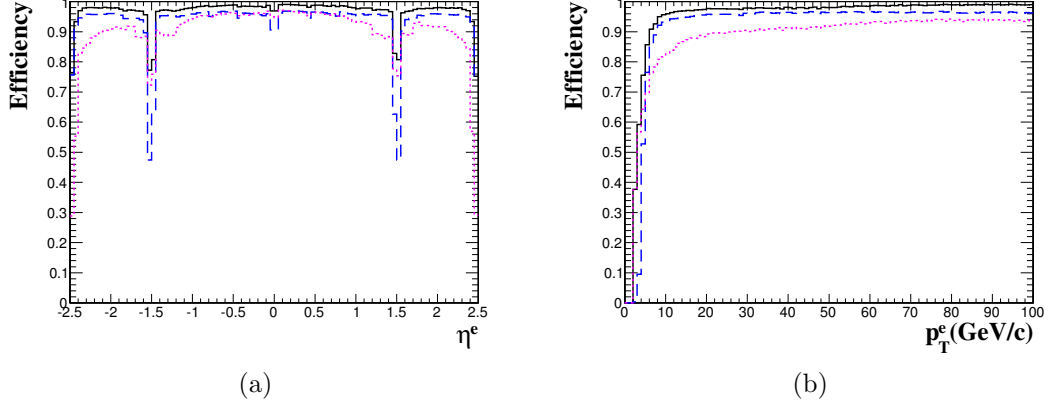


Figure 3.11 – Efficacité de reconstruction des pré-traces (ligne noire) en fonction du η^e (a) et du p_T^e (b) de l'électron généré. Les contributions propres aux algorithmes *ecaldriven* (ligne traits bleus) et *trackerdriven* (ligne pointillée rose) sont également affichées.

Concernant l'efficacité de reconstruction après pré-sélection (figure 3.12), l'évaluation a été faite sur le même échantillon en associant un électron reconstruit à un généré au sein d'un cône de rayon 0.05 autour de la direction du reconstruit tout en demandant que la charge de l'électron reconstruit corresponde à celle de l'électron généré. De même nous

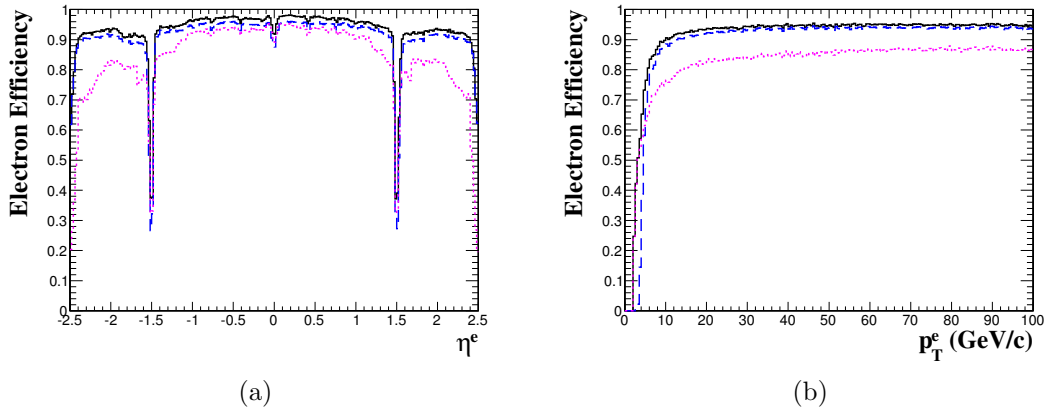


Figure 3.12 – Efficacité de reconstruction des électrons après pré-sélection (ligne noire) en fonction du η^e (a) et du p_T^e (b) de l'électron généré. Les contributions propres aux algorithmes *ecaldriven* (ligne traits bleus) et *trackerdriven* (ligne pointillée rose) sont également affichées.

observons une augmentation d'efficacité de reconstruction au niveau des zones de cracks et à bas p_T . Pour les régions de transition tonneau-bouchons cela s'explique par le bémol de l'approche *ecaldriven* relatif à la formation des super-agrégats dans ce cas particulier impliquant des différences de structures du ECAL entre les parties bouchons et tonneau. Comme nous l'avons détaillé dans le paragraphe 3.2.1, l'algorithme ne reconstruit pas de

la même manière ses super-agrégats suivant qu'ils se situent dans l'une ou l'autre des parties. Ainsi la zone de recouvrement pose problème du point de vue *ecaldriven*. A l'inverse, la méthode *trackerdriven* étant initiée par la reconstruction de la trace et n'établissant de correspondance avec un équivalent de super-agrégat qu'a posteriori, elle s'avère plus performante dans ces zones ; sans oublier que cet équivalent de super-agrégat est un ensemble d'agréats distincts mais reconstruits par un algorithme commun aux différentes parties du calorimètre électromagnétique et de ce fait peut s'affranchir du problème causé par les recouvrements tonneau-bouchons. Concernant l'amélioration à basse impulsion transverse, elle s'explique par la grande courbure que subissent des électrons de bas p_T soumis à un champ magnétique de 3.8 T et qui peut amener les pré-traces à se trouver hors du champ de recherche de l'algorithme *ecaldriven* basé sur la position de la trace extrapolée dans les couches les plus internes du trajectographe à partir de la position du super-agrégat.

En conclusion la combinaison des deux algorithmes permet d'améliorer les efficacités dans les zones de craks ainsi qu'à bas p_T . Pour la reconstruction des pré-traces, nous obtenons une efficacité voisine de 95% pour $p_T^e = 10 \text{ GeV.c}^{-1}$ et approchant 100% pour $p_T^e = 100 \text{ GeV.c}^{-1}$. Concernant l'efficacité de reconstruction, elle atteint $\simeq 90\%$ pour $p_T^e = 15 \text{ GeV.c}^{-1}$ et jusqu'à $\simeq 95\%$ pour $p_T^e = 100 \text{ GeV.c}^{-1}$.

L'évaluation de l'efficacité de reconstruction sur les données recueillies par le détecteur CMS est détaillée dans le chapitre 4.

Chapitre 4

Validation des algorithmes et mesure de l'efficacité de reconstruction des électrons dans CMS

Contenu du chapitre

4.1 Démarrage, certification des données et mesure des distributions des électrons	76
4.1.1 DQMOffline : certification des données	76
4.1.2 Mesure des distributions des électrons et comparaison avec le Monte-Carlo, données 2011	77
4.2 Sélection des événements et effet sur la mesure d'efficacité .	82
4.3 Mesure de l'efficacité de reconstruction des électrons	86
4.3.1 La méthode "tag-and-probe" et mesure d'efficacité par simple comptage	86
4.3.1.1 Efficacité et calcul des incertitudes statistiques	87
4.3.1.2 Résultats	87
4.3.2 Mesure de l'efficacité par une méthode d'ajustement	88
4.3.2.1 Fonctions de modélisation signal et bruit de fond . . .	89
4.3.2.2 Résultats obtenus avec les données 2010	91
4.3.2.3 Résultats obtenus avec les données 2011	94
4.3.2.4 Estimation des incertitudes systématiques	95
4.4 Résultats finaux et conclusion	100

Une reconstruction performante des objets électrons est nécessaire dans les analyses de canaux multileptoniques telles $ZZ \rightarrow 4l$ ($l=e, \mu$) que ce soit dans le but de mesurer des sections efficaces ou dans un objectif de découverte. En effet l'efficacité de reconstruction de $ZZ \rightarrow 4e$ est, en première approximation, celle d'un électron à la puissance quatre. Je vais donc présenter dans ce chapitre le travail que j'ai effectué sur la mesure de l'efficacité de reconstruction des électrons issus des données récoltées par le détecteur CMS et que j'ai réalisé en utilisant la technique "tag-and-probe" (marquer et sonder). Cette technique construit des paires *tag-probe* formant une masse invariante proche d'une résonance connue, par exemple celle du Z à $m_{tag-probe} \simeq 91.2 \text{ GeV.c}^{-2}$, requérant des critères de sélection très stricts sur le tag et au contraire beaucoup moins contraignants sur le probe. Elle permet de sélectionner un lot pur d'électrons par exemple et de mesurer ou présenter des caractéristiques particulières au probe. Ainsi dans un premier temps j'introduirai les outils que j'ai développés lors du démarrage en 2009 afin de réaliser une certification des données recueillies. Ensuite je détaillerai les coupures de sélection d'événements selon la méthode tag-and-probe, et leurs effets. Puis j'expliquerai plus précisément la technique et enfin je présenterai les efficacités de reconstruction des électrons pour les données acquises en 2010 ainsi que pour celles acquises en 2011, et effectuées par simple mesure "cut and count" (couper et mesurer) ou à l'aide d'une méthode d'ajustement appliquée sur le spectre de masse du Z .

4.1 Démarrage, certification des données et mesure des distributions des électrons

Au démarrage du LHC en 2009, coïncidant avec celui de ma thèse, j'ai travaillé à l'élaboration d'un programme d'analyse pour les électrons à incorporer dans l'outil de surveillance et de certification des données de CMS, le *DQMOOffline* (Data Quality Monitoring Offline). Je vais dans un premier temps décrire et illustrer ce programme d'analyse, puis je présenterai le travail de validation des algorithmes de reconstruction pour les électrons que j'ai réalisé sur une partie des données 2011 en utilisant la méthode tag-and-probe afin de sélectionner un échantillon pur d'électrons.

4.1.1 DQMOOffline : certification des données

A l'accumulation de données recueillies par CMS et le besoin de capacité de calcul et de stockage afin de traiter ces données, s'ajoute la nécessité d'en contrôler la qualité. Dans ce but j'ai élaboré, pour le cas particulier des électrons, un programme d'analyse spécifique au contrôle qualité hors ligne, le *DQMOOffline*. Le squelette de ce programme repose sur des boucles d'itérations sur des collections d'objets reconstruits, dans notre cas ceux participant à la reconstruction de l'électron et introduits dans le chapitre 3, à savoir :

- les électrons
- les super-agrégats
- les listes de traces reconstruites (traces générales, traces GSF)

afin de remplir en sortie des histogrammes pour chacune des variables permettant d'évaluer la qualité de ces objets reconstruits et qui dans le cas des électrons ont été définies dans

la partie 3.4.2. Notons que pour les objets intermédiaires que sont les traces ou les super-agrégats, des variables spécifiques sont considérées telles que le nombre de points d'impact et le χ^2 d'ajustement de la trace notamment.

Ensuite des sélections de plus en plus contraignantes permettent de définir différents lots de contrôle à la pureté de plus en plus grande :

- sélection 1 : tous les électrons reconstruits
- sélection 2 : électrons reconstruits avec $E_T > 10$ GeV
- sélection 3 : électrons reconstruits avec $E_T > 10$ GeV et coupure sur l'isolation trace $TkIso < 1$ GeV.
- sélection 4 : électrons *passing probes* issus d'une sélection tag-and-probe avec pour tag, un électron reconstruit $E_T > 10$ GeV et $TkIso < 1$ et pour probe un super-agrégat. Aucun critère de masse n'est en revanche appliqué : toutes les paires formées sont considérées.

Ce programme est alors intégré au *DQMOffline*, dont l'interface web permet de rendre compte des résultats en vue d'une certification. En effet, comme le montre la figure 4.1 dans le cas particulier de la sélection 4 sur des données recueillies en 2012, l'utilisateur peut directement visualiser les histogrammes de contrôle et diverses options s'offrent à lui comme la sélection de numéros de runs et d'événements précis, la superposition des histogrammes ou encore le zoom en vue d'en faciliter l'examen. Il peut ainsi procéder à la vérification de la qualité des électrons dans les données de CMS.

4.1.2 Mesure des distributions des électrons et comparaison avec le Monte-Carlo, données 2011

Dans le but de valider les algorithmes de reconstruction des électrons, j'ai développé un autre programme d'analyse reposant sur la technique *tag-and-probe* sélectionnant des événements $Z \rightarrow e^+e^-$ afin de comparer les données recueillies et reconstruites à l'attendu Monte-Carlo, signal et bruits de fond. Il s'agit ici non pas de procéder à une vérification grossière du fait des contraintes du DQMOffline, comme cela est fait avec l'outil précédent mais plutôt d'élaborer une sélection très pure afin d'avoir l'assurance de contrôler les caractéristiques de vrais électrons. L'échantillon d'électrons déterminé, je procède ensuite au contrôle qualité en comparant les distributions des variables électrons, introduites dans le chapitre 3, avec l'attendu Monte-Carlo.

La méthode tag-and-probe appliquée ici considère pour former le tag, un électron vérifiant les critères du point de fonctionnement WP60 (cf. table 4.1). Il s'agit d'un des points de fonctionnement ayant été déterminés et optimisés par la collaboration CMS pour la sélection des électrons dans les analyses W et Z [26]. Le probe est quant à lui défini comme un candidat électron vérifiant $p_T^e > 10$ GeV. c^{-1} et une coupure sur l'isolation relative trace (< 0.09 dans la partie tonneau et < 0.04 dans les parties bouchons).

Les variables électrons sont présentées dans les figures 4.2 à 4.6. Les distributions y figurent séparément pour les électrons situés dans le tonneau (EB) et dans les bouchons (EE) du calorimètre. La comparaison est effectuée entre données, avec une statistique de 240 pb $^{-1}$ correspondant à la première re-reconstruction effectuée en 2011, et échantillons Monte-Carlo (signal et bruits de fond).

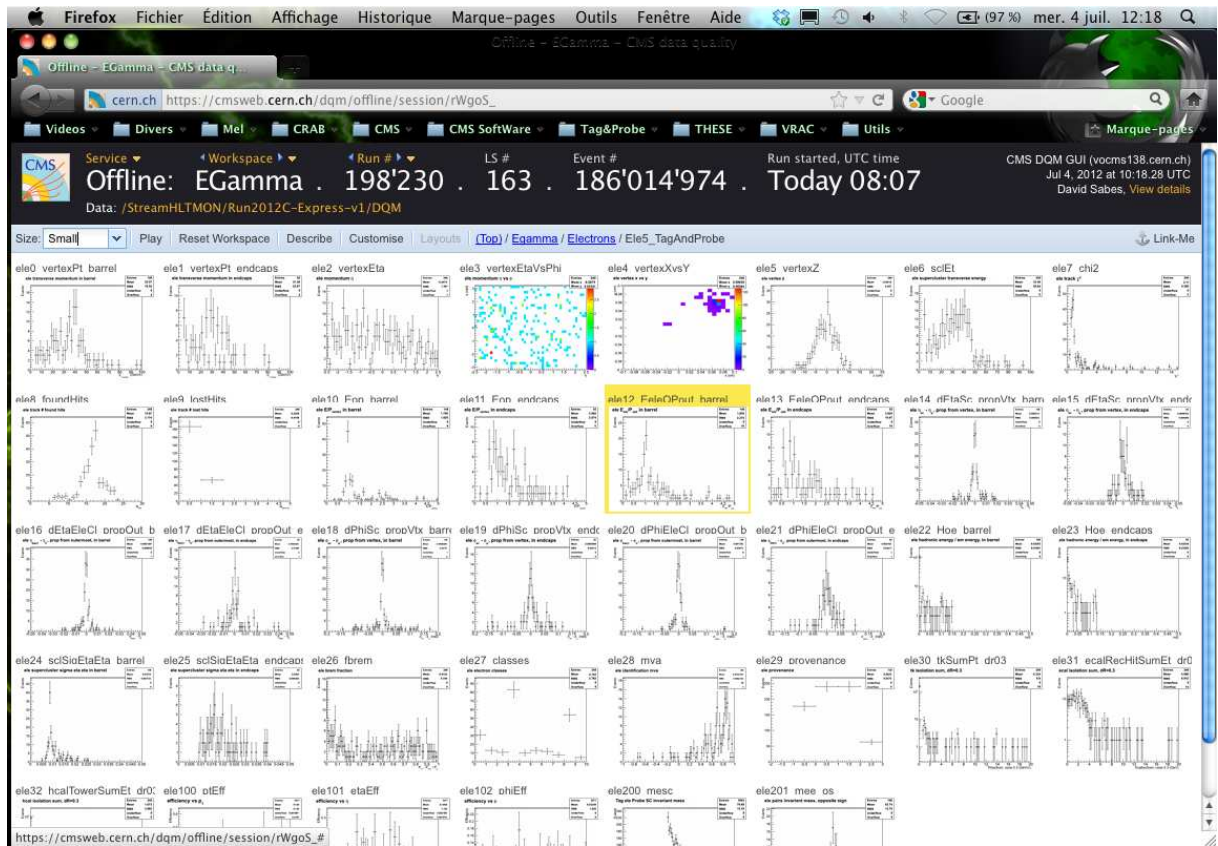


Figure 4.1 – Capture d’écran de l’interface web du contrôle de qualité hors champ dans le cas de la sélection tag-and-probe.

Les distributions montrent en général un bon accord entre les données et la simulation. Notons que le bruit de fond est négligeable avec la sélection utilisée. La figure 4.2 présente un bon accord même si une différence due à une calibration imparfaite dans ces premières reconstructions est visible au niveau du pic de la distribution en particulier pour les parties bouchons. La figure 4.4 illustre un problème d’alignement, qui sera corrigé par la suite, visible à une distribution décalée dans les données.

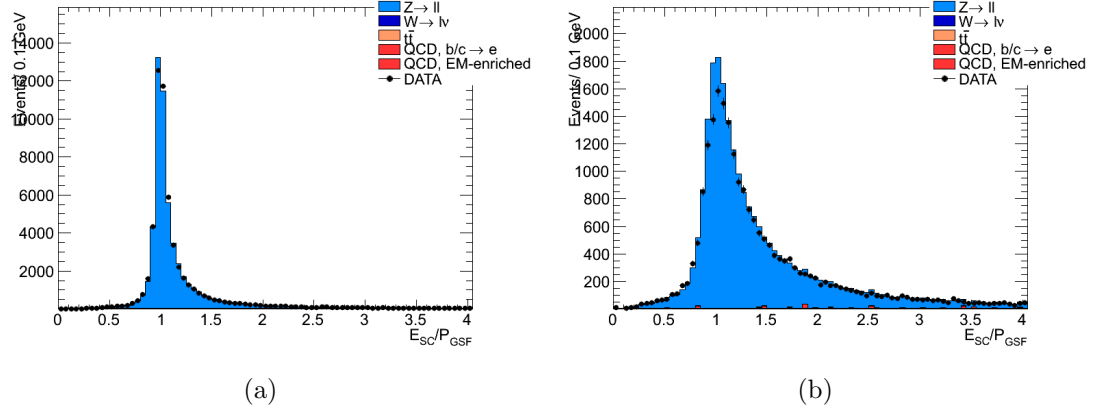


Figure 4.2 – Distributions de la variable E/p pour un électron dans le tonneau (a) ou les bouchons (b).

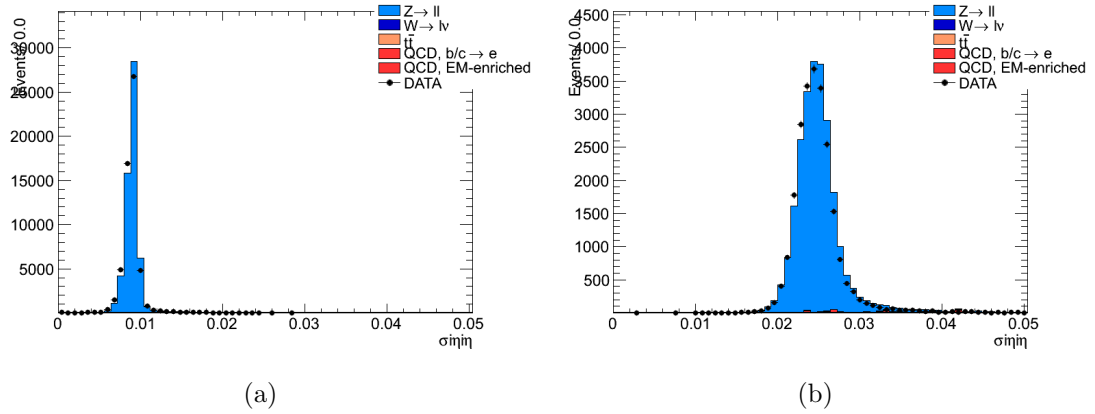
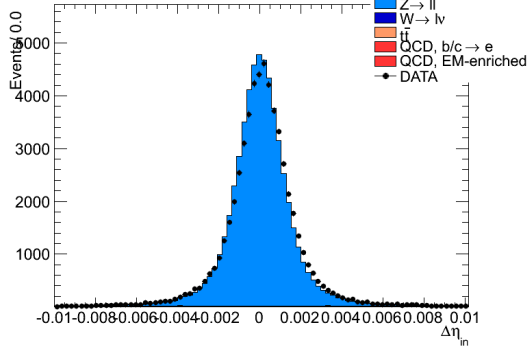
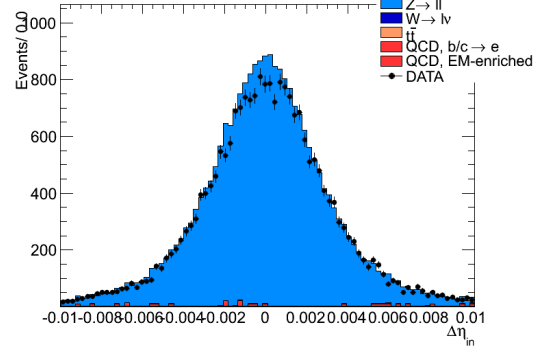


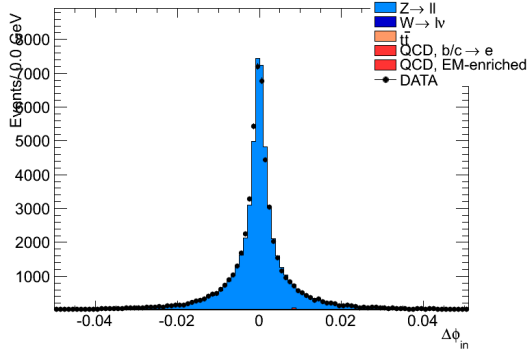
Figure 4.3 – Distributions de la variable $\sigma_{i\eta i\eta}$ pour un électron dans le tonneau (a) ou les bouchons (b).



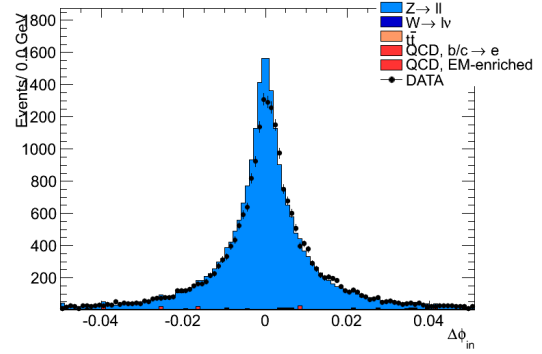
(a) $\Delta\eta$ (EB)



(b) $\Delta\eta$ (EE)

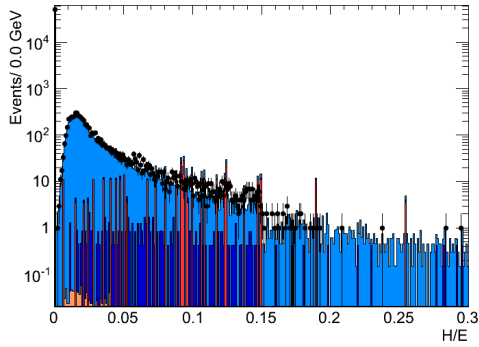


(c) $\Delta\phi$ (EB)

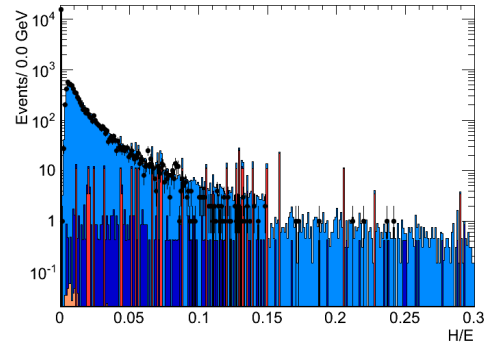


(d) $\Delta\phi$ (EE)

Figure 4.4 – Distributions des variables $\Delta\eta$ et $\Delta\phi$ pour un électron dans le tonneau (a,c) ou les bouchons (b,d).

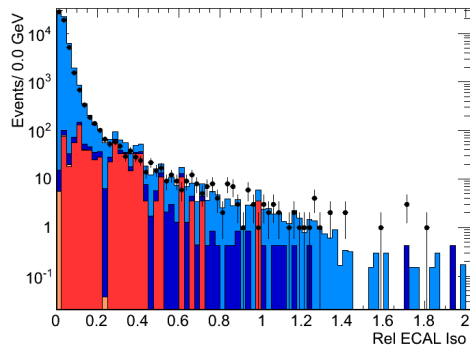


(a)

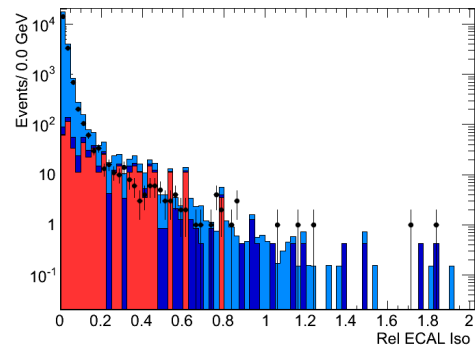


(b)

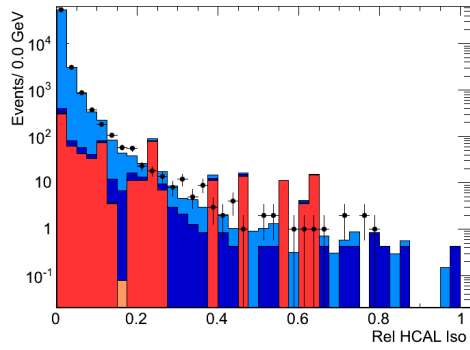
Figure 4.5 – Distributions de la variable H/E pour un électron dans le tonneau (a) ou les bouchons (b).



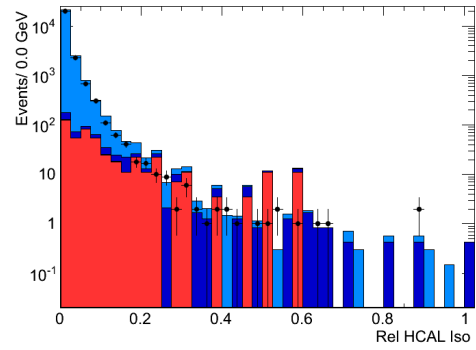
(a) ecalIso (EB)



(b) ecalIso (EE)



(c) hcalIso (EB)



(d) hcalIso (EE)

Figure 4.6 — *Distributions des différentes variables d'isolation relatives ecal et hcal pour un électron dans le tonneau (a,c) ou les bouchons (b,d).*

4.2 Sélection des événements et effet sur la mesure d'efficacité

Nous allons présenter dans cette section une étude Monte-Carlo de la sélection tag-and-probe et des effets des coupures sur la mesure. La méthode tag-and-probe appliquée à la mesure de l'efficacité de reconstruction considère pour former le tag, un électron vérifiant les critères de sélection d'un des points de fonctionnement suivants : WP95, WP80, WP60 (cf. table 4.1).

Le probe est un super-agrégat ; le passing probe, un super-agrégat associé par référence

	WP95		WP80		WP60	
	EB	EE	EB	EE	EB	EE
$\sigma_{i\eta i\eta}$	<0.01	<0.03	<0.01	<0.03	<0.01	<0.03
$ \Delta\eta $	<0.007	<0.01	<0.004	<0.007	<0.004	<0.005
$ \Delta\phi $	<0.8	<0.7	<0.06	<0.03	<0.025	<0.02
H/E	<0.15	<0.07	<0.04	<0.025	<0.025	< 0.025
trackIso	<0.15	<0.08	<0.09	<0.04	<0.04	<0.025
ecalIso	<2	<0.06	<0.07	<0.05	<0.04	<0.02
hcalIso	<0.12	<0.05	<0.10	<0.025	<0.03	<0.02

Table 4.1 – Coupures pour les différents points de fonctionnement servant à la sélection des électrons.

à un électron reconstruit. Les événements considérés vérifient en outre :

- $E_T > 20$ GeV et $|\eta| < 2.5$
- une énergie transverse manquante $TCmet$ inférieure à 15 GeV.
- Aucun jet avec $E_T > 10$ GeV dans l'événement.

Nous verrons plus loin que ces critères de sélection évolueront lorsque nous procéderons à la mesure sur les données 2011, du fait des nouvelles conditions du système de déclenchement, d'empilement...

De plus des coupures additionnelles sur le probe, en vue de diminuer le bruit de fond, ont été étudiées :

- $H/E < 0.15$
- $H/E < 0.15$ et l'isolation WP80
- $H/E < 0.15$, isolation WP80 et $\sigma_{i\eta i\eta}$ WP80.

En examinant la figure 4.7 présentant les effets sur le spectre de masse des différentes sélections appliquées, que ce soit au niveau du tag ou au niveau du probe, nous pouvons remarquer la réduction du bruit de fond apportée en considérant des critères plus stricts sur le tag, WP60 par exemple, en particulier lorsque le probe n'est lui soumis à aucune ou seulement à une faible sélection : cas sans coupure (no cut) ou H/E . Au contraire si nous faisons un choix moins sévère du point de fonctionnement WP95 sur le tag, les jeux de coupures additionnelles sur le probe, H/E + Isolation ou H/E + Isolation + $\sigma_{i\eta i\eta}$, se révèlent indispensables pour réduire le bruit de fond. Notons qu'en général l'apport de la coupure sur la variable $\sigma_{i\eta i\eta}$ quant à la réduction du bruit de fond n'est que mineur.

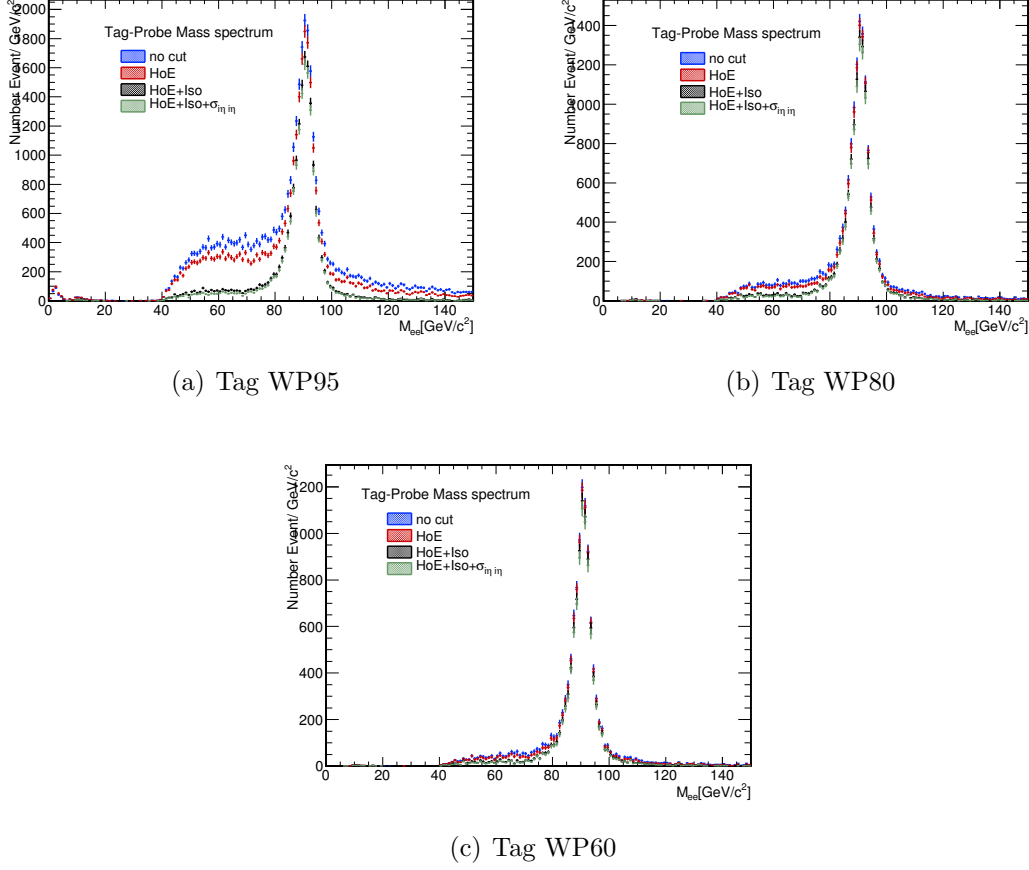


Figure 4.7 – Spectre de masse attendu pour différentes coupures appliquées sur le tag et le probe.

Sachant que nous devons appliquer le moins de sélection possible sur le probe afin d'éviter de biaiser la mesure, nous ne considérons plus que les points de fonctionnement WP80 et WP60 pour l'identification du tag. Evaluons donc, en chiffre, l'augmentation de la pureté attendue pour ces différents points de fonctionnement et autres coupures et que nous illustrons avec la figure 4.7. Plusieurs fenêtres de masse ont été étudiées : $60 < m_{e,SC} < 120$ $\text{GeV}\cdot\text{c}^{-2}$, $75 < m_{e,SC} < 105$ $\text{GeV}\cdot\text{c}^{-2}$ et $80 < m_{e,SC} < 100$ $\text{GeV}\cdot\text{c}^{-2}$.

Les résultats sont regroupés dans la table 4.2. Avec la table 4.2, nous remarquons que nous obtenons une très grande pureté ($>99\%$) avec des critères très stricts à la fois sur le tag (WP60) et sur le probe ($H/E + Iso + \sigma_{\eta\eta\eta\eta}$) dans le cas de la fenêtre $60 < m_{e,SC} < 120$ $\text{GeV}\cdot\text{c}^{-2}$. Pour des fenêtres plus étroites, $80 < m_{e,SC} < 100$ $\text{GeV}\cdot\text{c}^{-2}$, nous pouvons être moins contraignant sur les critères de sélection (tag WP80 et probe $H/E + Iso$) et garantir un niveau de pureté équivalent. Si nous ne requérons aucune sélection sur le probe, une pureté de 96.07% peut tout de même être atteinte avec un tag WP60 et une fenêtre de masse $80 < m_{e,SC} < 100$ $\text{GeV}\cdot\text{c}^{-2}$.

Étudions maintenant les efficacités de reconstruction attendues sur des échantillons de données simulées Monte-Carlo lorsque nous appliquons ces différentes sélections.

	60 - 120 GeV.c ⁻²		75 - 105 GeV.c ⁻²		80 - 100 GeV.c ⁻²	
	Tag WP80	Tag WP60	Tag WP80	Tag WP60	Tag WP80	Tag WP60
aucune coupure	79.28	89.88	87.68	94.35	90.92	96.07
H/E	85.77	93.35	92.13	96.46	94.26	97.41
H/E+Iso	97.59	98.80	98.91	99.32	99.11	99.40
H/E+Iso+ $\sigma_{i\eta i\eta}$	98.50	99.25	99.27	99.56	99.44	99.63

Table 4.2 – *Pureté (%) pour différentes sélections du tag et du probe ainsi que différentes fenêtres de masse.*

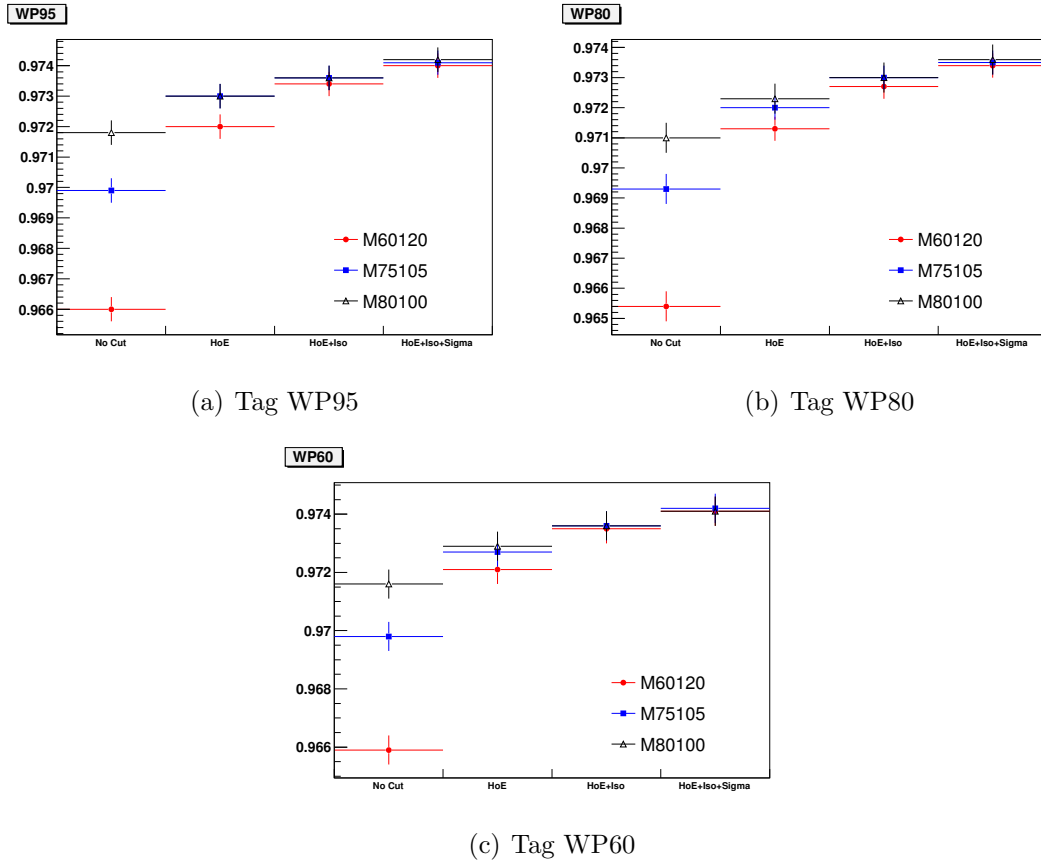


Figure 4.8 – *Effet des différentes sélections sur le tag et le probe ainsi que celui des fenêtres de masse sur l'efficacité de reconstruction Monte-Carlo (1=100%).*

Ces efficacités Monte-Carlo sont mesurées en appliquant la technique tag-and-probe tout en vérifiant que les électrons et les super-agrégats soient bien associés aux électrons de la vérité Monte-Carlo.

Les résultats illustrés par la figure 4.8 montrent que les différents points de fonctionnement utilisés pour définir le tag n'ont quasiment aucune incidence sur l'efficacité de reconstruction attendue. A contrario les sélections du probe affectent les efficacités, en particulier pour les cas les moins contraignants et la fenêtre de masse la plus large. No-

tons cependant que la variation induite reste inférieure à 1% d'après le Monte-Carlo. Ce dernier aspect s'explique par le fait que les fenêtres de masse plus strictes suppriment les candidats électrons ayant leur super-agrégat mal reconstruit, événements appartenant à la queue inférieure du spectre de masse et ayant le plus de risque d'échouer la reconstruction. Dans les cas $H/E + Iso$ ou $H/E + Iso + \sigma_{\eta\eta}$, leurs effets et ceux de la fenêtre de masse ne sont que minimales ($<0.1\%$).

Nous obtenons finalement, via tag-and-probe, une efficacité de reconstruction attendue par Monte-Carlo de $96.59 \pm 0.05\%$ pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$, dans le cas où nous n'appliquons aucune sélection sur le probe et quasiment indépendamment de la sélection sur le Tag. En contraignant la fenêtre de masse ($80 < m_{e,SC} < 100 \text{ GeV.c}^{-2}$), l'efficacité de reconstruction est de $97.16 \pm 0.05\%$.

Observons enfin l'effet de la sélection de l'événement sur l'efficacité de reconstruction, sélection caractérisée par :

- une énergie transverse manquante, ou Met (missing E_T) inférieure à 15 GeV
- Aucun jet avec $E_T > 10 \text{ GeV}$ dans l'événement.

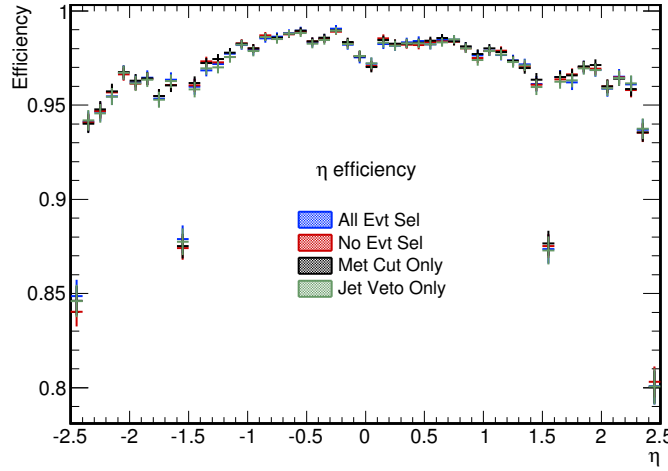


Figure 4.9 – *Effet de divers jeux de sélection des événements (aucune contrainte (rouge), coupure sur l'énergie transverse manquante seulement (noir), application du veto sur les jets seul (vert), coupure sur la MET et veto jets (bleu)) sur l'efficacité MC de reconstruction attendue en fonction de η pour une sélection tag WP60, aucune sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$.*

La figure 4.9 montre que les contraintes globales appliquées pour la sélection des événements sur lesquels sera ensuite mise en œuvre la technique tag-and-probe ne biaisent pas la mesure de l'efficacité de reconstruction.

4.3 Mesure de l'efficacité de reconstruction des électrons

La mesure de l'efficacité de reconstruction des électrons, que nous allons exposer maintenant, concerne dans un premier temps les données collectées par le LHC au cours de l'année 2010, à une énergie dans le centre de masse $\sqrt{s_{pp}} = 7$ TeV et correspondant à une luminosité intégrée $\mathcal{L}_{\text{int}} \sim 36 \text{ pb}^{-1}$. Ensuite nous présenterons les mêmes mesures faites sur une partie des données 2011 ($\mathcal{L}_{\text{int}} \sim 240 \text{ pb}^{-1}$).

4.3.1 La méthode "tag-and-probe" et mesure d'efficacité par simple comptage

La technique tag-and-probe utilise une résonance connue dans un spectre de masse donné, dans notre cas la résonance du boson Z dans le spectre de masse e^+e^- située à $91.2 \text{ GeV}.c^{-2}$, afin de sélectionner un lot pur d'événements et sonder l'efficacité d'un critère particulier en utilisant la contrainte de masse. Le marqueur, ou "tag", est un électron vérifiant des contraintes de sélection très strictes, d'identification et d'isolation par exemple, tandis que la sonde, ou "probe", est sélectionnée uniquement en formant avec le tag une paire donnant une masse invariante proche de la résonance. Nous mesurons alors l'efficacité d'un critère en comptant le nombre de probes passant ce critère :

$$Eff = \frac{N_{pp}}{N_{tot}} = \frac{N_{pp}}{N_{pp} + N_{fp}} \quad (4.1)$$

avec

- $N_{tot} = N_{pp} + N_{fp}$, le nombre total de probes
- N_{pp} , le nombre de probes passant le critère (les "passing probes")
- N_{fp} , le nombre de probes qui échouent (les "failing probes")

Appliquée à la mesure de l'efficacité de reconstruction cette technique prend pour tag, un électron vérifiant des critères de sélection très sévères qui en assurent la pureté. Le probe est un super-agrégat. Le critère dont nous voulons mesurer l'efficacité, i.e. la reconstruction, se mesure en comptant la fraction des cas pour lesquels le probe super-agrégat est celui d'un électron reconstruit.

En n'appliquant aucune sélection sur le probe, les paires tag-probe sont potentiellement contaminées par du bruit de fond. Une technique d'ajustement modélisant les distributions pour le signal et le bruit de fond peut être utilisée afin d'en extraire la composante lors de la mesure de l'efficacité. Dans certains cas la présence de bruit de fond se révèle trop importante et nécessite l'application de coupures sur le probe ce qui limite la mesure : nous ne mesurons alors l'efficacité dans les données que sur une sous-partie.

Nous allons par la suite présenter la mesure de l'efficacité de reconstruction des électrons obtenue par simple comptage des quantités issues de la sélection tag-and-probe : nous recenserons les quantités N_{pp} , N_{fp} et N_{tot} , définies dans la formule 4.1, parmi les événements vérifiant la sélection établie dans la partie 4.2. Aucune soustraction du bruit de fond ne sera effectuée. Notons également que cette mesure par simple comptage nécessite d'adapter l'estimation des incertitudes statistiques.

4.3.1.1 Efficacité et calcul des incertitudes statistiques

Notons que les populations N_{pp} et N_{tot} sont fortement corrélées, N_{pp} étant une sous-partie de N_{tot} . L'erreur sur l'efficacité peut être estimée en traitant le nombre de *passing probes* comme une variable à distribution binomiale. Ce qui nous donne :

$$d\epsilon^2 = \frac{\epsilon(1 - \epsilon)}{N_{tot}} \quad (4.2)$$

Une autre manière d'évaluer l'incertitude sur l'efficacité est d'utiliser la formule de propagation des erreurs sur le rapport :

$$\epsilon = \frac{N_{pp}}{N_{pp} + N_{fp}}. \quad (4.3)$$

pour lequel N_{pp} et N_{fp} sont indépendants.

Nous obtenons alors la relation :

$$d\epsilon^2 = \frac{(1 - \epsilon)^2 dN_p^2 + \epsilon^2 dN_{fp}^2}{(N_{pp} + N_{fp})^2} \quad (4.4)$$

qui pour des erreurs Poissonniennes redonne la formule 4.2.

Notons que dans le cas spécifique des efficacités mesurées par la technique "cut-and-count", les erreurs données par la formule binomiale ne sont pas appropriées aux cas limites. Comme l'efficacité de reconstruction est proche de 100%, il est préférable d'utiliser l'intervalle de confiance Clopper-Pearson ([8] chapitre 33) fixant les limites inférieure et supérieure :

$$d\epsilon_{inf} = \frac{N_{pp} F_F^{-1}[\alpha_{inf}; 2N_{pp}, 2(N_{tot} - N_{pp} + 1)]}{N_{tot} - N_{pp} + 1 + N_{pp} F_F^{-1}[\alpha_{inf}; 2N_{pp}, 2(N_{tot} - N_{pp} + 1)]} \quad (4.5)$$

$$d\epsilon_{sup} = \frac{(N_{pp} + 1) F_F^{-1}[1 - \alpha_{sup}; 2(N_{pp} + 1), 2(N_{tot} - N_{pp})]}{N - N_{pp} + (N_{pp} + 1) F_F^{-1}[(1 - \alpha_{sup}); 2(N_{pp} + 1), 2(-N_{tot} N_{pp})]} \quad (4.6)$$

Avec :

- $\alpha_{inf} = \alpha_{sup} = 0.1585$. Valeurs correspondant à un niveau de confiance de 68% équivalent à $\pm 1\sigma$.
- F_F^{-1} , quantile de la distribution de Fisher.

4.3.1.2 Résultats

Les efficacités mesurées sur les données sont présentées par la figure 4.10 et comparées avec l'attendu MC dans les tables 4.3, 4.4 et 4.5, et ce pour différentes sélections du tag, du probe et de fenêtres de masse. Les erreurs citées sont statistiques seulement. La présence de bruit de fond altère clairement la mesure dans les cas les moins contraignants de sélection au vue de l'écart notable avec l'attendu MC. Lorsque nous n'appliquons aucune coupure sur le probe, seule la sélection la plus stricte sur le tag, WP60, et la fenêtre de masse la plus étroite, $80 < m_{e,SC} < 100 \text{ GeV.c}^{-2}$, fournissent une mesure non biaisée de l'efficacité de reconstruction.

La méthode de simple comptage sans soustraction de bruit de fond peut fournir une

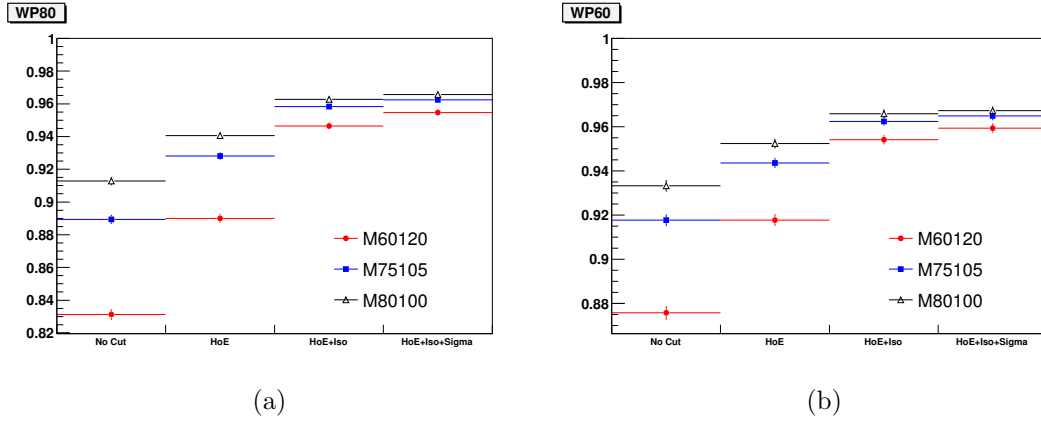


Figure 4.10 – Efficacités de reconstruction ($1=100\%$) des électrons issus des données obtenues par simple comptage pour le cas tag WP80 (a) et tag WP60 (b) et en fonction des différentes sélections sur le probe ainsi que les diverses fenêtres de masse.

mesure d'efficacité seulement dans les cas les plus stricts de sélection du tag et du probe si nous voulons avoir un résultat pour la fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$. Si au contraire nous voulons une mesure non biaisée par des coupures sur le probe, il nous faudrait considérer une fenêtre de masse très étroite avec une sélection très contraignante du tag. Cependant il est possible d'obtenir des mesures non altérées par le bruit de fond en soustrayant ce dernier estimé par ajustement du spectre de masse comme cela est détaillé dans la partie suivante.

	Tag WP80		Tag WP60	
	data	MC	data	MC
aucune coupure	$83.12^{+0.31}_{-0.31}$	$96.54^{+0.05}_{-0.05}$	$87.58^{+0.31}_{-0.32}$	$96.59^{+0.05}_{-0.05}$
H/E cut	$89.00^{+0.27}_{-0.27}$	$97.13^{+0.04}_{-0.04}$	$91.78^{+0.27}_{-0.27}$	$97.21^{+0.05}_{-0.05}$
H/E+Iso	$94.65^{+0.21}_{-0.21}$	$97.27^{+0.04}_{-0.04}$	$95.41^{+0.21}_{-0.22}$	$97.35^{+0.05}_{-0.05}$
H/E+Iso+ $\sigma_{in\eta}$	$95.47^{+0.20}_{-0.20}$	$97.34^{+0.04}_{-0.04}$	$95.94^{+0.21}_{-0.22}$	$97.41^{+0.05}_{-0.05}$

Table 4.3 – Efficacité de reconstruction (en %) obtenue par simple comptage pour différents jeux de sélections du tag et du probe et pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$. Les erreurs mentionnées ici sont statistiques.

4.3.2 Mesure de l'efficacité par une méthode d'ajustement

Une technique de soustraction du bruit de fond est de modéliser la forme du spectre de masse par une superposition de deux courbes, l'une propre au signal et l'autre modélisant le bruit de fond. Pour cela plusieurs fonctions peuvent être utilisées. Ainsi nous allons dans une première partie détailler la modélisation utilisée, ensuite nous exposerons les résultats : efficacité intégrée, fonction de η et fonction de p_T , obtenues par utilisation de fits et enfin nous présenterons l'évaluation des erreurs systématiques de ces mesures.

	Tag WP80		Tag WP60	
	data	MC	data	MC
aucune coupure	88.94 ^{+0.28} _{-0.28}	96.93 ^{+0.05} _{-0.05}	91.77 ^{+0.28} _{-0.28}	96.98 ^{+0.05} _{-0.05}
H/E	92.81 ^{+0.23} _{-0.24}	97.20 ^{+0.04} _{-0.05}	94.36 ^{+0.23} _{-0.24}	97.27 ^{+0.05} _{-0.05}
H/E+Iso	95.83 ^{+0.19} _{-0.20}	97.30 ^{+0.04} _{-0.05}	96.24 ^{+0.20} _{-0.21}	97.36 ^{+0.05} _{-0.05}
H/E+Iso+ $\sigma_{in\eta}$	96.24 ^{+0.18} _{-0.19}	97.35 ^{+0.04} _{-0.05}	96.49 ^{+0.20} _{-0.21}	97.42 ^{+0.05} _{-0.05}

Table 4.4 – Efficacité de reconstruction (en %) obtenue par simple comptage pour différents jeux de sélections du tag et du probe et pour une fenêtre de masse $75 < m_{e,SC} < 105 \text{ GeV}.c^{-2}$. Les erreurs mentionnées ici sont statistiques.

	Tag WP80		Tag WP60	
	data	MC	data	MC
aucune coupure	91.28 ^{+0.26} _{-0.27}	97.10 ^{+0.05} _{-0.05}	98.05 ^{+0.16} _{-0.17}	97.16 ^{+0.05} _{-0.05}
H/E	94.06 ^{+0.22} _{-0.23}	97.23 ^{+0.05} _{-0.05}	98.67 ^{+0.13} _{-0.15}	97.29 ^{+0.05} _{-0.05}
H/E+Iso	96.27 ^{+0.18} _{-0.19}	97.30 ^{+0.05} _{-0.05}	98.95 ^{+0.12} _{-0.14}	97.36 ^{+0.05} _{-0.05}
H/E+Iso+ $\sigma_{in\eta}$	96.56 ^{+0.18} _{-0.19}	97.36 ^{+0.05} _{-0.05}	98.96 ^{+0.12} _{-0.14}	97.41 ^{+0.05} _{-0.05}

Table 4.5 – Efficacité de reconstruction (en %) obtenue par simple comptage pour différents jeux de sélections du tag et du probe et pour une fenêtre de masse $80 < m_{e,SC} < 100 \text{ GeV}.c^{-2}$. Les erreurs mentionnées ici sont statistiques.

4.3.2.1 Fonctions de modélisation signal et bruit de fond

Plusieurs fonctions peuvent ajuster correctement la forme du spectre de masse du signal obtenu de l'échantillon MC. Usuellement la fonction Breit-Wigner convoluée à une Crystal Ball sert à modéliser ce signal. Cependant j'ai constaté qu'une modélisation basée sur une fonction Voigtienne, fonction Breit-Wigner convoluée avec une fonction gaussienne, elle-même convoluée à une Crystal Ball, permettait un meilleur ajustement. Les paramètres libres de la Voigtienne sont :

- σ , noté sv_{pp} dans les figures 4.14 à 4.16. Ce paramètre est contraint à être le même dans le cas des probes échouant le critère de reconstruction, les *failing probes* et dans le cas de ceux qui réussissent, les *passing probes*.
- la moyenne, notée mv_{pp} . De même ce paramètre est le même pour les cas *failing probes* et *passing probes*.
- la largeur, notée $width_{pp}$. Paramètre identique pour les *failing probes* et *passing probes*.

La fonction Crystal Ball est quant à elle définie par :

$$f(x; \alpha, n, \bar{x}, \sigma) = \exp\left(-\frac{(x - \bar{x})^2}{2\sigma^2}\right) \text{ pour } \frac{x - \bar{x}}{\sigma} \geq -|\alpha| \quad (4.7)$$

Et

$$f(x; \alpha, n, \bar{x}, \sigma) = A \times \left(B - \frac{(x - \bar{x})}{\sigma}\right)^{-n} \text{ pour } \frac{x - \bar{x}}{\sigma} < -|\alpha| \quad (4.8)$$

Avec

$$A = \left(\frac{n}{|\alpha|}\right)^n \times \exp(-|\alpha|^2/2) ; B = \frac{n}{|\alpha|} - |\alpha| \quad (4.9)$$

Et ses paramètres libres à ajuster sont :

- α , noté α_{pp} pour les *passing probes* et α_{fp} pour les *failing probes*.
- n , noté n_{pp} pour les *passing probes* et n_{fp} pour les *failing probes*.
- la moyenne $\bar{x}$ fixée à $\bar{x} = 0$.
- σ , noté σ_{pp} pour les *passing probes* et σ_{fp} pour les *failing probes*.

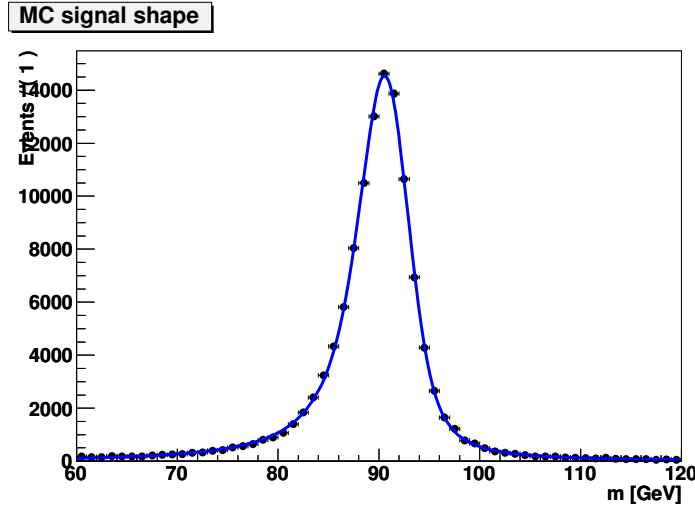


Figure 4.11 – Illustration d’un fit par Voigtienne convoluée avec une Crystal Ball de la distribution MC de la masse invariante formée par une paire tag-probe, cas tag WP60 et aucune sélection sur le probe.

Le choix de cette combinaison de fonctions s’explique par plusieurs aspects. Une résonance idéale d’un boson Z se désintégrant en paire électron-positron suit une modélisation selon une distribution de Breit-Wigner. Dans notre cas, la masse invariante n’étant pas construite par une paire e^-e^+ mais par une paire électron-super-agrégat, le pic de résonance ne bénéficie plus de la précision apportée par la reconstruction électron, notamment grâce à la combinaison E-p (cf. 3.4.2), quant à l’estimation de l’impulsion du candidat reconstruit super-agrégat. Ceci a pour conséquence d’élargir le pic de résonance et la convolution entre une Breit-Wigner et une gaussienne, c’est-à-dire une Voigtienne, peut très bien décrire cet effet. Par ailleurs certains électrons, notamment ceux très affectés par le rayonnement de freinage ou ceux situés à proximité de zones de cracks, contribuent à augmenter la queue de la distribution à basse masse : ce phénomène peut être modélisé par la convolution avec une Crystal Ball. Cette dernière combinaison de fonctions ajuste parfaitement l’allure d’une résonance caractéristique d’un boson Z dans le spectre de masse $m_{e,SC}$ comme cela peut être constaté sur la figure 4.11.

Cependant l’amplitude et l’allure du spectre de masse, mais surtout de sa queue de distribution à basse masse, varient selon les critères de sélection appliqués sur le tag et sur le probe. Elles diffèrent également entre *passing probes* et *failing probes*. Par conséquent les paramètres de la Crystal Ball seront laissés libres lors de l’ajustement sur les données.

Le bruit de fond est quant à lui modélisé par une fonction exponentielle possédant un seul paramètre : sa pente notée c_1 sur les figures 4.14 à 4.16. Notons dès à présent que dans le cas particulier de l'estimation en fonction de E_T du probe, se produit un phénomène de pallier visible sur les distributions propres aux bins hautes énergies transverses et sera modélisé en associant cette fonction exponentielle avec une fonction complémentaire à la fonction "erreur" :

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \quad (4.10)$$

Dans le but d'estimer les incertitudes systématiques relatives à la mesure de l'efficacité de reconstruction par technique d'ajustement, d'autres fonctions seront utilisées :

- Breit-Wigner convoluée avec une Crystal Ball pour la composante signal
- Une loi de puissance pour la composante bruit de fond

4.3.2.2 Résultats obtenus avec les données 2010

Roofit [27] a été ici utilisé pour ajuster simultanément les distributions issues des échantillons *passing probes* et *failing probes* en implémentant notamment l'efficacité comme paramètre d'ajustement. L'option "extended" pour tenir compte des fluctuations du nombre total d'entrées pour chaque catégorie a été activée. Enfin les erreurs sur l'efficacité sont évaluées par Minos qui tient compte des corrélations éventuelles entre paramètres et ont été validées via une comparaison avec la formule de propagation des erreurs.

La statistique considérée pour la fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$ est donnée en table 4.7. Les résultats sont regroupés dans les tables 4.6 et 4.8. Nous y observons un bon accord entre mesure et attendu MC. Bien que la plupart des valeurs s'écarte au maximum de 2σ de l'attendu, l'efficacité de reconstruction mesurée sur les données est en général inférieure à l'attendu. Dans le cas d'un tag WP60, d'aucune sélection sur le probe et d'une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$, nous obtenons une efficacité mesurée de $95.61^{+0.78}_{-0.95}\%$ à comparer avec l'attendu de $96.59^{+0.05}_{-0.05}\%$.

Comme cela a été présenté dans la partie 4.2, nous n'observons aucune dépendance significative en fonction des points de fonctionnement utilisés pour le tag ce qui nous permet de conclure que la modélisation du bruit de fond est correcte ou qu'il est trop faible pour avoir un effet, le comportement avec soustraction de ce bruit de fond restant conforme à l'attendu. Une légère augmentation de l'efficacité mesurée en fonction de critères de sélection du probe plus stricts peut être également observée, les variations restant de l'ordre de 0.5%.

Les rapports entre efficacité mesurée sur les données et attendu MC pour des fenêtres de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$ et $80 < m_{e,SC} < 100 \text{ GeV.c}^{-2}$ ainsi que diverses sélections du tag et du probe sont présentés sur la figure 4.12.

Par ailleurs, l'efficacité a également été mesurée en fonction de η , en divisant les limites d'acceptance pour la reconstruction ($0 < |\eta| < 2.5$) en cinq bins : deux dans la partie tonneau ($0 < |\eta| < 0.78$ et $0.78 < |\eta| < 1.444$), deux pour les parties bouchons ($1.566 < |\eta| < 2.0$ et $2.0 < |\eta| < 2.5$) et un dernier spécifique aux régions de transition tonneau-bouchons ($1.444 < |\eta| < 1.566$). Les résultats sont regroupés dans la table 4.9 et présentés dans la figure 4.13.

	Tag WP80		Tag WP60	
	data	MC	data	MC
aucune coupure	95.36 ^{+0.44} _{-0.44}	96.54 ^{+0.05} _{-0.05}	95.61 ^{+0.78} _{-0.95}	96.59 ^{+0.05} _{-0.05}
H/E	96.93 ^{+0.39} _{-0.49}	97.13 ^{+0.04} _{-0.04}	97.39 ^{+0.22} _{-0.23}	97.21 ^{+0.05} _{-0.05}
H/E+Iso	96.97 ^{+0.22} _{-0.22}	97.27 ^{+0.04} _{-0.04}	96.94 ^{+0.25} _{-0.27}	97.35 ^{+0.05} _{-0.05}
H/E+Iso+ $\sigma_{i\eta i\eta}$	96.98 ^{+0.17} _{-0.17}	97.34 ^{+0.04} _{-0.04}	96.94 ^{+0.23} _{-0.25}	97.41 ^{+0.05} _{-0.05}

Table 4.6 – Efficacités de reconstruction (en %) obtenues par ajustement simultané des échantillons passing probes et failing probes pour différentes sélections de tag et de probe et pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.

	Tag WP80		Tag WP60	
	passing probes	failing probes	passing probes	failing probes
aucune coupure	12237	2470	9749	1372
H/E	12237	1507	9749	870
H/E+Iso	11190	632	9118	438
H/E+Iso+ $\sigma_{i\eta i\eta}$	10803	512	8811	372

Table 4.7 – Nombre de passing probes et failing probes dans les données pour différentes sélections de tag et de probe et pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.

	Tag WP80		Tag WP60	
	data	MC	data	MC
aucune coupure	96.33 ^{+0.39} _{-0.42}	97.10 ^{+0.05} _{-0.05}	96.43 ^{+0.36} _{-0.36}	97.16 ^{+0.05} _{-0.05}
H/E	96.95 ^{+0.29} _{-0.31}	97.23 ^{+0.05} _{-0.05}	97.01 ^{+0.27} _{-0.28}	97.29 ^{+0.05} _{-0.05}
H/E+Iso	96.93 ^{+0.21} _{-0.21}	97.30 ^{+0.05} _{-0.05}	96.97 ^{+0.22} _{-0.23}	97.36 ^{+0.05} _{-0.05}
H/E+Iso+ $\sigma_{i\eta i\eta}$	97.08 ^{+0.20} _{-0.21}	97.36 ^{+0.05} _{-0.05}	96.95 ^{+0.22} _{-0.23}	97.41 ^{+0.05} _{-0.05}

Table 4.8 – Efficacités de reconstruction (en %) obtenues par ajustement simultané des échantillons passing probes et failing probes pour différentes sélections de tag et de probe et pour une fenêtre de masse $80 < m_{e,SC} < 100 \text{ GeV}.c^{-2}$.

	data	pp	fp	MC
$0 < \eta < 0.78$	97.94 ^{+0.42} _{-0.42}	3898	453	98.37 ^{+0.06} _{-0.07}
$0.78 < \eta < 1.444$	96.0 ^{+1.3} _{-1.3}	2875	369	97.62 ^{+0.08} _{-0.08}
$1.444 < \eta < 1.566$	85.2 ^{+2.9} _{-2.9}	232	51	88.17 ^{+0.51} _{-0.53}
$1.566 < \eta < 2.0$	96.8 ^{+1.1} _{-1.1}	1437	226	96.27 ^{+0.13} _{-0.14}
$2.0 < \eta < 2.5$	91.8 ^{+1.6} _{-2.2}	1307	273	93.17 ^{+0.18} _{-0.19}

Table 4.9 – Efficacités de reconstruction (en %) en fonction de η obtenues par ajustement simultané des échantillons passing probes et failing probes pour un tag WP60, aucune sélection sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Les nombres de passing probes (pp) et failing probes (fp) sont également cités.

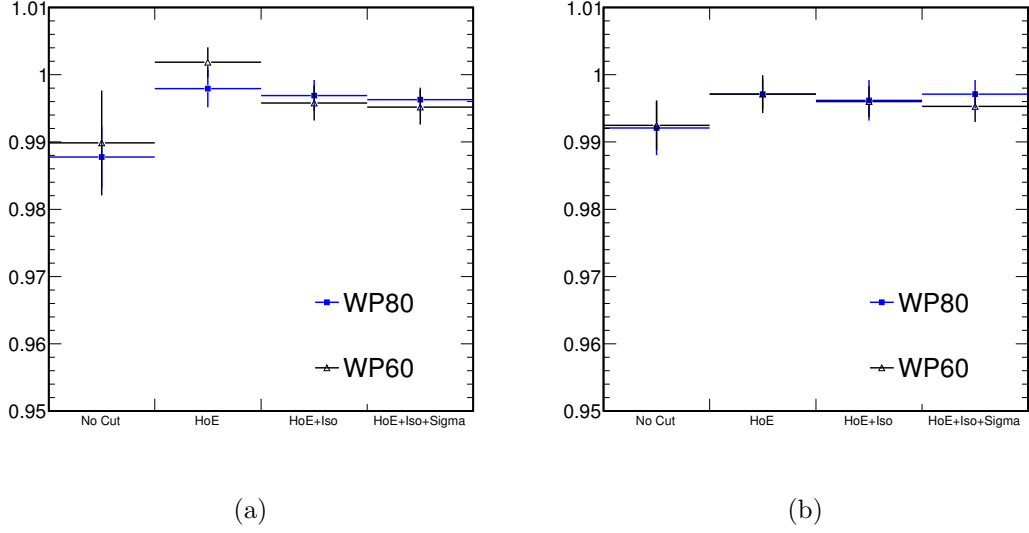


Figure 4.12 – Rapport des efficacités de reconstruction des électrons mesurées et attendues pour le cas $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$ (a) et $80 < m_{e,SC} < 100 \text{ GeV.c}^{-2}$ (b), avec tag WP80 (bleu) et tag WP60 (noir) et en fonction des différentes sélections sur le probe.

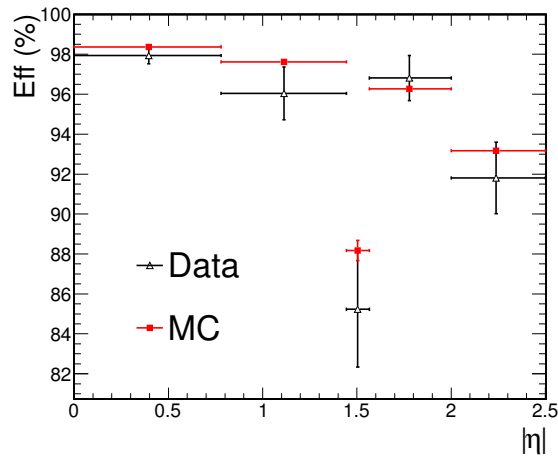


Figure 4.13 – Comparaison des efficacités mesurées (noir) et attendues (rouge) en fonction de η et obtenues pour une sélection tag WP60, aucune coupure sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV.c}^{-2}$.

4.3.2.3 Résultats obtenus avec les données 2011

J'ai également procédé à la mesure de l'efficacité de reconstruction des électrons pour une partie des données 2011 ($\mathcal{L}_{\text{int}} \sim 240 \text{ pb}^{-1}$). Cependant comme les coupures imposées par le système de déclenchement ont changé ainsi que les conditions d'empilement, nous avons dû adapter notre sélection des événements :

- énergie transverse manquante $< 25 \text{ GeV}$.
- Tag : électron $E_T > 20 \text{ GeV}$; $|\eta| < 2.5$; WP60
- Probe :
 - $E_T > 20 \text{ GeV}$ dans le cas de l'efficacité fonction de η ; $E_T > 10 \text{ GeV}$ dans le cas de l'efficacité fonction de E_T .
 - $|\eta| < 2.5$
 - $H/E < 0.15$ dans la partie tonneau ; $H/E < 0.1$ dans les parties bouchons

Les résultats en fonction de η sont regroupés dans la table 4.10. Nous pouvons remarquer une large différence avec les mesures 2010 pour l'efficacité dans les zones de transition tonneau-bouchons. Cela s'explique par le fait que nous appliquons maintenant une coupure H/E sur notre probe.

Les efficacités en fonction de E_T sont regroupées dans la table 4.11.

	data	pp	fp	MC
$0 < \eta < 0.78$	$97.94^{+0.11}_{-0.11}$	35316	6867	$98.08^{+0.03}_{-0.04}$
$0.78 < \eta < 1.444$	$97.35^{+0.17}_{-0.19}$	26142	4685	$97.60^{+0.04}_{-0.04}$
$1.444 < \eta < 1.566$	$94.18^{+0.57}_{-0.57}$	2319	339	$95.74^{+0.17}_{-0.18}$
$1.566 < \eta < 2.0$	$95.24^{+0.25}_{-0.25}$	12781	2204	$96.00^{+0.07}_{-0.07}$
$2.0 < \eta < 2.5$	$93.66^{+0.31}_{-0.32}$	10948	1906	$93.48^{+0.1}_{-0.1}$

Table 4.10 – Efficacités de reconstruction (en %) en fonction de η obtenues par ajustement simultané des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV} \cdot c^{-2}$. Les nombres de passing probes (pp) et failing probes (fp) sont également cités. Données 2011.

	data	pp	fp	MC
$10 < E_T < 15$	$92.2^{+2.2}_{-2.2}$	2533	5451	$87.81^{+0.45}_{-0.45}$
$15 < E_T < 20$	$94.3^{+2.0}_{-2.0}$	3262	3187	$91.40^{+0.25}_{-0.25}$
$20 < E_T < 30$	$93.37^{+0.79}_{-0.79}$	13194	3723	$94.87^{+0.08}_{-0.08}$
$30 < E_T < 40$	$96.25^{+0.33}_{-0.24}$	27743	2114	$96.50^{+0.04}_{-0.04}$
$40 < E_T < 50$	$97.32^{+0.11}_{-0.11}$	31177	1191	$97.23^{+0.04}_{-0.04}$
$E_T > 50$	$97.96^{+0.18}_{-0.18}$	9597	335	$97.83^{+0.06}_{-0.06}$

Table 4.11 – Efficacités de reconstruction (en %) en fonction de E_T obtenues par ajustement simultané des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV} \cdot c^{-2}$. Les nombres de passing probes (pp) et failing probes (fp) sont également cités. Données 2011.

4.3.2.4 Estimation des incertitudes systématiques

La principale source d'incertitude systématique provient de la modélisation du signal et surtout de la modélisation du bruit de fond. La table 4.12 présente une comparaison entre différentes combinaisons de fonctions modélisant le signal (Voigtienne convoluée avec Crystal Ball ou Breit-Wigner convoluée avec Crystal Ball) et le bruit de fond (exponentielle ou loi de puissance) et ce pour un tag WP60 et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Ces résultats sont illustrés par les figures 4.15 et 4.16.

Nous avons également exploré l'option de fixer certains paramètres de la Voigtienne

	LorCB - Exp	VoigtCB - Puiss	VoigtCB-Exp
aucune coupure	$95.57^{+0.82}_{-1.03}$	$95.06^{+0.78}_{-0.90}$	$95.61^{+0.78}_{-0.95}$
H/E	$97.16^{+0.26}_{-0.27}$	$96.57^{+0.43}_{-0.53}$	$97.39^{+0.22}_{-0.23}$
H/E+Iso	$96.94^{+0.24}_{-0.27}$	$96.85^{+0.25}_{-0.28}$	$96.94^{+0.25}_{-0.27}$
H/E+Iso+ $\sigma_{i\eta i\eta}$	$97.30^{+0.21}_{-0.21}$	$96.85^{+0.24}_{-0.27}$	$96.94^{+0.23}_{-0.25}$

Table 4.12 – Efficacités de reconstruction (en %) obtenues par ajustement simultané de différentes allures (signal et bruit de fond) des échantillons passing probes et failing probes pour un tag WP60, aucune sélection sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.

et de la Breit-Wigner afin de réduire le nombre de paramètres libres. Les résultats sont réunis dans la table 4.13. La Voigtienne avec ses paramètres fixés aux valeurs officielles du PDG ne donne pas un ajustement correct ni, par conséquent, une bonne mesure de l'efficacité. Au contraire, la Breit-Wigner avec ses paramètres fixés aux valeurs officielles du PDG donne un ajustement correct même si son χ^2 reste légèrement moins bon que celui obtenu par une Voigtienne aux paramètres libres qui restera donc le choix pour déterminer, avec le moins de biais possible, l'efficacité de reconstruction des électrons.

De cette étude nous estimons l'incertitude systématique par somme quadratique des

	Lor. param fixes	Voigt param fixes	Voigt $\bar{x}$ fixe	Voigt param libres
aucune coupure	$96.22^{+0.58}_{-0.69}$	$97.15^{+0.12}_{-0.12}$	$95.53^{+0.41}_{-0.41}$	$95.61^{+0.78}_{-0.95}$
H/E	$96.92^{+0.38}_{-0.52}$	$97.32^{+0.22}_{-0.22}$	$96.88^{+0.38}_{-0.53}$	$97.39^{+0.22}_{-0.23}$
H/E+Iso	$96.99^{+0.23}_{-0.26}$	$97.02^{+0.24}_{-0.27}$	$97.11^{+0.22}_{-0.23}$	$96.94^{+0.25}_{-0.27}$
H/E+Iso+ $\sigma_{i\eta i\eta}$	$96.99^{+0.22}_{-0.23}$	$97.03^{+0.22}_{-0.22}$	$96.95^{+0.23}_{-0.24}$	$96.94^{+0.23}_{-0.25}$

Table 4.13 – Efficacités de reconstruction (en %) obtenues suivant diverses conditions sur les paramètres par ajustement simultané des échantillons passing probes et failing probes pour un tag WP60, aucune sélection sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.

différences obtenues par les diverses modélisations et conditions d'ajustement utilisées en prenant comme référence la valeur donnée par la combinaison Voigtienne convoluée avec Crystal Ball + exponentielle. Cette erreur systématique, dans le cas où nous n'appliquons

aucune coupure sur le probe s'élève à 0.55% .

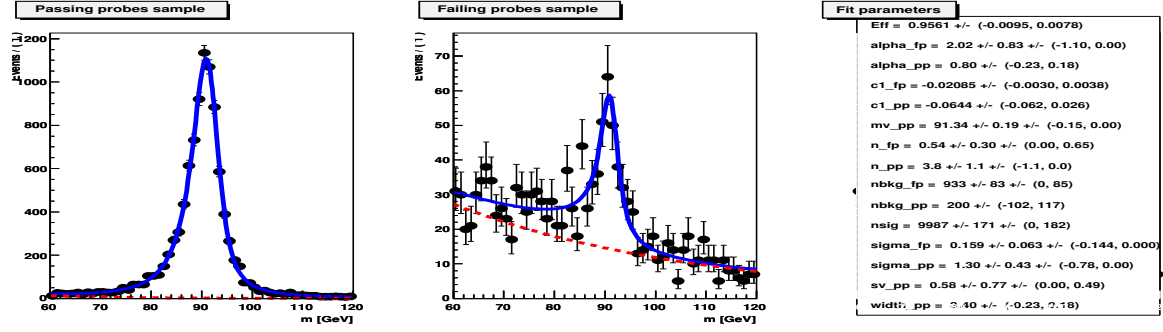
Nous avons procédé de même pour évaluer les erreurs systématiques pour les données 2011. Les résultats en fonction de η sont regroupés dans la table 4.14 et obtenus par somme quadratique. En ce qui concerne l'efficacité en fonction de E_T , la systématique se résume à la variation de forme du bruit de fond ; les résultats sont regroupés dans la table 4.15.

	VoigtCB - Err	VoigtCB - Exp	LorCB - Exp	Err syst (%)
$0 < \eta < 0.78$	$97.94^{+0.11}_{-0.11}$	$97.0^{+0.2}_{-0.2}$	$96.96^{+0.24}_{-0.25}$	± 0.98
$0.78 < \eta < 1.444$	$97.35^{+0.17}_{-0.19}$	$96.52^{+0.12}_{-0.12}$	$96.42^{+0.55}_{-0.29}$	± 1.22
$1.444 < \eta < 1.566$	$94.18^{+0.57}_{-0.57}$	$94.57^{+0.56}_{-0.56}$	$94.3^{+1.0}_{-1.0}$	± 0.12
$1.566 < \eta < 2.0$	$95.24^{+0.25}_{-0.25}$	$95.27^{+0.49}_{-0.40}$	$95.38^{+0.32}_{-0.32}$	± 0.14
$2.0 < \eta < 2.5$	$93.66^{+0.31}_{-0.32}$	$93.54^{+0.34}_{-0.33}$	$93.48^{+0.33}_{-0.44}$	± 0.14

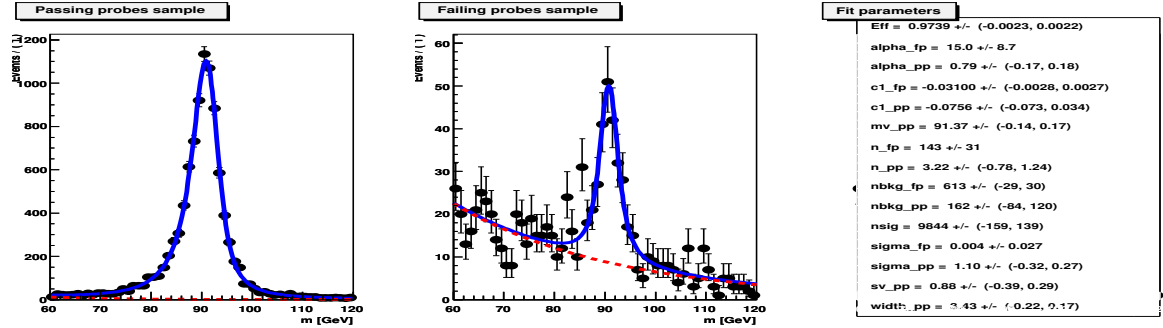
Table 4.14 – Efficacités de reconstruction (en %) en fonction de η obtenues par ajustement simultané de différentes formes (signal et bruit de fond) des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Données 2011

	VoigtCB - Exp	VoigtCB - Err	Err syst (%)
$10 < E_T < 15$	$93.2^{+2.4}_{-2.4}$	$92.2^{+2.2}_{-2.2}$	± 1.0
$15 < E_T < 20$	$92.9^{+2.4}_{-2.6}$	$94.3^{+2.0}_{-2.0}$	± 1.4
$20 < E_T < 30$	$93.29^{+0.76}_{-0.76}$	$93.37^{+0.79}_{-0.79}$	± 0.08
$30 < E_T < 40$	$95.51^{+0.26}_{-0.31}$	$96.25^{+0.33}_{-0.24}$	± 0.74
$40 < E_T < 50$	$97.14^{+0.1}_{-0.1}$	$97.32^{+0.11}_{-0.11}$	± 0.18
$E_T > 50$	$97.69^{+0.22}_{-0.28}$	$97.96^{+0.18}_{-0.18}$	± 0.27

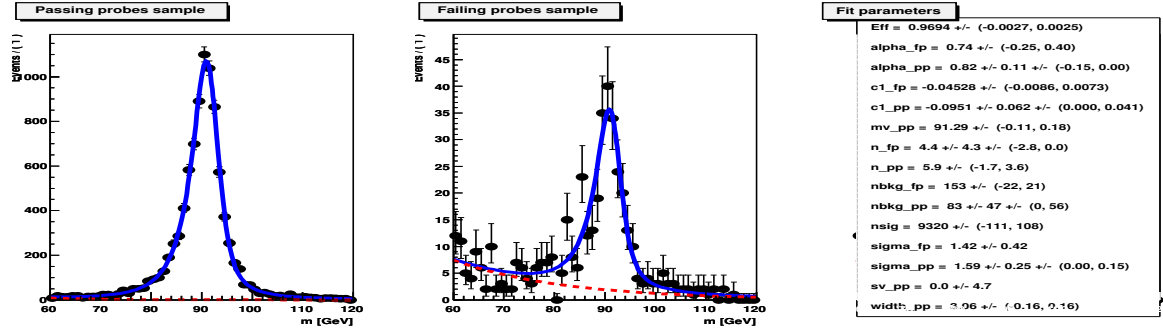
Table 4.15 – Efficacités de reconstruction (en %) obtenues par ajustement simultané de différentes formes (signal et bruit de fond) des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Données 2011



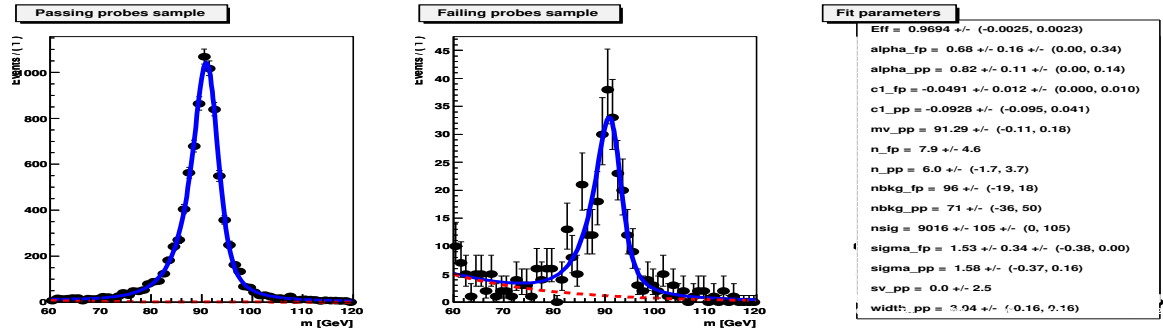
(a)



(b)

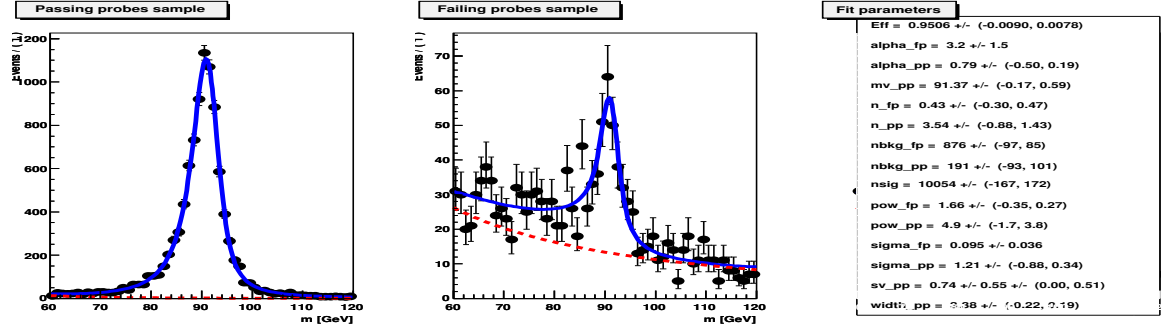


(c)

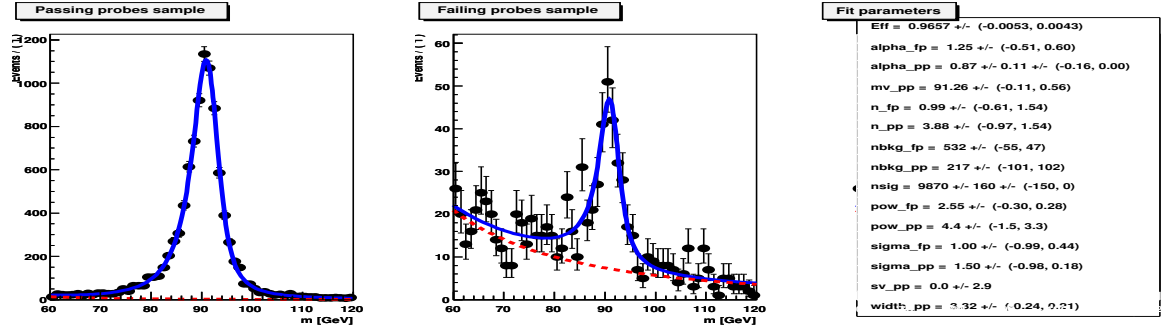


(d)

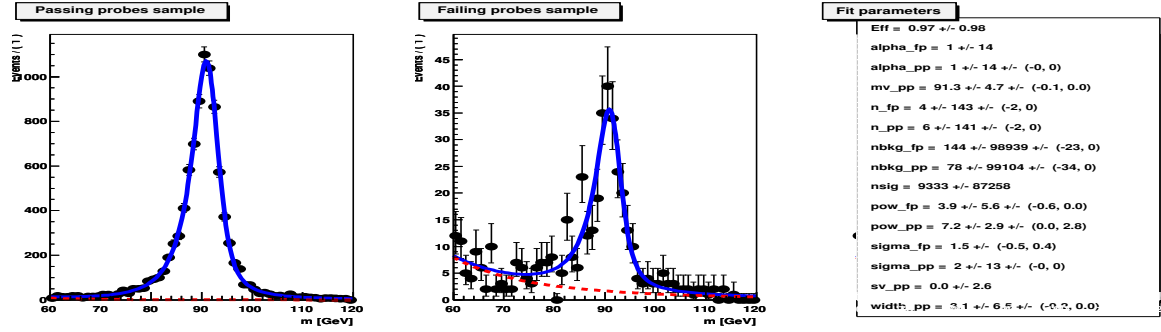
Figure 4.14 – *Fit simultané des distributions de masse invariante pour les passing probes et failing probes obtenues pour un tag WP60 et différentes sélections sur le probe : aucune coupure (a), H/E (b), $H/E+Iso$ (c) et $H/E+Iso+\sigma_{in\eta}$ (d). Le fit a été réalisé pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.*



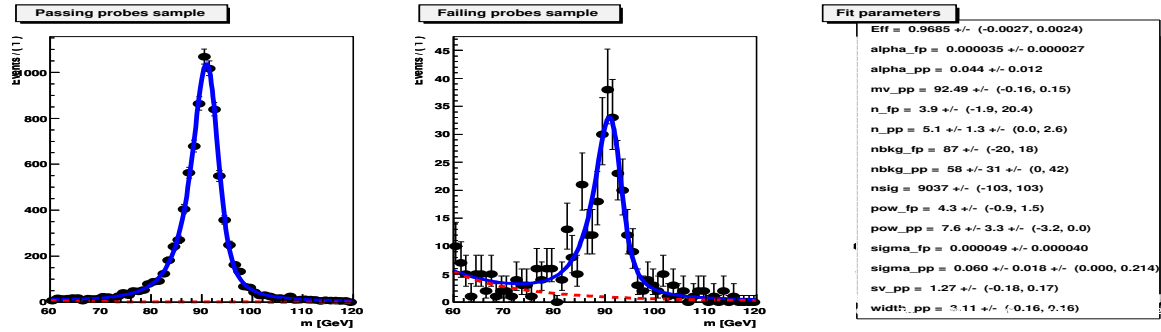
(a)



(b)



(c)



(d)

Figure 4.15 – *Signal Voigt(x)CB Bruit de fond Puiss - tag WP60 et diverses coupures sur le probe aucune coupure (a), H/E (b), H/E+Iso (c) et H/E+Iso+ σ_{inq} (d). Le fit a été réalisé pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV} \cdot c^{-2}$.*

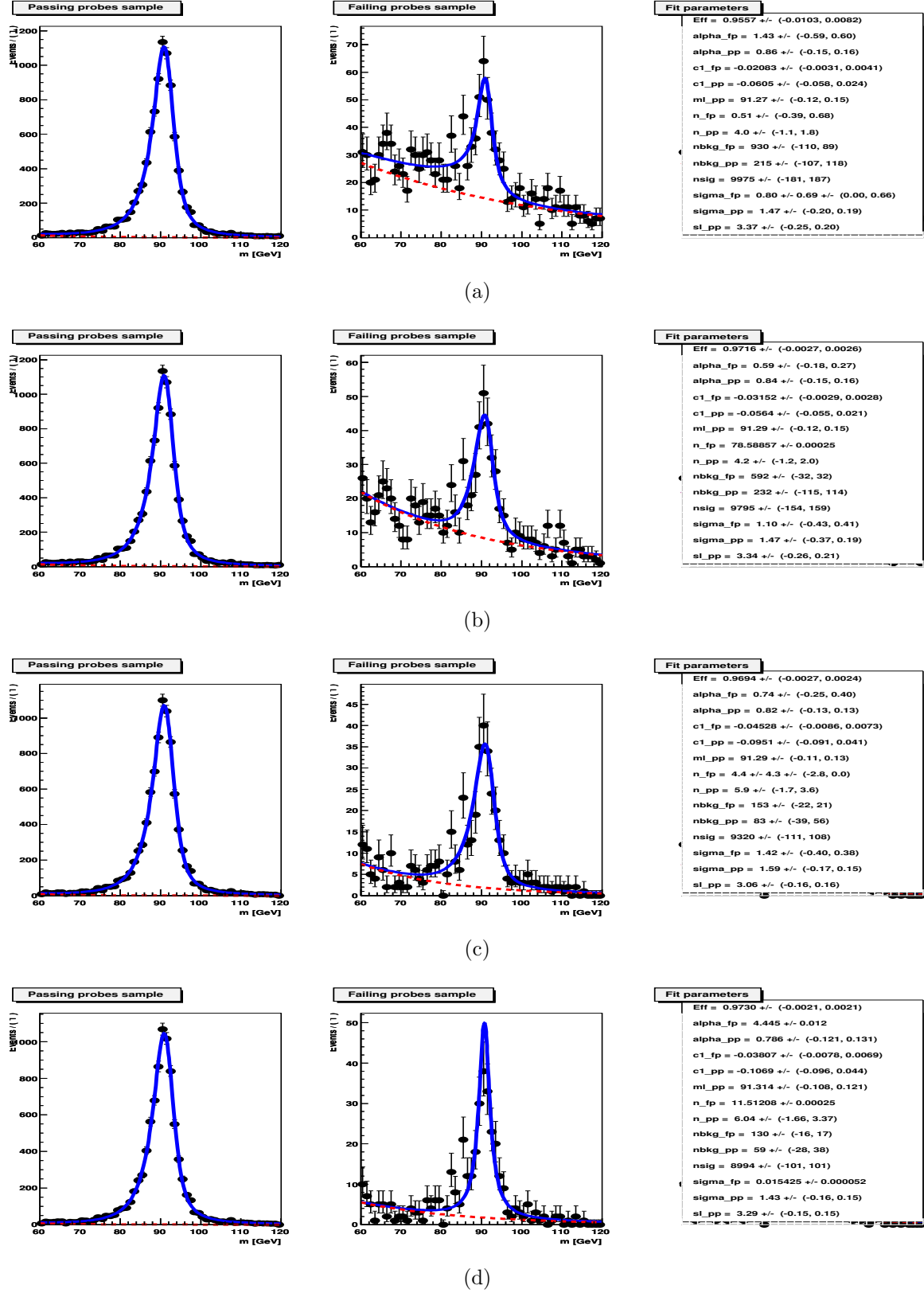


Figure 4.16 – Signal Lorentz(x)CB Bruit de fond Exp - tag WP60 et diverses coupures sur le probe aucune coupure (a), H/E (b), $H/E+Iso$ (c) et $H/E+Iso+\sigma_{iniq}$ (d). Le fit a été réalisé pour une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$.

4.4 Résultats finaux et conclusion

En conclusion des parties précédentes, j'ai mesuré grâce à la méthode tag-and-probe les efficacités de reconstruction des électrons sur les données 2010 et une partie des données 2011. J'ai vérifié que les sélections sur l'événement et sur le tag ne biaisaient pas la mesure et garantissaient une pureté suffisante. Sur les distributions de masse invariante $m_{e,SC}$, nous pouvons constater que la mesure obtenue par fit nous donne des résultats en meilleur accord avec l'attendu Monte-Carlo que ceux obtenus par simple comptage après sélection. Afin d'éviter tout biais nous n'avons appliqué aucune coupure sur le probe, excepté lorsque cela était imposé par le système de déclenchement (cas des données 2011). Nous obtenons au final pour les données 2010 les résultats de la table 4.16 :

	data	MC
$0 < \eta < 0.78$	$97.94^{+0.42}_{-0.42}(\text{stat})^{+0.55}_{-0.55}(\text{syst})$	$98.37^{+0.06}_{-0.07}$
$0.78 < \eta < 1.444$	$96.0^{+1.3}_{-1.3}(\text{stat})^{+0.55}_{-0.55}(\text{syst})$	$97.62^{+0.08}_{-0.08}$
$1.444 < \eta < 1.566$	$85.2^{+2.9}_{-2.9}(\text{stat})^{+0.55}_{-0.55}(\text{syst})$	$88.17^{+0.51}_{-0.53}$
$1.566 < \eta < 2.0$	$96.8^{+1.1}_{-1.1}(\text{stat})^{+0.55}_{-0.55}(\text{syst})$	$96.27^{+0.13}_{-0.14}$
$2.0 < \eta < 2.5$	$91.8^{+1.6}_{-2.2}(\text{stat})^{+0.55}_{-0.55}(\text{syst})$	$93.17^{+0.18}_{-0.19}$

Table 4.16 – Efficacités de reconstruction (en %) en fonction de η obtenues par ajustement simultané des échantillons passing probes et failing probes pour un tag WP60, aucune sélection sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Les erreurs statistiques et systématiques sont ici indiquées. Données 2010.

Pour les données 2011, nous obtenons les résultats regroupés dans les tables 4.17 et 4.18, et présentés dans la figure 4.17.

	data	MC
$0 < \eta < 0.78$	$97.94^{+0.11}_{-0.11}(\text{stat})^{+0.98}_{-0.98}(\text{syst})$	$98.08^{+0.03}_{-0.04}$
$0.78 < \eta < 1.444$	$97.35^{+0.17}_{-0.19}(\text{stat})^{+1.22}_{-1.22}(\text{syst})$	$97.6^{+0.04}_{-0.04}$
$1.444 < \eta < 1.566$	$94.18^{+0.57}_{-0.57}(\text{stat})^{+0.12}_{-0.12}(\text{syst})$	$95.74^{+0.17}_{-0.18}$
$1.566 < \eta < 2.0$	$95.24^{+0.25}_{-0.25}(\text{stat})^{+0.14}_{-0.14}(\text{syst})$	$96.00^{+0.07}_{-0.07}$
$2.0 < \eta < 2.5$	$93.66^{+0.31}_{-0.32}(\text{stat})^{+0.14}_{-0.14}(\text{syst})$	$93.48^{+0.1}_{-0.1}$

Table 4.17 – Efficacités de reconstruction (en %) en fonction de η obtenues par ajustement simultané de différentes formes (signal et bruit de fond) des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Les erreurs statistiques et systématiques sont ici indiquées. Données 2011.

	data	MC
$10 < E_T < 15$	$92.2^{+2.2}_{-2.2}(\text{stat}) \quad {}^{+1.0}_{-1.0}(\text{syst})$	$87.81^{+0.45}_{-0.45}$
$15 < E_T < 20$	$94.3^{+2.0}_{-2.0}(\text{stat}) \quad {}^{+1.4}_{-1.4}(\text{syst})$	$91.40^{+0.25}_{-0.25}$
$20 < E_T < 30$	$93.37^{+0.79}_{-0.79}(\text{stat}) \quad {}^{+0.08}_{-0.08}(\text{syst})$	$94.87^{+0.08}_{-0.08}$
$30 < E_T < 40$	$96.25^{+0.33}_{-0.24}(\text{stat}) \quad {}^{+0.74}_{-0.74}(\text{syst})$	$96.50^{+0.04}_{-0.04}$
$40 < E_T < 50$	$97.32^{+0.11}_{-0.11}(\text{stat}) \quad {}^{+0.18}_{-0.18}(\text{syst})$	$97.23^{+0.04}_{-0.04}$
$E_T > 50$	$97.96^{+0.18}_{-0.18}(\text{stat}) \quad {}^{+0.27}_{-0.27}(\text{syst})$	$97.83^{+0.06}_{-0.06}$

Table 4.18 – Efficacités de reconstruction (en %) obtenues par ajustement simultané de différentes formes (signal et bruit de fond) des échantillons passing probes et failing probes pour un tag WP60, $H/E < 0.15$ (tonneau) et $H/E < 0.1$ (bouchons) sur le probe et une fenêtre de masse $60 < m_{e,SC} < 120 \text{ GeV}.c^{-2}$. Les erreurs statistiques et systématiques sont ici indiquées. Données 2011.

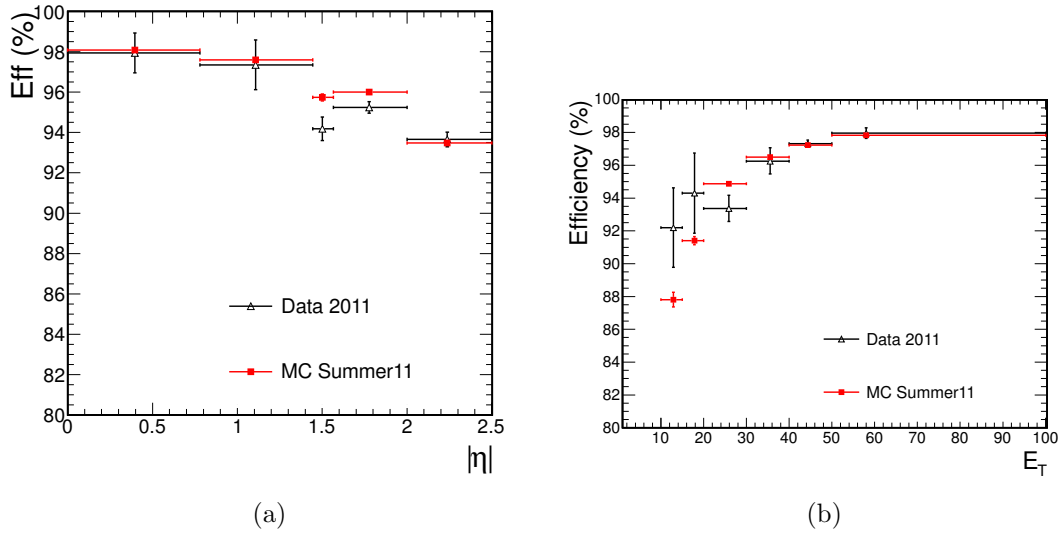


Figure 4.17 – Efficacités de reconstruction (%) des électrons issus des données 2011 (noir) et attendues MC (rouge) en fonction de η (a) et en fonction de E_T (b).

Ces résultats constituent les mesures officielles de la collaboration CMS pour les données 2010 et 2011 et ont été utilisés par la collaboration pour les analyses et les résultats publics.

Chapitre 5

Développement d'un algorithme de récupération des radiations dans l'état final

Contenu du chapitre

5.1	Modélisation de l'émission de FSR, comparaison et choix du générateur	105
5.2	Caractéristiques des photons de FSR	106
5.3	Récupération des photons FSR lointains	107
5.3.1	Étude au niveau générateur	107
5.3.1.1	Sélection	107
5.3.1.2	Résultats	108
5.3.2	Étude sur les événements reconstruits	109
5.3.2.1	Sélection	109
5.3.2.2	Performances	111
5.4	Récupération de la composante FSR dans un cône étroit autour du lepton	113
5.4.1	Sélection	113
5.4.2	Performances	114
5.5	Validation sur les données.	115

L'émission de photons dans l'état final [28] dénommée en anglais *Final State Radiation* (FSR) constitue une correction d'électrodynamique quantique qui, appliquée au canal de désintégration $ZZ \rightarrow 4l$, concerne un, et parfois deux, des leptons issus de la désintégration d'un boson Z. Cet effet affecte le spectre de masse du boson Z, et donc a fortiori celui des di-bosons ZZ, comme nous pouvons le constater sur la figure 5.1, mais altère également leurs distributions cinématiques. Dans le cas de l'analyse $ZZ \rightarrow 4l$, que nous présenterons dans le chapitre 6, les coupures de sélection dans le spectre de masse di-lepton pour sélectionner des Z sur leur couche de masse imposent $m_{ll} > 60 \text{ GeV.c}^{-2}$. Cette coupure cause une perte due aux FSR de 1.41% des événements. Nous pouvons comprendre l'enjeu que représente la correction des radiations dans l'état final : récupérer la composante FSR, et ainsi reconstituer pleinement la masse du Z et ses distributions cinématiques, permet d'augmenter la précision d'une mesure de masse ou encore de rendre plus sensible une analyse basée sur les déviations de ces distributions par rapport à une hypothèse Modèle Standard (cf chapitre 6). J'ai donc réalisé une étude ayant pour objectif d'élaborer un algorithme de récupération des photons FSR en fin d'analyse ZZ, dans le but d'assurer la correction dans le spectre de masse des Z tout en garantissant une pureté élevée afin d'éviter de biaiser cette masse reconstruite par les paires de leptons en y ajoutant un photon issu du bruit de fond et satisfaisant aux contraintes apportées par l'isolation nécessaire à l'analyse ZZ. Ainsi dans une première partie je présenterai l'effet des FSR en fonction de deux générateurs Monte-Carlo. Puis j'introduirai les caractéristiques attendues des photons de FSR en opposition avec celles des photons de bruit de fond afin d'en dégager des coupures de sélection. Ensuite j'élaborerai un premier algorithme de récupération des FSR lointains puis un deuxième, ciblant les FSR proches des leptons et permettant d'obtenir une grande pureté, chacun répondant à des contraintes propres, tout en détaillant leurs performances respectives.

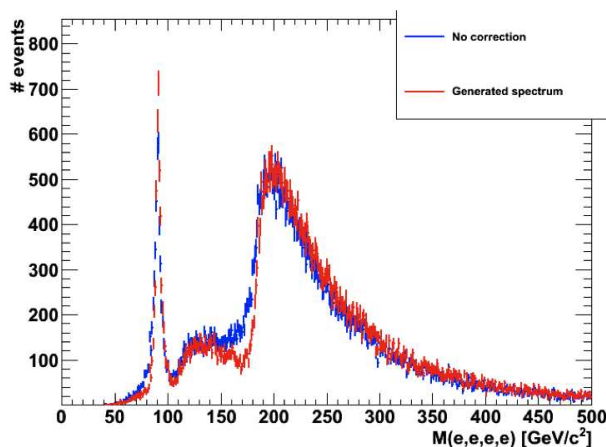


Figure 5.1 – Spectre de masse $4e$ avant effet FSR (rouge) et après (bleu).

5.1 Modélisation de l'émission de FSR, comparaison et choix du générateur

J'ai étudié l'effet de l'émission de photons dans l'état final pour le canal $ZZ \rightarrow 4l$ à l'aide de deux générateurs. D'une part le générateur au second ordre (NLO) Powheg, et de l'autre le générateur au premier ordre plus n jets, Sherpa. Plus précisément dans le cas de Sherpa, les échantillons ont été générés avec $n \leq 1$. La comparaison des distributions obtenues après effet de FSR par chaque générateur est présentée dans la figure 5.2, pour l'ensemble du spectre de masse formé par le système quatre muons et dans le cas spécifique où nous sélectionnons deux Z sur leur couche de masse par $60 < m_{\mu^+\mu^-} < 120 \text{ GeV.c}^{-2}$.

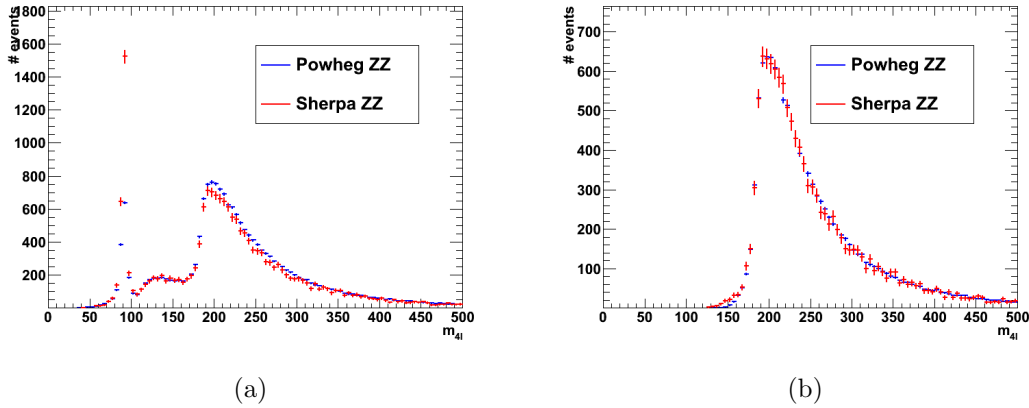


Figure 5.2 – Comparaison des distributions obtenues par le générateur Powheg (bleu) et Sherpa (rouge) pour l'ensemble du spectre quatre-muons (a) ou pour $ZZ \rightarrow 4\mu$ (b). Distributions normalisées au nombre total d'événements.

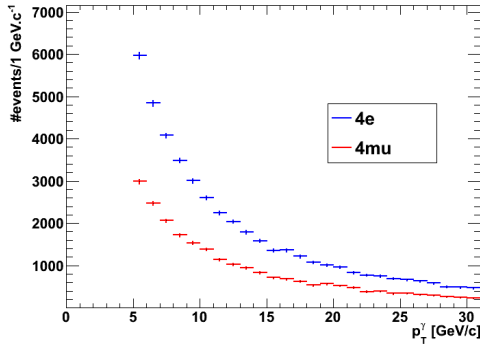
Concernant le premier cas nous pouvons voir un bon accord entre la prédiction Powheg et celle de Sherpa : la différence essentielle ne concerne pas la modélisation de l'effet FSR mais réside dans la modélisation de la désintégration $Z \rightarrow 4l$ pour laquelle nous obtenons un plus large excès avec Sherpa. Cependant avec l'analyse ZZ nous requérons $60 < m_{ll} < 120 \text{ GeV.c}^{-2}$ et nous obtenons donc la figure 5.2 (b). Nous constatons là encore un bon accord entre les deux générateurs, signe que la physique et les radiations dans l'état final sont modélisées de manière comparable.

Le générateur Powheg présente l'avantage de référencer toutes les particules, y compris intermédiaires, intervenant dans les différents processus physiques à l'aide leur *pdgId* respectif tandis que Sherpa ne référence que les particules de l'état initial ou de l'état final. Par exemple pour le cas $pp \rightarrow ZZ \rightarrow 4l$, les bosons Z n'apparaissent pas avec le générateur Sherpa ; nous n'avons que les gluons et quarks initiaux et les leptons et photons finaux. Afin de faciliter cette étude et ne constatant que peu de différence entre les deux générateurs en ce qui concerne l'effet de l'émission de FSR, je poursuivrai avec le générateur Powheg.

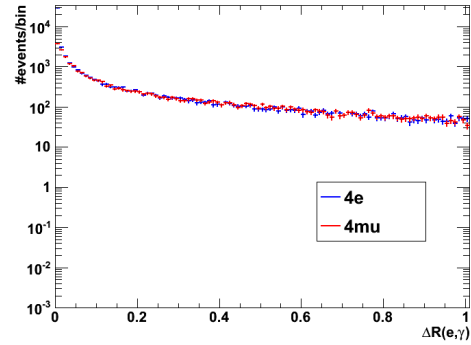
5.2 Caractéristiques des photons de FSR

Cette étude est réalisée dans le cadre du canal de désintégration $ZZ \rightarrow 4l$ et consiste en une comparaison des caractéristiques des photons de FSR avec celles des photons issus du bruit de fond. Soulignons tout d'abord qu'il existe plusieurs sources de photons de bruit de fond. Il peut s'agir :

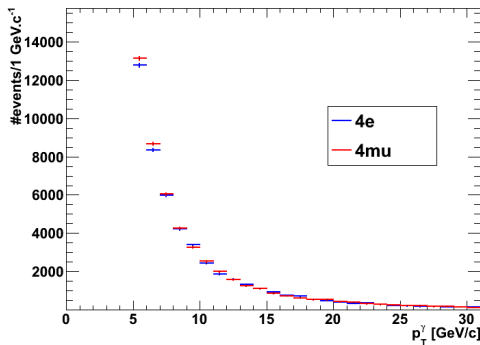
- des événements sous-jacents : ce sont là des photons provenant d'autres produits des collisions proton-proton ; le plus souvent ils sont issus de la désintégration de pions neutres.
- du phénomène d'empilement : ces photons sont semblables à ceux des événements sous-jacents, car comme eux il proviennent d'une interaction proton-proton.
- des radiations dans l'état initial (ou ISR : Initial State Radiation) : les photons sont émis par les quarks des processus Drell-Yann.
- les autres : cette catégorie regroupe toutes les sources de photons restantes, des photons de bremsstrahlung en passant par les photons contenus dans des jets ou encore les jets faussement reconstruits en photon.



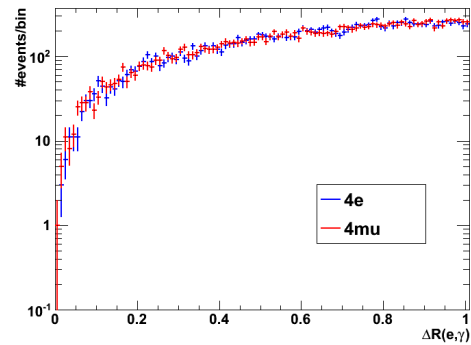
(a)



(b)



(c)



(d)

Figure 5.3 – Comparaison des distributions de p_T (a) et de dR (b) des photons FSR (a,b) et des photons de bruit de fond (c,d) dans les cas $Z \rightarrow e^+e^-$ (bleu) ou $Z \rightarrow \mu^+\mu^-$ (rouge).

Je vais tout d'abord comparer les caractéristiques, au niveau générateur, de photons (FSR et bruit de fond) dans le cas $Z \rightarrow e^+e^-$ avec ceux issus du $Z \rightarrow \mu^+\mu^-$. Comme nous pouvons le constater avec la figure 5.3, le phénomène de radiation dans l'état final est plus important dans le canal électron que dans le canal muon, l'excès concerne essentiellement les FSR colinéaires aux leptons, tandis que pour les photons de bruit de fond, les caractéristiques sont similaires dans les deux canaux. Les coupures de sélection seront donc communes aux deux cas. En nous concentrant sur une étude comparative entre photons FSR et photons de bruit de fond, nous remarquons certaines caractéristiques propres aux photons FSR. En observant la distribution en fonction de p_T sur figure 5.4 nous constatons que les photons de bruit de fond sont moins "durs" que les photons de FSR même si 91% des photons FSR sont de $p_T < 10 \text{ GeV.c}^{-1}$.

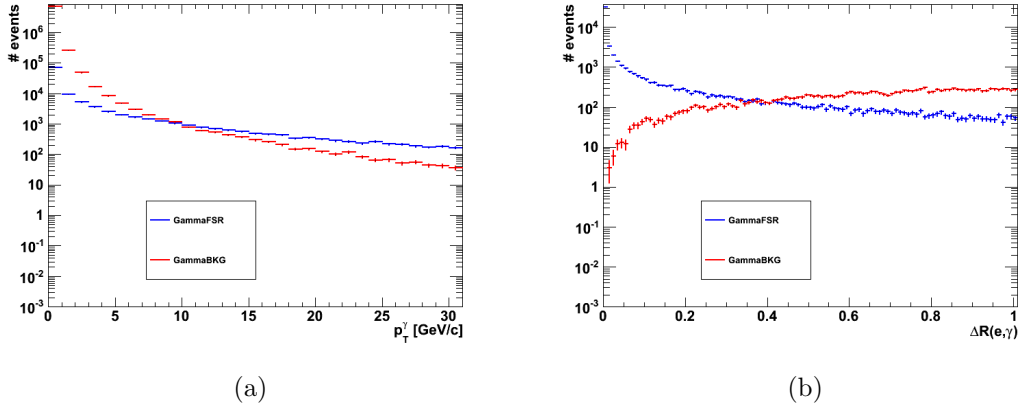


Figure 5.4 – Comparaison des photons FSR (bleu) et photons du bruit de fond (rouge) pour les distributions en fonction p_T (a) et le dR (b); dans ce dernier cas une coupure à $p_T > 5 \text{ GeV}$ est effectuée.

5.3 Récupération des photons FSR lointains

5.3.1 Étude au niveau générateur

5.3.1.1 Sélection

J'élabore ici un premier algorithme de récupération de la composante FSR, en étudiant ses effets sur la vérité Monte-Carlo avant processus de reconstruction; le but étant de concilier pureté des événements corrigés et maximisation du nombre d'événements corrigés. Fort des observations faites dans la partie précédente sur des distributions avant effet de la reconstruction, nous pouvons déjà introduire deux coupures de sélection. Les photons de FSR prédominent à haut p_T , au contraire des photons de bruit de fond, mais sont cependant essentiellement concentrés à des $p_T < 10 \text{ GeV.c}^{-1}$: une coupure à 5 GeV.c^{-1} permet d'éliminer une très grande proportion de bruit de fond. De plus les photons FSR sont principalement colinéaires aux leptons nous entraînant à ne garder dans notre algorithme que les photons ayant un $dR = \sqrt{(\eta_l - \eta_{\gamma_{FSR}})^2 + (\phi_l - \phi_{\gamma_{FSR}})^2} < 0.7$. En

observant la figure 5.4 (b) cette coupure pourrait être plus stricte mais des contraintes d'isolation au niveau de la sélection ZZ empêchant la recherche des photons FSR dans toute la zone interne modifieront la géométrie de ce cône de sélection, comme nous le verrons plus loin. Le fait d'associer au lepton, les photons FSR se trouvant à l'intérieur d'un cône de rayon $dR < 0.7$ autour de ce lepton est correct dans 90.31% des cas. Nous pouvons cependant souligner que ce chiffre n'est pas uniformément réparti en fonction de dR mais se dégrade au fur et à mesure que le photon FSR est éloigné du lepton. Nous obtenons plus précisément une association correcte pour :

- $dR < 0.1$ dans 99.80% des cas.
- $0.1 < dR < 0.3$ dans 94.33% des cas.
- $0.3 < dR < 0.7$ dans 79.76% des cas.

En testant ces premières coupures dans un algorithme simple se contentant de corriger la masse dilepton par un photon vérifiant $p_T > 5 \text{ GeV.c}^{-1}$ et situé dans un cône $dR < 0.7$ autour d'un des leptons de la paire, nous voyons apparaître, figure 5.5 (a), un effet de sur-correction à haute masse. Ce problème est résolu (cf. figure 5.5 (b)) en imposant un critère supplémentaire de contrainte de masse : le candidat FSR ne sera considéré que s'il corrige la masse m_{ll} en la rapprochant de la masse nominale $m_Z = 91.18 \text{ GeV.c}^{-2}$. Ce critère n'est cependant approprié que pour les Z sur leur couche de masse et donc ne concernera que les candidats satisfaisant $60 < m_{ll} < 120 \text{ GeV.c}^{-2}$.

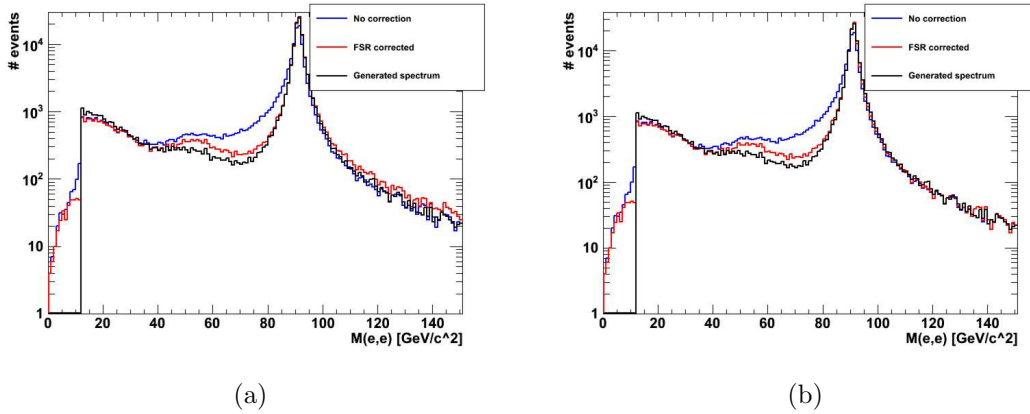


Figure 5.5 – Spectre de masse du Z sans correction de FSR (bleu) vs après correction (rouge) via algorithme de récupération vs généré (noir), i.e. sans effet de FSR. Sans contrainte de masse (a) et avec (b).

5.3.1.2 Résultats

A ce stade de l'élaboration, nous avons un algorithme optimisé pour récupérer la composante FSR jusqu'à $dR < 0.7$ avec une contrainte de masse qui supprime l'effet de sur-correction mais également augmente la pureté : cet effet de sur-correction concerne

particulièrement les photons de bruit de fond qui, s'ils sont ajoutés à la masse m_{ll} , la biaise forcément vers des valeurs supérieures à celle initiale créée par la physique. L'effet est parfaitement clair dans les résultats regroupés en table 5.1, où nous pouvons constater l'augmentation en termes de pureté, apportée par la contrainte de masse, de plus de 15% dans le cas 4e et près de 25% dans le cas 4μ .

Contrainte de sélection	$ZZ \rightarrow 4e$	$ZZ \rightarrow 4\mu$
association au lepton le plus proche	2.6	2.5
$E_T > 5$	45.0	31.5
$E_T > 5 + dR_{cut}$	83.8	72.0
$E_T > 5 + dR_{cut} + \text{contrainte de masse}$	97.6	96.2

Table 5.1 – *Tableau des puretés (%) avant reconstruction pour les cas $ZZ \rightarrow 4e$ et $ZZ \rightarrow 4\mu$.*

Le processus de reconstruction que nous considérerons complètement dans la section suivante, introduit les notions d'isolation afin de supprimer notamment les nouveaux bruits de fond engendrés par la reconstruction, à savoir les jets faussement reconstruits comme photons. L'isolation ECAL est calculée, comme nous l'avons vue dans le chapitre 3, en formant un cône d'isolation de $dR < 0.3$ autour des leptons. Par conséquent tout rayonnement FSR donnant une contribution non négligeable à l'intérieur de ce cône est susceptible de faire identifier, à tort, le lepton comme non isolé et supprimer l'événement. Or comme cet algorithme de récupération des photons de FSR est fait pour être positionné en fin d'analyse, le cas précédemment décrit sera forcément éliminé en amont du processus de correction FSR. Ce nouvel élément nous conduit à ne considérer que les photons situés dans un anneau autour du lepton et dimensionné par $0.3 < dR < 0.7$.

Les nouvelles performances sur la pureté sont regroupées dans la table 5.2. Nous remarquons que nous sommes réduits à une zone plus contaminée en bruit de fond où là encore la contrainte de masse joue un rôle majeur en augmentant la pureté de plus de 40%. Cet algorithme présente un taux de correction, i.e. la fraction d'événements corrigés après sélection ZZ , de 2.05% des cas $ZZ \rightarrow 4e$ et 2.17% des cas $ZZ \rightarrow 4\mu$.

Contrainte de sélection	$ZZ \rightarrow 4e$	$ZZ \rightarrow 4\mu$
association au lepton le plus proche	2.6	2.5
$E_T > 5$	45.0	31.5
$E_T > 5 + dR_{cut} \ 0.3-0.7$	35.7	37.3
$E_T > 5 + dR_{cut} \ 0.3-0.7 + \text{contrainte de masse}$	79.3	84.3

Table 5.2 – *Tableau des puretés (%) avant reconstruction en ajoutant les contraintes d'isolation pour les cas $ZZ \rightarrow 4e$ et $ZZ \rightarrow 4\mu$.*

5.3.2 Étude sur les événements reconstruits

5.3.2.1 Sélection

A l'issue de la reconstruction des leptons, de nouvelles contraintes apparaissent et nécessitent d'adapter l'algorithme en conséquence comme nous avons déjà commencé à le

voir dans la section précédente. Dans le cas des électrons, le super-agrégat permet de collecter automatiquement 59% des photons de FSR. Cependant comme nous l'avons déjà dit, l'existence d'un cône d'isolation ($dR < 0.3$), nécessaire afin d'obtenir une meilleure réjection du bruit de fond, et qui contrôle la présence d'activité dans le calorimètre électromagnétique (EcallIso) en prenant soin d'enlever la composante du lepton, supprime les événements avec un photon à l'intérieur de ce cône. Nous sommes donc obligés de ne considérer que les photons tels que $0.3 < dR < 0.7$. Ce dernier critère nous place dans une zone plus contaminée par le bruit de fond d'autant plus que la reconstruction introduit une nouvelle source de photons de bruit de fond, à savoir les jets faussement reconstruits comme des photons. Par conséquent nous ajoutons aux précédents de nouveaux critères de sélection :

- isolation sur les traces autour du photon : $TkIso < 2$ GeV.
- ne considérer que les photons qui corrigent m_{ll} à l'intérieur d'une fenêtre 91.18 ± 10 GeV. c^{-2} .
- lorsqu'il existe plusieurs photons candidats à la correction d'un seul Z, les prendre par ordre croissant en dR et ne garder que le premier qui vérifie tous les critères de sélection précédents.

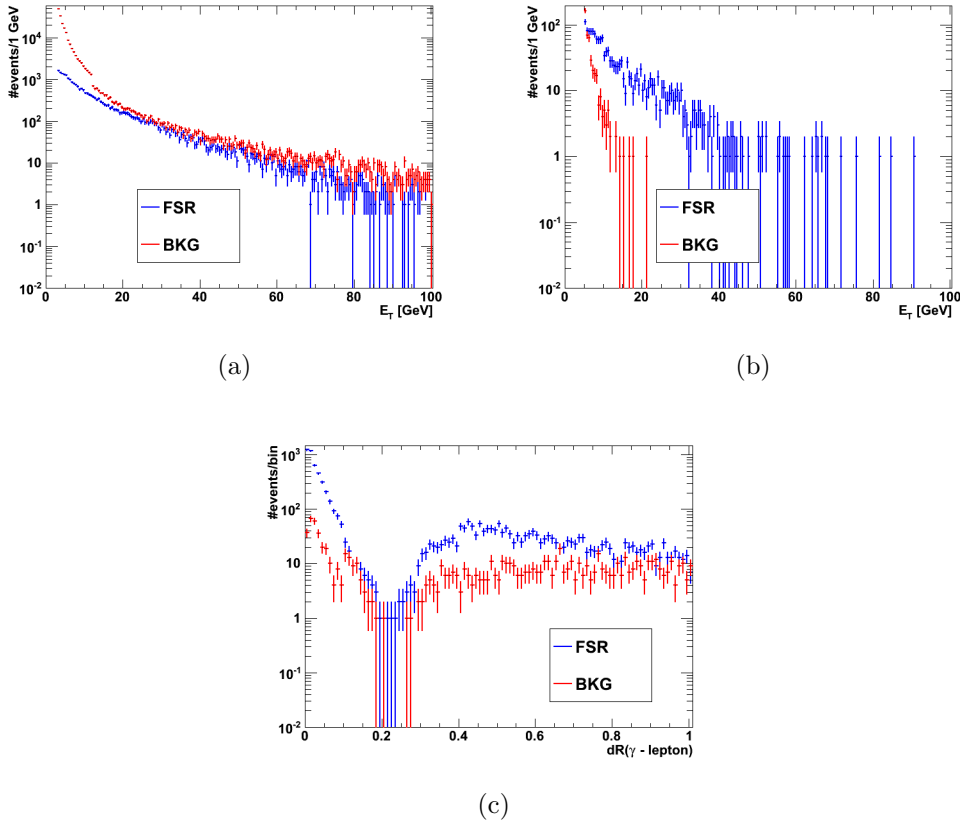


Figure 5.6 – Comparaison des distributions des photons FSR (bleu) avec les photons du bruit de fond (rouge) : en fonction de E_T avant (a) et après (b) sélection. Comparaison en fonction de dR (c) pour toutes les coupures appliquées exceptée celle sur la variable dR elle-même.

Les effets de cette sélection en termes de réduction du bruit de fond sont présentés dans la figure 5.6. Avant sélection, la figure 5.6 (a) présente un saut dans la distribution des photons du bruit de fond. Ceci est dû au fait que la collection des photons reconstruits ne comprend que les super-agrégats vérifiant $E_T > 10$ GeV et que pour rechercher des photons de FSR jusqu'à $E_T = 5$ GeV nous avons ajouté la collection des super-agrégats pour la fenêtre $4 < E_T < 12$ GeV en prenant soin d'éviter de considérer deux fois le même super-agrégat pour $10 < E_T < 12$ GeV. Or au niveau de la reconstruction des photons, une coupure $H/E > 0.15$ pour réduire le bruit de fond est appliquée au contraire des super-agrégats. Ceci explique le saut à 12 GeV plus important dans la distribution des photons de bruit de fond. Avec la figure 5.6 (b), nous voyons clairement la réduction du bruit de fond apportée par la sélection. La figure 5.6 (c) présente la distribution en fonction de dR après application de toutes les coupures de sélection exceptée celle sur dR . Cette figure met en évidence l'effet du cône d'isolation qui creuse les distributions des photons jusqu'à $dR=0.3$. Elle montre également que la région $0.3 < dR < 0.7$ est une zone où le bruit de fond demeure important malgré la sélection. Les performances de cet algorithme sont détaillées dans la section suivante.

5.3.2.2 Performances

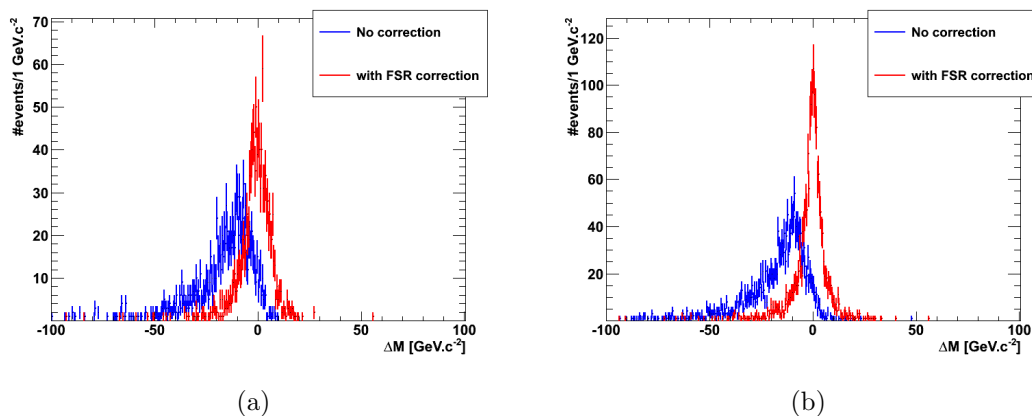


Figure 5.7 – Comparaison de la différence Δm entre masse reconstruite et masse générée pour le canal $ZZ \rightarrow 4e$ (a) et $ZZ \rightarrow 4\mu$ (b) dans le cas corrigé de la perte FSR (rouge) et non corrigé (bleu) pour des événements sélectionnés par l'algorithme.

Nous pouvons observer l'effet de la correction apportée par ce premier algorithme avec la figure 5.7. Elle y présente le Δm entre masse reconstruite, corrigée ou non de la composante FSR, et masse générée pour des événements sélectionnés par l'algorithme. Ainsi l'avantage de la correction est très visible avec un Δm beaucoup plus piqué et centré en 0.

Les performances de ce premier algorithme en termes de pureté et de pourcentage d'événements di-bosons corrigés sont présentées dans les tables 5.3 et 5.4. Nous remarquons que même avec l'ensemble des critères de sélection, nous n'atteignons qu'une pureté de 74.43% dans le cas $ZZ \rightarrow 4e$ et de 78.59% dans le cas $ZZ \rightarrow 4\mu$ ce qui est relativement faible. La pureté

inférieure dans le canal électron s'explique par le fait que les électrons sont plus affectés par le phénomène de bremsstrahlung que les muons, ce qui compromet l'association au plus proche lepton.

Par ailleurs en déterminant la pureté de la sélection par rapport au nombre de ver-

	$ZZ \rightarrow 4e$	
Contrainte de sélection	Pureté (%)	% corrigés
$E_T > 5$	12.0	non calculée
$E_T > 5 + \text{TkIso}$	22.7	22.8
$E_T > 5 + \text{TkIso} + \text{dRcut}$	48.2	3.5
$E_T > 5 + \text{TkIso} + \text{dRcut} + \text{contrainte et fenêtre de masse}$	74.4	1.0

Table 5.3 – Tableau des pureté et pourcentage d'événements corrigés attendus pour le cas $ZZ \rightarrow 4e$.

	$ZZ \rightarrow 4\mu$	
Contrainte de sélection	Pureté (%)	% corrigés
$E_T > 5$	20.7	non calculée
$E_T > 5 + \text{TkIso}$	36.0	25.2
$E_T > 5 + \text{TkIso} + \text{dRcut}$	49.2	3.5
$E_T > 5 + \text{TkIso} + \text{dRcut} + \text{contrainte et fenêtre de masse}$	78.6	1.0

Table 5.4 – Tableau des pureté et pourcentage d'événements corrigés attendus pour le cas $ZZ \rightarrow 4\mu$.

tex, comme cela est fait dans la figure 5.8, nous pouvons remarquer que la pureté se dégrade avec le nombre de vertex. Ainsi ce premier algorithme se révèle très sensible au phénomène d'empilement. De plus ces performances ont été étudiées avec un échantillon Monte-Carlo présentant les caractéristiques d'empilement prévues pour l'année 2011 au LHC et beaucoup plus faibles que l'année 2012 et que les années à venir où l'augmentation de luminosité intégrée instantanée entraîne irrémédiablement un accroissement du phénomène d'empilement. Ce dernier point, ainsi que la faible pureté, m'ont amené à développer un nouvel algorithme.

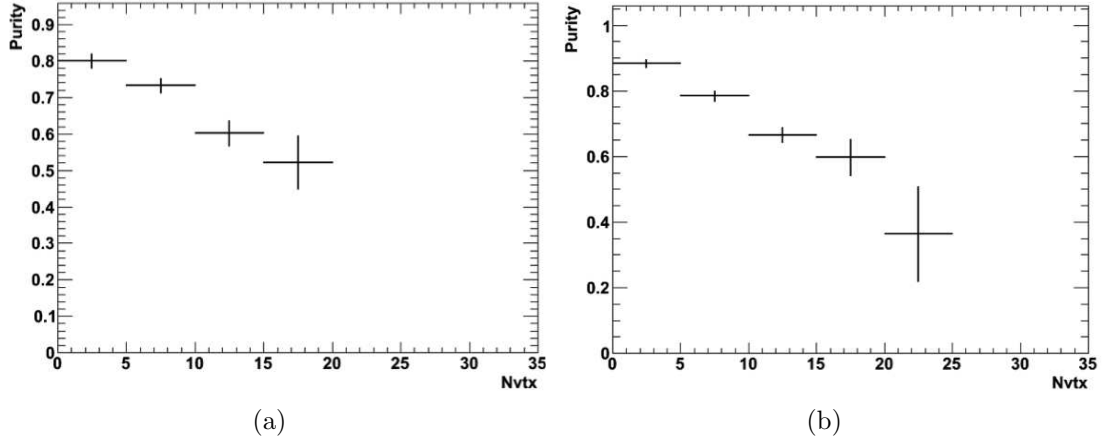


Figure 5.8 – Puret  en fonction du nombre de vertex dans le cas $ZZ \rightarrow 4e$ (a) et $ZZ \rightarrow 4\mu$ (b).

5.4 R cup ration de la composante FSR dans un c ne  troit autour du lepton

5.4.1 S lection

Afin d viter la d pendance au ph nom ne d'empilement et de satisfaire les contraintes de puret , je me suis int ress  aux photons proches du lepton. En effet en observant de nouveau la figure 5.4, nous constatons clairement que les photons colin aires au lepton sont tr s majoritairement des photons FSR. Comme nous l'avons d taill  dans le chapitre 3 et d j   voqu  un peu plus haut, les  lectrons poss dent un super-agr gat qui r cup re par construction les photons FSR colin aires dans une proportion de 59%. Si nous voulons d velopper un algorithme de r cup ration de la composante FSR dans le cas  lectron nous devons revoir l'isolation, or celle-ci a  t  optimis e pour rejeter au mieux le bruit de fond. Si nous prenons maintenant le cas des muons, ces-derniers pr sentent un veto c ne ($dR < 0.07$)   l'int rieur du c ne d'isolation. Ce veto c ne d finit une zone dans laquelle tout d p t d' nergie n'est pas consid r  pour le calcul de l'isolation et par cons quent constitue une nouvelle fen tre de recherche de photons FSR.

L'algorithme ne concerne donc que les muons. Il requiert pour les photons :

- $p_T > 5 \text{ GeV}/c$.
- $dR < 0.07$. Cette coupure implique une zone de recherche tr s pure en photons de FSR : il n'est plus n cessaire d'ajouter une coupure sur l'isolation trace (TkIso) autour du photon.
- ne consid rer que les photons qui corrigent m_{ll}   l'int rieur d'une fen tre $91.18 \pm 10 \text{ GeV}.c^{-2}$.
- lorsqu'il existe plusieurs photons candidats   la correction d'un seul Z, les prendre par ordre croissant en dR et ne garder que le premier qui v rifie tous les crit res de s lection pr c dents.

La figure 5.9 présente les distributions des photons de FSR et de bruit de fond avant et après sélection. Nous remarquons que nous réduisons considérablement la proportion de bruit fond tout en conservant une quantité importante de photons FSR.

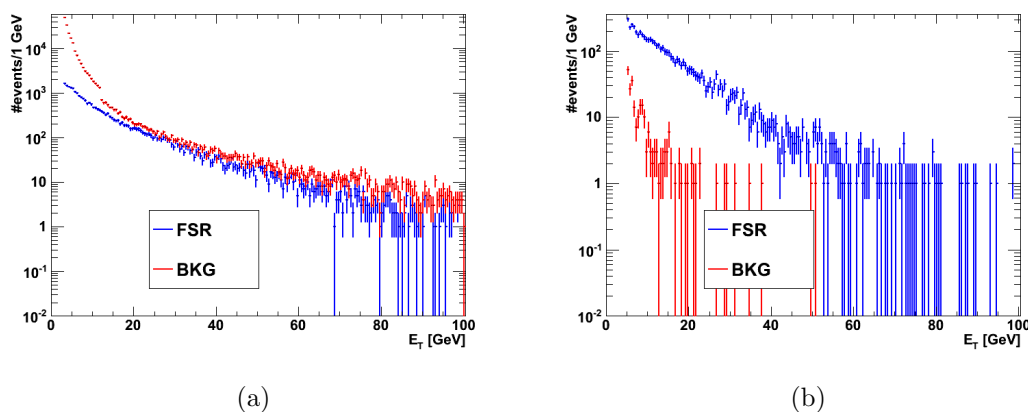


Figure 5.9 – Comparaison des distributions des photons FSR (bleu) avec les photons du bruit de fond (rouge) : en fonction de E_T avant (a) et après (b) sélection.

5.4.2 Performances

Les performances de ce nouvel algorithme de récupération de la composante FSR sont très supérieures au précédent comme nous pouvons le voir avec les résultats de la table 5.5. La pureté de la sélection est de plus totalement indépendante du phénomène d'empilement comme l'illustre la figure 5.10.

Contrainte de sélection	Pureté (%)	% corrigés
Sélection complète	95.6	4.2

Table 5.5 – Tableau des pureté et pourcentage d'événements corrigés attendus pour le cas $ZZ \rightarrow 4\mu$ avec le second algorithme de récupération des FSR dans un cône étroit autour des leptons.

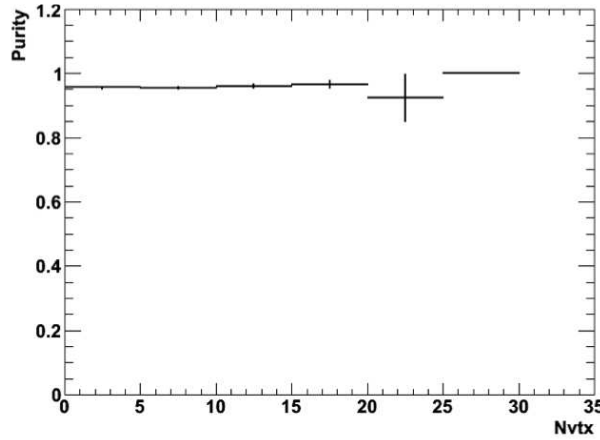


Figure 5.10 – Pureté en fonction du nombre de vertex dans le cas $ZZ \rightarrow 4\mu$ obtenue avec le second algorithme de récupération de FSR.

5.5 Validation sur les données.

Une validation sur les données 2011 a été effectuée dans le cas de l'algorithme de récupération des photons FSR lointains, i.e. situés dans un cône $0.3 < dR < 0.7$ autour du lepton. La sélection tag-and-probe pour former un lot pur de leptons, requiert des électrons vérifiant $p_T > 7 \text{ GeV.c}^{-1}$ et $|\eta| < 2.5$ ou des muons satisfaisant $p_T > 5 \text{ GeV.c}^{-1}$ et $|\eta| < 2.4$ et ne garde que les paires "tag-probe" de même saveur, de signes opposés et formant une masse invariante $50 < m_{ll} < 120 \text{ GeV.c}^{-2}$. De plus le lepton de plus haut p_T doit vérifier $p_T > 20 \text{ GeV.c}^{-1}$ tandis que le second lepton doit satisfaire $p_T > 10 \text{ GeV.c}^{-1}$. Des coupures sur la somme des isolations combinées de chaque lepton, $R_{iso,1} + R_{iso,2} < 0.35$, et sur le paramètre d'impact de chaque lepton, $|SIP3D| < 4$, sont également appliquées. L'algorithme de récupération de la composante FSR intervient sur ces événements sélectionnés. Les échantillons Monte-Carlo considérés sont générés par Powheg, excepté pour le bruit de fond QCD qui est généré par Pythia. Plus de détails sur ces échantillons Monte-Carlo sont donnés dans le chapitre 6 et plus précisément dans la section 6.1.2.2. Nous contrôlons en premier lieu le bon accord entre les données et l'attendu Monte-Carlo avec la figure 5.11.

Ensuite nous vérifions l'amélioration apportée par l'algorithme de récupération des photons de FSR uniquement sur les cas sélectionnés par l'algorithme, et le bon accord entre données et simulation. avec la figure 5.12. Avant correction, les spectres de masse présentent une large queue à basse masse, alors qu'après correction la distribution présente un pic plus important à 91.2 GeV.c^{-2} et est étroitement centrée autour de cette valeur, signe de la performance de la correction. Nous conservons sur les données l'amélioration et le bon accord avec la prédiction Monte-Carlo ce qui nous permet de valider l'algorithme de récupération des photons issus des radiations dans l'état final.

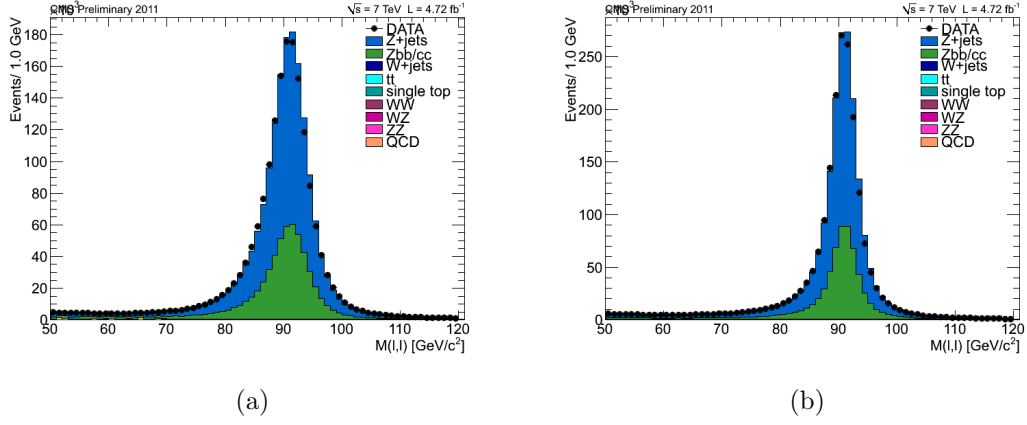


Figure 5.11 – Comparaison des spectres de masse di-électrons (a) et di-muons (b) entre données et simulations, normalisée à une luminosité de 4.72 fb^{-1} .

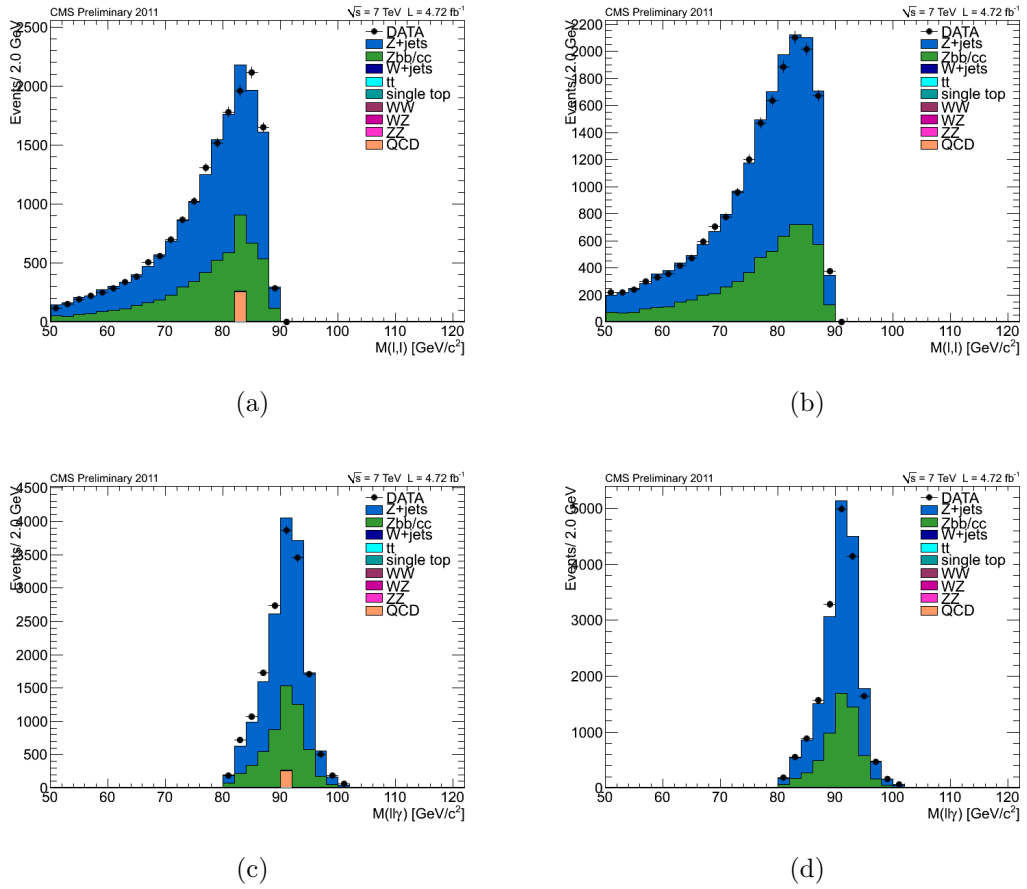


Figure 5.12 – Comparaison entre données et simulations des spectres de masse di-électrons, (a) et (c) et di-muons, (b) et (d), sélectionnés par l'algorithme de récupération des photons FSR, avant, (a) et (b), et après correction, (c) et (d). Les distributions sont normalisées à une luminosité de 4.72 fb^{-1} .

Chapitre 6

Analyse ZZ et couplages anormaux trilinéaires neutres

Contenu du chapitre

6.1	Mesure de di-bosons ZZ	118
6.1.1	États finaux $ZZ \rightarrow 4l$	118
6.1.2	Données expérimentales et échantillons Monte-Carlo	120
6.1.2.1	Données expérimentales CMS	120
6.1.2.2	Échantillons Monte-Carlo	121
6.1.3	Sélection et performances	122
6.1.3.1	Sélection des événements	122
6.1.3.2	Performances de sélection	124
6.1.4	Contrôle des bruits de fond	126
6.1.5	Distributions différentielles pour les candidats ZZ.	127
6.2	Recherche de couplages trilinéaires neutres	134
6.2.1	Méthodologie	134
6.2.1.1	Choix et validation du générateur	134
6.2.1.2	Méthode de repondération des distributions	135
6.2.1.3	Choix de la variable discriminante	137
6.2.2	Interprétation statistique, modèles simulés et résultats	138
6.2.2.1	Méthode CL_s	139
6.2.2.2	Limites d'exclusion	141

L'analyse ZZ que nous allons détailler dans ce chapitre est une sélection particulière d'événements développée et optimisée principalement par le groupe du Laboratoire Leprince-Ringuet en vue de la recherche du boson de Higgs et de nouvelle physique dans les canaux multileptons. Je présenterai dans une première partie le canal d'étude puis les données et les échantillons Monte-Carlo utilisés, j'aborderai ensuite la stratégie de sélection afin d'isoler le signal ainsi que ses performances et enfin je terminerai sur l'évaluation du bruit de fond à l'issue de la sélection.

J'aborderai ensuite, dans une deuxième partie, ce qui constitue l'analyse physique de ma thèse : la recherche, avec les données recueillies par le détecteur CMS au cours des années 2010-2011, de couplages anormaux trilineaires neutres (aTGC, pour anomalous Triple Gauge Couplings dans la dénomination anglaise) de type ZZZ. Ces couplages sont, comme cela a été détaillé dans le chapitre 1, décrits par deux coefficients f_4^Z et f_5^Z dans le Lagrangien effectif. Nous détaillerons dans un premier temps la stratégie de cette étude et enfin nous terminerons avec les résultats obtenus.

6.1 Mesure de di-bosons ZZ

6.1.1 États finaux $ZZ \rightarrow 4l$

Le canal $pp \rightarrow ZZ \rightarrow 4l$ ($l=e,\mu$) se caractérise par une signature claire constituée par quatre leptons isolés donnant un très bon rapport signal sur bruit. Cependant il s'agit d'un processus de production rare auquel s'ajoute un faible rapport de branchement des bosons Z en paires de leptons :

- La section efficace de production $\sigma_{pp \rightarrow ZZ} = 6.4 \pm 0.6 \text{ pb}$.
- Rapport de branchement (ou BR : *Branching Ratio*) pour chacun des Z : $BR_{Z \rightarrow e^+e^-} = BR_{Z \rightarrow \mu^+\mu^-} = 3.36\%$ ce qui nous donne pour deux Z : $(3.36\%)^2$ soit environ 0.45%.
- Trois états finaux $e^+e^-e^+e^-$, $\mu^+\mu^-\mu^+\mu^-$ et $e^+e^-\mu^+\mu^-$, ce dernier comptant pour deux du point de vue de la combinatoire pour calculer le rapport de branchement total $BR_{tot} \simeq 1.8\%$

La figure 6.1 présente les principaux processus de production de di-bosons ZZ dans le Modèle Standard. La table 6.1 présente les sections efficaces $\sigma \times BR$ de production de di-bosons dans différents canaux. Nous remarquons que le canal $ZZ \rightarrow 4l$ est le plus rare. Penchons nous maintenant sur la topologie de ce signal : les événements caractéristiques

Processus	$\sigma \times BR$
$ZZ \rightarrow 4l$	0.095 pb
$WZ \rightarrow 3l\nu$	0.868 pb
$WW \rightarrow 2l2\nu$	4.88 pb

Table 6.1 – Tableau comparatif des $\sigma \times BR$ pour divers signaux di-bosons.

possèdent deux paires de leptons isolés de même saveur et de signe opposé ($(e^+e^-e^+e^-)$, $(\mu^+\mu^-\mu^+\mu^-)$ et $(e^+e^-\mu^+\mu^-)$) formant chacune une masse reconstruite proche de celle

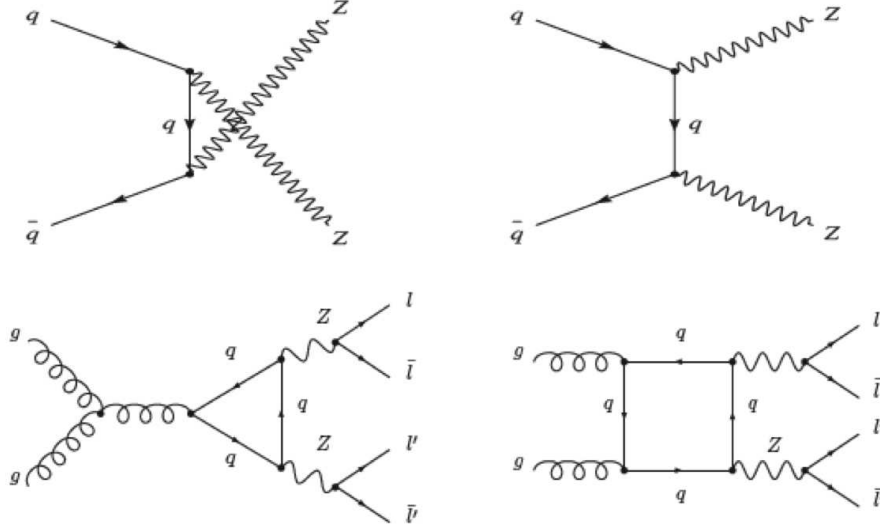


Figure 6.1 – Diagrammes de Feynman des principaux processus de production Modèle Standard de ZZ.

d'un boson Z sur sa couche de masse : $m_{2l} \simeq m_Z = 91.2 \text{ GeV.c}^{-2}$.

Les bruits de fond ayant une topologie similaire présentent des sections efficaces de production largement supérieures à celle de la production de di-bosons ZZ et doivent par conséquent être réduits, notamment en appliquant des contraintes sur les leptons, comme nous l'aborderons plus loin dans le paragraphe dédié à la sélection des événements. Les principaux bruits de fond sont :

1. les bruits de fond instrumentaux définis comme tels car impliquant des jets faussement reconstruits comme leptons.
 - WZ+jets : les produits de désintégration de WZ présentent naturellement trois leptons isolés ; combinés à un jet additionnel, nous obtenons une signature proche de celle recherchée.
 - W+Jets et Z+jets : ces bruits de fond présentent un ou deux leptons isolés qui associés à deux faux leptons issus des jets fournissent une topologie similaire à celle du signal.
 - QCD : la production de multi-jets faussement identifiés comme leptons peut également fausser la topologie recherchée.
2. les bruits de fond réductibles présentant de vrais leptons issus des désintégrations de bosons électrofaibles ainsi que des leptons issus de désintégrations de quarks lourds :
 - $Zb\bar{b}$ et $Zc\bar{c}$
 - $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+\nu bl^-\bar{\nu}\bar{b}$

6.1.2 Données expérimentales et échantillons Monte-Carlo

6.1.2.1 Données expérimentales CMS

Les données expérimentales analysées sont celles recueillies par CMS au cours des années 2010 et 2011, pour une énergie de collision de 7 TeV et correspondant à une luminosité intégrée de 5.02 fb^{-1} . La luminosité est connue avec une précision de 2.2%. Les données CMS considérées dans cette analyse sont issues de jeux de données définis par la combinaison de chemins de déclenchement de haut niveau (HLT Triggers). Le menu de déclenchement évoluant au cours des différentes périodes d'acquisition, les lots de données traités dans cette analyse reflètent ces modifications et sont, de manière générale, issus d'une sélection demandant une paire d'électrons ou de muons au dessus d'un certain seuil en p_T . (cf table 6.2).

Un filtrage commun à tous les canaux 4 leptons a été appliqué sur ces lots de données dans un objectif de réduction des quantités à analyser. Il requiert :

- au moins deux candidats leptons reconstruits
- le lepton dont l'impulsion transverse est la plus grande doit vérifier $p_{T,1} > 20 \text{ GeV.c}^{-1}$ et l'autre lepton : $p_{T,2} > 7 \text{ GeV.c}^{-1}$ dans le cas des électrons et $p_{T,2} > 5 \text{ GeV.c}^{-1}$ dans le cas des muons
- une masse invariante formée par ces deux leptons telle que $m_{1,2} > 40 \text{ GeV.c}^{-2}$.

Période	Nom de l'échantillon	Nom du Trigger
2010A	/Mu/Run2010A-Apr21ReReco-v1	DoubleMu3
2010A	/EG/Run2010A-Apr21ReReco-v1	Ele10_LW "OR" Ele15_SW
2010B	/Mu/Run2010B-Apr21ReReco-v1	DoubleMu3
2010B	/Electron/Run2010B-Apr21ReReco-v1	Ele17_SW_CaloEleId "OR" Ele17_SW_TightEleId "OR" Ele17_SW_TighterEleIdIso1
2011A	/DoubleMu/Run2011A-05Jul2011ReREco-ECAL-v1	Mu13_Mu8
2011A	/DoubleElectron/Run2011A-05Jul2011ReREco-ECAL-v1	Ele17_CaloIdL_CaloIsoVL Ele8_CaloIdL_CaloIsoVL
2011A	/DoubleMu/Run2011A-05Aug2011-v1	Mu13_Mu8
2011A	/DoubleElectron/Run2011A-05Aug2011-v1	Ele17_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL Ele8_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL
2011A	/DoubleMu/Run2011A-03Oct2011-v1	Mu13_Mu8
2011A	/DoubleElectron/Run2011A-03Oct2011-v1	Ele17_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL Ele8_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL
2011B	/DoubleMu/Run2011B-PromptReco-v1	Mu13_Mu8
2011B	/DoubleElectron/Run2011B-PromptReco-v1	Ele17_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL Ele8_CaloIdT_CaloIsoVL_TrIdVL_TrIsoVL

Table 6.2 – Détails des échantillons de données expérimentales et des noms des configurations du système de déclenchement.

6.1.2.2 Échantillons Monte-Carlo

Les échantillons de données Monte-Carlo sont les résultats de simulations des phénomènes physiques et de la procédure complète de reconstruction des différentes particules et autres objets physiques. Ces échantillons sont utilisés pour la mise en place et l'optimisation de la sélection des événements ZZ en vue d'être appliquée ensuite à l'analyse des données expérimentales. Ils sont également utilisés pour les comparaisons entre mesures expérimentales et prédictions Monte-Carlo. Cependant l'évaluation du bruit de fond sera préférentiellement extraite des données à l'aide des "régions de contrôles" où les mesures faites dans une région dominée par le bruit de fond sont extrapolées à la région propre au signal.

Les échantillons de signal se distinguent en deux composantes : la contribution principale $q\bar{q} \rightarrow ZZ$, et $gg \rightarrow ZZ$. Les bruits de fond, pour lesquels les jets hadroniques ou les leptons secondaires issus des désintégrations de quarks lourds sont faussement reconstitués et identifiés comme leptons primaires, se distinguent en plusieurs contributions comme nous l'avons déjà introduit en début de chapitre. La composante QCD, celle issue des désintégrations des autres di-bosons (WW, WZ, Z γ) ainsi que celle constituée par la production du top solitaire sont également considérées. Les différents échantillons Monte-Carlo utilisés pour cette étude sont regroupés dans la table 6.3.

Processus	MC generator	$\sigma_{(N)NLO}$	Nom de l'échantillon
signal ZZ			
$q\bar{q} \rightarrow ZZ \rightarrow 4e(4\mu, 4\tau)$	POWHEG	15.34 fb	ZZTo4e(4mu,4tau)
$q\bar{q} \rightarrow ZZ \rightarrow 2e2\mu$	POWHEG	30.68 fb	ZZTo2e2mu
$q\bar{q} \rightarrow ZZ \rightarrow 2e(2\mu)2\tau$	POWHEG	30.68 fb	ZZTo2e(2mu)2tau
$gg \rightarrow ZZ \rightarrow 2\ell 2\ell'$	gg2ZZ	3.48 fb	GluGluToZZTo2L2L
$gg \rightarrow ZZ \rightarrow 4\ell$	gg2ZZ	1.74 fb	GluGluToZZTo4L
Autres di-bosons			
$WZ \rightarrow 3\ell\nu$	Madgraph	0.868 pb	WZJetsTo3LNU
$t\bar{t}$ and single t			
$t\bar{t} \rightarrow \ell^+\ell^-\nu\bar{\nu}b\bar{b}$	POWHEG	17.32 pb	TTTo2L2Nu2B
Z/W + jets ($q = d, u, s, c, b$)			
Z + jets, $m_{\ell\ell} > 50$	MadGraph	3048 pb	DYJetsToLL*M-50
QCD multi-jets inclusifs, $\hat{p}_T^{\min}$ binné			
$b, c \rightarrow e + X$	PYTHIA		QCD_Pt-XXtoYY_BCtoE
EM-enriched	PYTHIA		QCD_Pt-XXtoYY_EMEnriched
MU-enriched	PYTHIA		QCD_Pt-XXtoYY_MuPt5Enriched

Table 6.3 – Échantillons de données simulées Monte-Carlo pour le signal et les différents bruits de fond ; Z signifie Z, Z*, γ^* ; ℓ désigne e, μ or τ .

Le générateur Monte-Carlo multi-usage PYTHIA [29] est utilisé pour les procédés incluant la production de multi-jets QCD. Il sert soit dans le cas de la génération au premier ordre (LO : Leading Order) d'un processus dur donné, soit dans le but de modéliser la fragmentation des quarks et gluons, l'hadronisation, les désintégrations ou d'ajouter des événements sous-jacents lorsque des processus durs sont générés à des ordres supérieurs par d'autres générateurs. Cette dernière option est utilisée dans les cas faisant usage du générateur Madgraph(MadEvent) [30] ou de POWHEG [31] générateur à l'ordre 2 (NLO : Next to Leading Order), i.e. pour le signal et les principaux bruits de fond. Dans le cas particulier de $gg \rightarrow ZZ$, nous utilisons le générateur gg2ZZ [32].

6.1.3 Sélection et performances

6.1.3.1 Sélection des événements

La sélection ne retient tout d'abord que les événements comptant au moins quatre leptons, électrons ou muons, vérifiant :

- $|\eta_e| < 2.5$, $p_T^e > 7 \text{ GeV.c}^{-1}$, trackIso relative < 0.7 et critère identification WP60 dans le cas électron.
- $|\phi_\mu| < 2.4$, $p_T^\mu > 5 \text{ GeV.c}^{-1}$, trackIso relative < 0.7 et "global muons", i.e. identification reposant sur l'association entre trace muon du trajectographe et traces laissées dans les chambres à muons.

Les étapes de sélection sont ensuite, dans cet ordre :

1. premier candidat Z formé par une paire de candidats leptons de charge opposée et de même saveur (e^+e^- ou $\mu^+\mu^-$) vérifiant $m_{1,2} > 50 \text{ GeV.c}^{-2}$, $p_{T,1} > 20 \text{ GeV.c}^{-1}$ et $p_{T,2} > 10 \text{ GeV.c}^{-1}$. Chaque lepton doit avoir une isolation relative combinée (cf. chapitre 3) telle que $R_{iso,i} < 0.275$. Une coupure sur la signification du paramètre d'impact est également effectuée : $|SIP_{3D} = \frac{IP}{\sigma_{3D}}| < 4$ pour chacun des leptons de la paire, avec IP , le paramètre d'impact du lepton sélectionné, défini en dimension 3 au point le plus proche du vertex primaire d'interaction, et σ_{3D} son incertitude. Enfin le candidat Z ayant une masse reconstruite la plus proche de la masse nominale de 91.2 GeV.c^{-2} est conservé et noté Z_1 .
2. Au moins un candidat lepton additionnel sans critère de charge ni de saveur.
3. Un deuxième candidat lepton ayant la même saveur que le précédent et de charge opposée.
4. Choix du meilleur candidat "4l" : la seconde paire de leptons, notée Z_2 , parmi toutes les combinaisons de paires l^+l^- restantes vérifiant $m_{Z_2} > 12 \text{ GeV.c}^{-2}$ doit former avec la première paire une masse $m_{4l} > 100 \text{ GeV.c}^{-2}$. Dans les cas particuliers 4e et 4 μ , trois des quatre combinaisons de paires de signe opposées doivent vérifier $m_{ll} > 12 \text{ GeV.c}^{-2}$. Enfin si il existe plus d'une combinaison formant le candidat Z_2 , nous conservons celui formé à partir des leptons de plus haut p_T .
5. Chacun des quatre leptons doit vérifier $R_{iso,i} < 0.275$.
6. Coupure sur le paramètre d'impact : chaque lepton doit vérifier $|SIP_{3D} = \frac{IP}{\sigma_{3D}}| < 4$.

7. Sélection "haute masse". Il s'agit des coupures cinématiques finales afin de sélectionner deux candidats Z sur leur couche de masse en requérant $60 < m_{Z_1} < 120 \text{ GeV} \cdot \text{c}^{-2}$ et $60 < m_{Z_2} < 120 \text{ GeV} \cdot \text{c}^{-2}$.
8. Sur les paires de leptons vérifiant l'étape précédente, l'algorithme de récupération de FSR effectue une recherche de photons et son association à l'un des Z le cas échéant.

Dans la première étape les seuils en p_T requis se trouvent sur le plateau d'efficacité haute des algorithmes utilisés pour le déclenchement. Précisons que cette première étape sert à la sélection des événements Z+X utilisés pour l'analyse et le contrôle des bruits de fond Z+jets, $Zb\bar{b}/c\bar{c}$ et $t\bar{t}$. Les quatre premières étapes ont été ainsi établies et optimisées dans le but de réduire les composantes des bruits de fond QCD et Z+jets tout en préservant au maximum le signal. Le choix d'imposer $m_{4l} > 100 \text{ GeV} \cdot \text{c}^{-2}$ a pour objectif de s'affranchir du cas particulier de la désintégration $Z \rightarrow 4l$ [33], qui correspond à l'émission d'un photon ou d'un Z virtuel par l'un des leptons, émission se matérialisant ensuite en une paire de leptons. Les étapes suivantes servent finalement à réduire les composantes des bruits de fond $Zb\bar{b}/c\bar{c}$, $t\bar{t}$ et di-bosons WZ.

6.1.3.2 Performances de sélection

Les résultats obtenus pour les différentes étapes de la sélection et pour les divers canaux sont présentés dans la figure 6.2 et regroupés dans la table 6.4. Comme il a été montré [34] que les contributions QCD sont nulles ou négligeables dès l'étape 2, elles ne sont pas considérées ici.

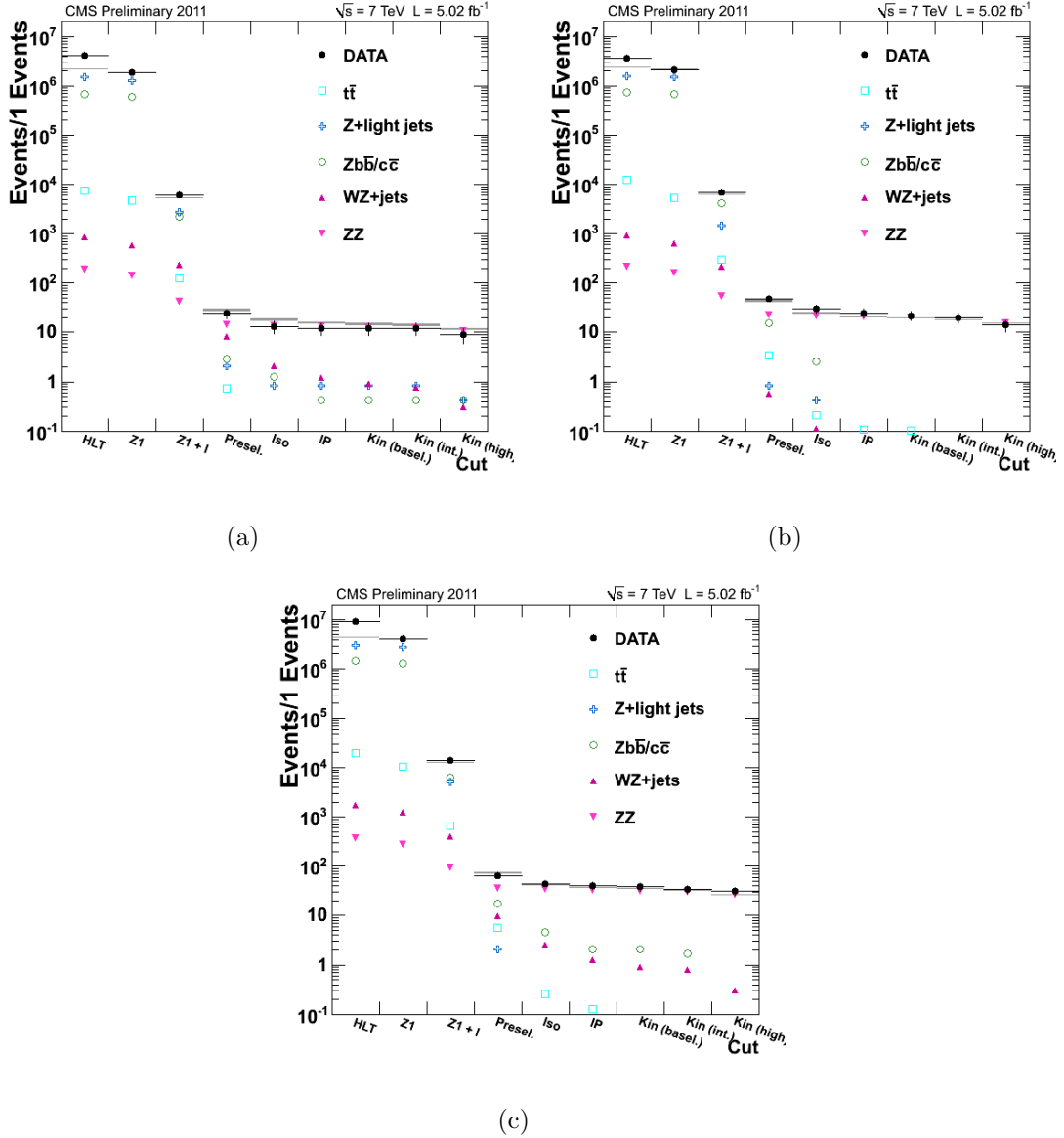


Figure 6.2 – Rendement des différentes étapes de sélection pour les canaux $4e$ (a), 4μ (b) et $2e2\mu$ (c). La luminosité considérée correspond à $\mathcal{L}_{\text{int}} = 5.02 \text{ fb}^{-1}$.

Etape	$t\bar{t}$	Z+jets	Z $b\bar{b}/c\bar{c}$	WZ+jets	ZZ	Total	Data
Trigger	1.23×10^4	$1.6 \times 10^6 \pm 822$	$7.45 \times 10^5 \pm 561$	916 ± 1.83	212 ± 0.246	$2.35 \times 10^6 \pm 995$	3.63×10^6
Z ₁	5.46×10^3	$1.49 \times 10^6 \pm 794$	$6.93 \times 10^5 \pm 542$	673 ± 1.57	161 ± 0.214	$2.19 \times 10^6 \pm 961$	2.16×10^6
Z ₁ + ℓ	305 ± 6.78	$1.53 \times 10^3 \pm 25.2$	$4.15 \times 10^3 \pm 41.9$	213 ± 0.88	54.3 ± 0.1	$6.22 \times 10^3 \pm 48.9$	6.98×10^3
4 ℓ	3.46 ± 0.171	0.846 ± 0.598	15.2 ± 2.54	0.57 ± 0.05	22.3 ± 0.06	42.4 ± 2.61	47
Iso	0.21 ± 0.0421	0.423 ± 0.423	2.54 ± 1.04	0.11 ± 0.02	21.2 ± 0.05	24.5 ± 1.12	30
IP	0.109 ± 0.0291	0	0	0.05 ± 0.01	20.8 ± 0.05	21 ± 0.06	25
Final	0.0168 ± 0.0119	0	0	0.02 ± 0.01	15.4 ± 0.05	15.4 ± 0.05	14

(a)

Etape	$t\bar{t}$	Z+jets	Z $b\bar{b}/c\bar{c}$	WZ+jets	ZZ	Total	Data
Trigger	$7.43 \times 10^3 \pm 7.91$	$1.48 \times 10^6 \pm 791$	$6.83 \times 10^5 \pm 538$	847 ± 1.76	193 ± 0.234	$2.17 \times 10^6 \pm 956$	4.17×10^6
Z ₁	$4.82 \times 10^3 \pm 6.36$	$1.28 \times 10^6 \pm 735$	$5.94 \times 10^5 \pm 501$	590 ± 1.47	140 ± 0.199	$1.88 \times 10^6 \pm 889$	1.89×10^6
Z ₁ + ℓ	125 ± 1.03	$2.71 \times 10^3 \pm 33.9$	$2.18 \times 10^3 \pm 30.4$	229 ± 0.92	42 ± 0.08	$5.3 \times 10^3 \pm 45.5$	6.08×10^3
4 ℓ	0.723 ± 0.078	2.12 ± 0.946	2.96 ± 1.12	8.17 ± 0.17	14.3 ± 0.04	28.3 ± 1.48	24
Iso	0.0505 ± 0.0206	0.846 ± 0.598	1.27 ± 0.733	2.12 ± 0.09	13.7 ± 0.04	18 ± 0.95	13
IP	0.0252 ± 0.0146	0.846 ± 0.598	0.423 ± 0.423	1.19 ± 0.07	13.1 ± 0.04	15.6 ± 0.74	12
Final	0.0084 ± 0.0084	0.423 ± 0.423	0.423 ± 0.423	0.31 ± 0.03	10.5 ± 0.04	11.7 ± 0.6	12

(b)

Etape	$t\bar{t}$	Z+jets	Z $b\bar{b}/c\bar{c}$	WZ+jets	ZZ	Total	Data
Trigger	$1.97 \times 10^4 \pm 12.9$	$3.07 \times 10^6 \pm 1140$	$1.43 \times 10^6 \pm 777$	1760 ± 2.54	368 ± 0.318	$4.52 \times 10^6 \pm 1380$	8.89×10^6
Z ₁	$1.03 \times 10^4 \pm 6.36$	$2.77 \times 10^6 \pm 1080$	$1.29 \times 10^6 \pm 738$	1240 ± 2.13	278 ± 0.277	$4.07 \times 10^6 \pm 1310$	4.04×10^6
Z ₁ + ℓ	652 ± 2.34	$5.08 \times 10^3 \pm 46.4$	$6.41 \times 10^3 \pm 52.1$	403 ± 1.22	94 ± 0.19	$1.26 \times 10^4 \pm 69.8$	1.43×10^3
4 ℓ	5.75 ± 0.22	2.12 ± 0.95	17.8 ± 2.74	9.73 ± 0.19	36.4 ± 0.119	71.8 ± 2.92	64
Iso	0.261 ± 0.0468	0	4.65 ± 1.4	2.53 ± 0.10	34.2 ± 0.115	41.6 ± 1.41	44
IP	0.126 ± 0.0326	0	2.12 ± 0.946	1.28 ± 0.07	33.1 ± 0.11	36.7 ± 0.96	40
Final	0.0252 ± 0.015	0	1.69 ± 0.85	0.30 ± 0.03	26.2 ± 0.10	26.5 ± 0.11	31

(c)

Table 6.4 – Nombre d'événements dans les cas 4μ (a), $4e$ (b) and $2e2\mu$ (c) à l'issue des différentes étapes de sélection de l'analyse et pour une luminosité de 5.02 fb^{-1} .

6.1.4 Contrôle des bruits de fond

La topologie "leptons isolés" caractéristique de la désintégration du boson Z permet, grâce à l'application de critères d'isolation contraignants, d'obtenir en fin de sélection très peu de bruit de fond. En observant les performances de sélection détaillées dans la partie 6.1.3.2, nous remarquons que le bruit de fond restant se limite essentiellement aux composantes Z+jets et $Zb\bar{b}/c\bar{c}$: en effet dans ces cas un jet ou un lepton non isolés peut être faussement reconstruit et identifié comme un lepton isolé. La trop faible proportion de ces événements dans les simulations n'en permet pas une estimation précise dans la zone de signal. Par ailleurs nous pensons que les simulations Monte-Carlo pour ces bruits de fond ne décrivent pas forcément correctement les données. Par conséquent la proportion de bruit de fond va être estimée à partir des données : technique dite "data-driven".

La procédure d'estimation du bruit de fond à partir des données consiste à choisir une large région de contrôle du bruit de fond en dehors de l'espace de phase propre au signal et de la peupler en relâchant les critères de sélections des événements. Nous contrôlons que le taux d'événements varie en accord avec l'attendu Monte-Carlo. La région de contrôle doit être soigneusement déterminée car les critères de sélection faiblement contraignants utilisés peuvent nous éloigner de la région signal et rendre l'extrapolation plus difficile.

La région de contrôle Z+X (X= jet, $b\bar{b}/c\bar{c}$) est constituée par le sous-ensemble d'événements vérifiant la première étape de la sélection, c'est-à-dire la détermination du premier candidat Z_1 . De ce sous-ensemble, nous gardons les événements avec deux leptons additionnels de même saveur et de même signe, formant une masse invariante telle que $60 < m_{Z_2} < 120$ GeV.c⁻² et $m_{4l} > 100$ GeV.c⁻². Afin de déterminer le nombre d'événements de bruit de fond dans la région signal, N_{B_i} , le taux de mauvaise identification F_{l_i} est appliqué aux événements N_i vérifiant tous les critères de sélection, contrainte d'opposition de charge sur la paire formant Z_2 incluse, à l'exception cependant des critères d'isolation laissés libres pour cette deuxième paire.

Trois régions de contrôle sont ainsi définies pour chaque canal de désintégration $4e$, 4μ , $2e2\mu$:

- Catégorie 0 : cette catégorie contient tous les événements tels que les deux leptons formant le Z_2 ne vérifient pas la contrainte d'isolation. Cette catégorie est dominée par les événements Z+jets et la proportion de bruit de fond est estimée par :

$$N_{B_0} = \frac{N_0 F(l_1) F(l_2)}{1 - F(l_1) F(l_2)} \quad (6.1)$$

- Catégorie 1 : cette catégorie contient tous les événements tels que seul le premier lepton formant le Z_2 ne vérifie pas la contrainte d'isolation. La contribution de bruit de fond pour cette catégorie est donnée par :

$$N_{B_1} = \frac{N_1 F(l_1)}{1 - F(l_1)} \quad (6.2)$$

- Catégorie 2 : cette catégorie est similaire à la catégorie 1 avec pour différence que seul le second lepton formant le Z_2 ne vérifie pas la contrainte d'isolation. Ainsi nous

obtenons de la même manière la proportion de bruit de fond pour cette catégorie par :

$$N_{B_2} = \frac{N_2 F(l_2)}{1 - F(l_2)} \quad (6.3)$$

L'estimation globale du bruit de fond est obtenue en additionnant tous les bruits de fond. Notons également que la catégorie 0 contribue pour chacune des catégories 1 et 2 et par conséquent doit être soustraite afin d'éviter un double comptage. L'estimation finale est donnée par :

$$N_{bkg \text{ total}} = N_{cat1} + N_{cat2} - N_{cat0} \quad (6.4)$$

et les résultats sont regroupés dans la table 6.5

Canal de désintégration	Catégorie 0	Catégorie 1	Catégorie 2	Bruit de fond total estimé
4μ	0.053	0.577	0	$0.52^{+0.26}_{-0.26}$
4e	0.221	0.172	0.302	$0.25^{+0.14}_{-0.14}$
$2e2\mu$	0.249	0.552	0.276	$0.58^{+0.18}_{-0.18}$

Table 6.5 – *Estimation du bruit de fond pour chaque catégorie ainsi que l'estimation globale pour chacun des canaux de désintégration, normalisée à la luminosité $\mathcal{L}_{\text{int}} = 5.02 \text{ fb}^{-1}$. Les erreurs citées sont statistiques.*

6.1.5 Distributions différentielles pour les candidats ZZ.

Nous présentons dans cette partie les résultats obtenus pour les distributions cinématiques telles que la masse invariante du système ZZ ou encore l'impulsion transverse et la rapidité des candidats Z, obtenues en fin de sélection. Notons que ce sont les variables que je considérerai dans mon étude des couplages anormaux détaillée dans la partie suivante. Les distributions sont également présentées après application de l'algorithme de récupération des FSR dans le cas 4μ et pour la combinaison de tous canaux.

La forme du bruit de fond Z+X ne peut être déterminée par son allure en fin d'analyse : en raison du faible nombre d'événements, ces derniers ne sont pas régulièrement distribués selon le spectre de masse ou l'impulsion transverse, pouvant provoquer des accumulations non significatives dues aux fluctuations statistiques. Par conséquent et de manière comparable à la procédure d'estimation du bruit de fond, nous nous plaçons dans une zone où sa contribution est forte : en fin de première étape de la sélection. Nous ajustons alors l'allure de sa distribution à l'aide d'une modélisation. Pour les distributions de masse invariante 4 leptons, d'impulsion transverse du Z de plus grand p_T (le "leading Z") ou de plus petit p_T ("subleading Z"), nous utilisons une fonction de Landau, tandis que pour les distributions de rapidité il s'agit d'une fonction gaussienne. La forme est ensuite normalisée au bruit de fond total présenté dans la table 6.5. Nous obtenons finalement, pour une luminosité intégrée de 5.02 fb^{-1} , les distributions regroupées dans les figures 6.3 à 6.7.

Nous observons un bon accord dans l'ensemble entre les données et la prédiction du signal ZZ. Nous avons en effet, tous canaux confondus, $54.5 \pm 0.3(\text{stat}) \pm 4.8(\text{syst})$

événements attendus. L'influence de la récupération des photons de FSR est très faiblement visible sur les prédictions Monte-Carlo. Un cas est cependant observé dans les données dans le canal $ZZ \rightarrow 4\mu$ et ses caractéristiques sont détaillées dans la table 6.6. Les nombres d'événements attendus et observés sont présentés dans la table 6.7.

	avant correction FSR	après correction FSR		$\gamma_{FSR} E_T [\text{GeV}]$	$\gamma_{FSR} dR$
$m_{Z_1} [\text{GeV}.c^{-2}]$	88.05	88.05			
$m_{Z_2} [\text{GeV}.c^{-2}]$	74.84	95.69		15.34	0.046
$m_{4l} [\text{GeV}.c^{-2}]$	199.21	220.759			

Table 6.6 – *Effet observé sur les données de la correction FSR et caractéristiques du photon de FSR.*

Canal de désintégration	attendu ZZ	Bruit de fond estimé	Total attendu	Observés
4μ	$15.40 \pm 0.05 \pm 1.43$	$0.52 \pm 0.26 \pm 0.25$	$15.92 \pm 0.26 \pm 1.45$	14
$4e$	$10.50 \pm 0.04 \pm 0.95$	$0.25 \pm 0.14 \pm 0.07$	$10.75 \pm 0.14 \pm 0.95$	9
$2e2\mu$	$26.20 \pm 0.1 \pm 2.41$	$0.58 \pm 0.18 \pm 0.23$	$26.78 \pm 0.17 \pm 2.41$	31

Table 6.7 – *Événements signal ZZ attendus, estimation du bruit de fond par la technique data-driven, et nombre d'événements observés pour chacun des canaux de désintégration. Les nombres correspondent à une luminosité intégrée de 5.02 fb^{-1} . Les premières erreurs sont statistiques et les secondes sont systématiques.*

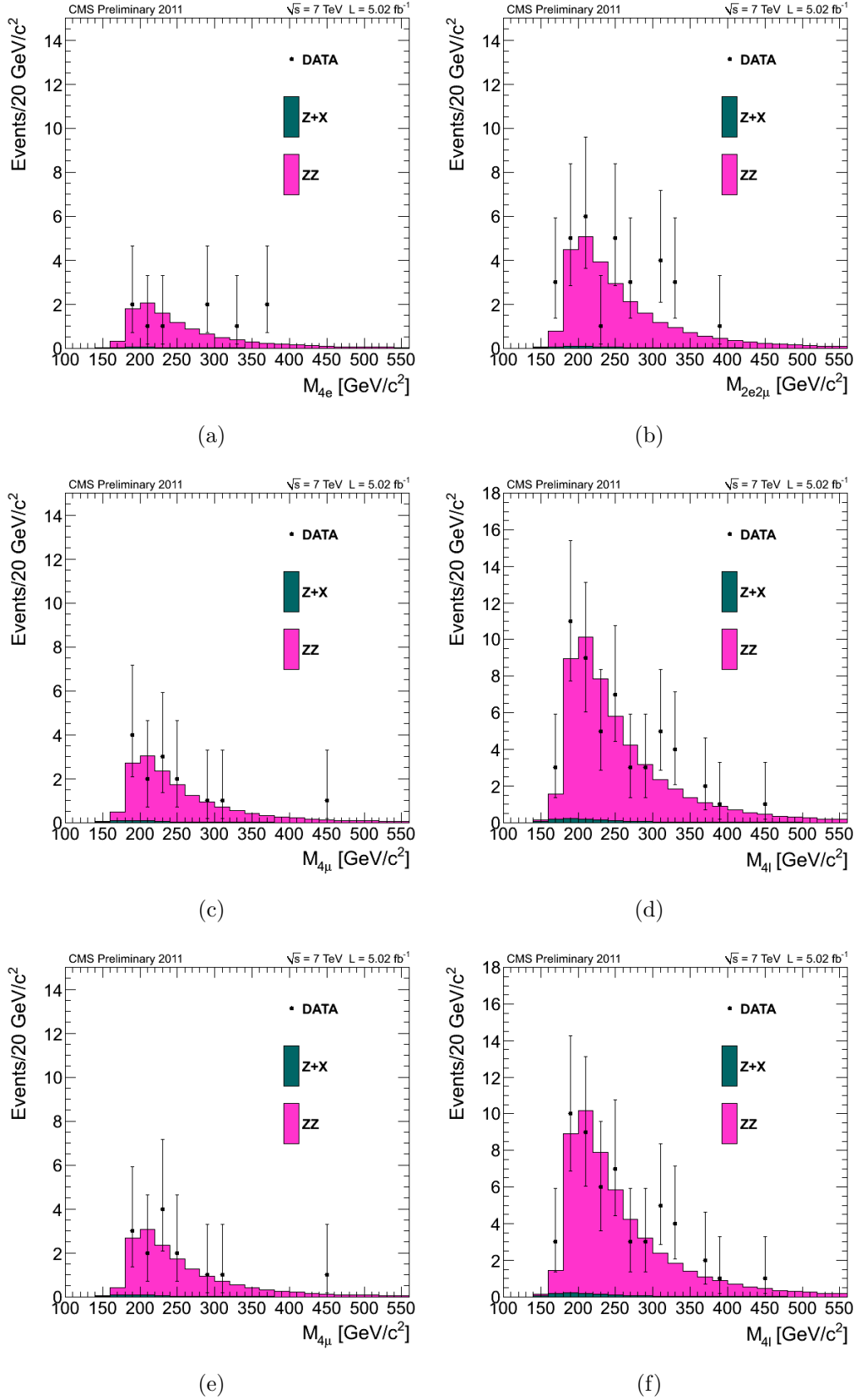


Figure 6.3 – Distribution de la masse invariante 4 leptons pour les canaux (a) $4e$, (b) $2e2\mu$, (c) 4μ et (d) tous canaux, (e) et (f) sont les versions corrigées de FSR pour 4μ et tous canaux respectivement.

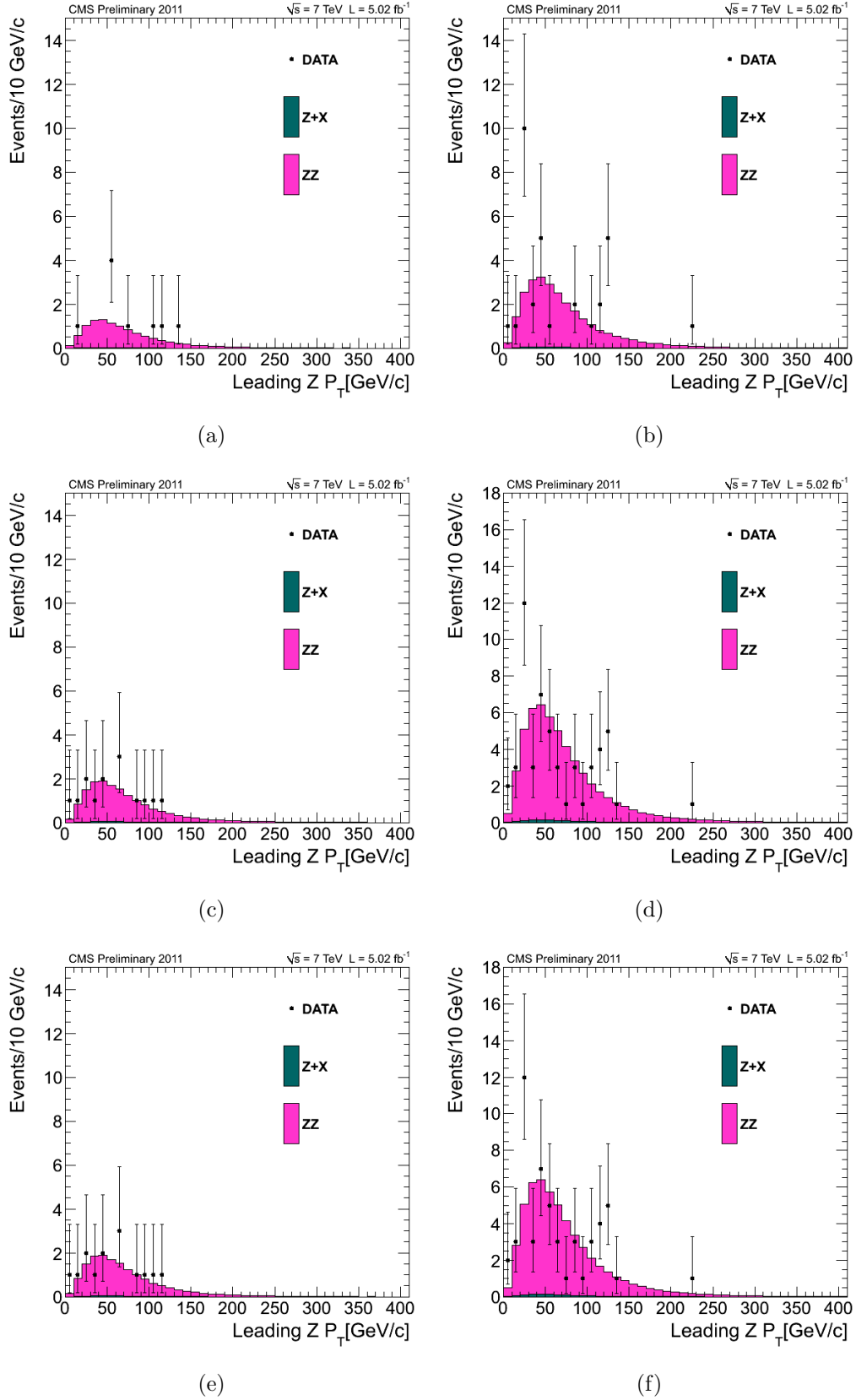


Figure 6.4 – Distribution de l'impulsion transverse du Z de plus haut p_T pour les canaux (a) $4e$, (b) $2e2\mu$, (c) 4μ et (d) tous canaux, (e) et (f) sont les versions corrigées de FSR pour 4μ et tous canaux respectivement.

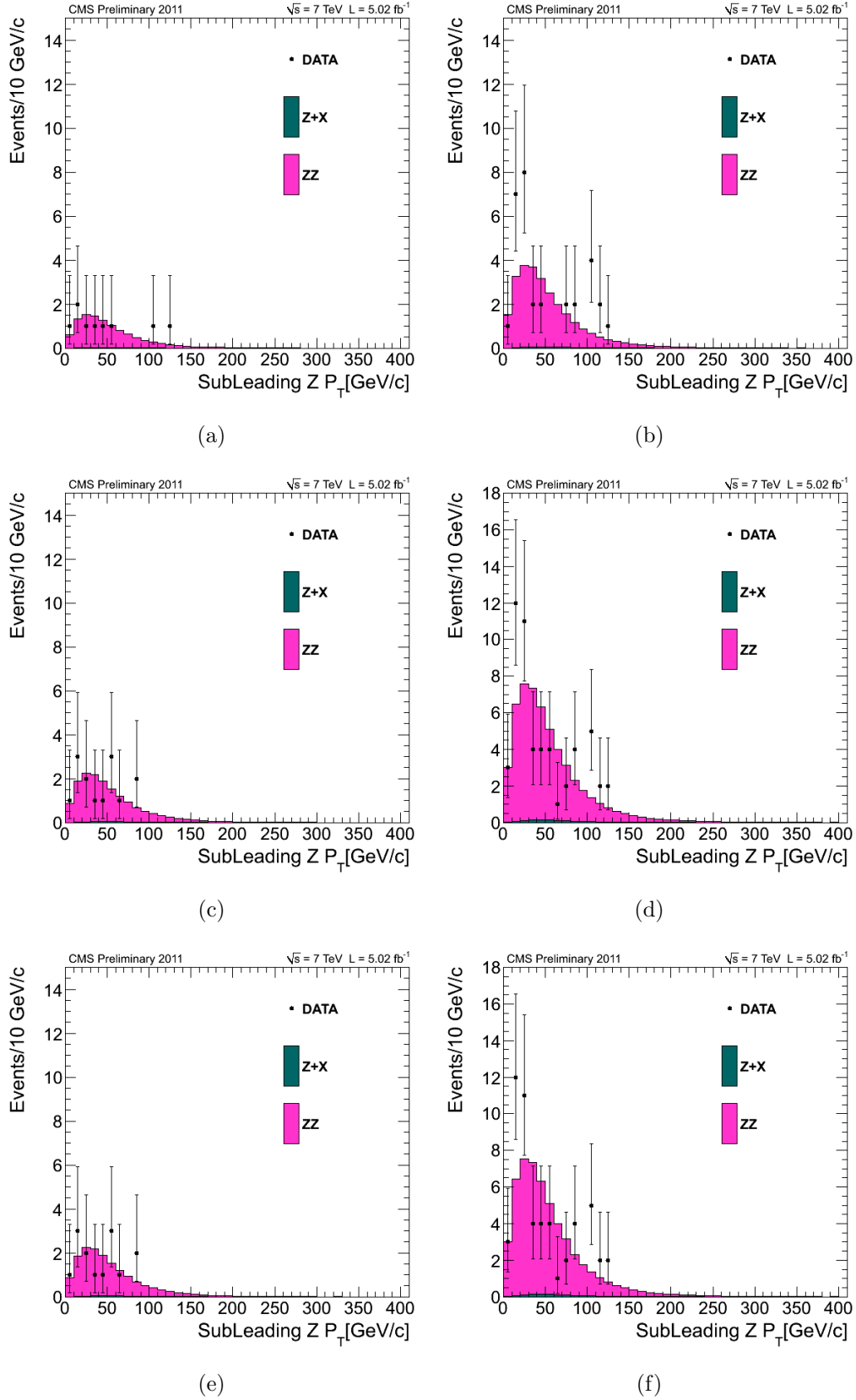


Figure 6.5 – Distribution de l'impulsion transverse du Z de plus bas p_T pour les canaux (a) $4e$, (b) $2e2\mu$, (c) 4μ et (d) tous canaux, (e) et (f) sont les versions corrigées de FSR pour 4μ et tous canaux respectivement.

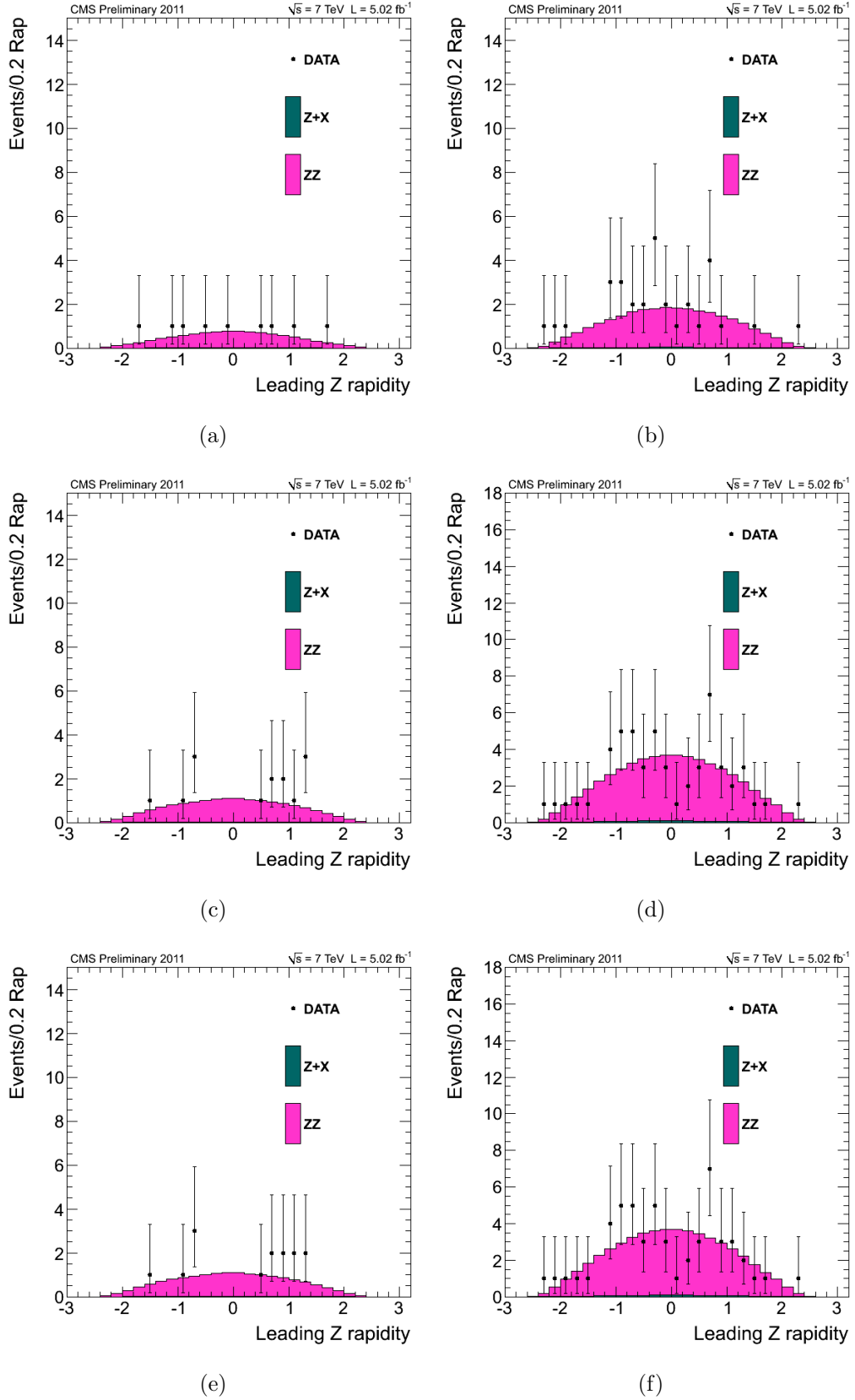


Figure 6.6 – Distribution de la rapidité du Z de plus haut p_T pour les canaux (a) $4e$, (b) $2e2\mu$, (c) 4μ et (d) tous canaux, (e) et (f) sont les versions corrigées de FSR pour 4μ et tous canaux respectivement.

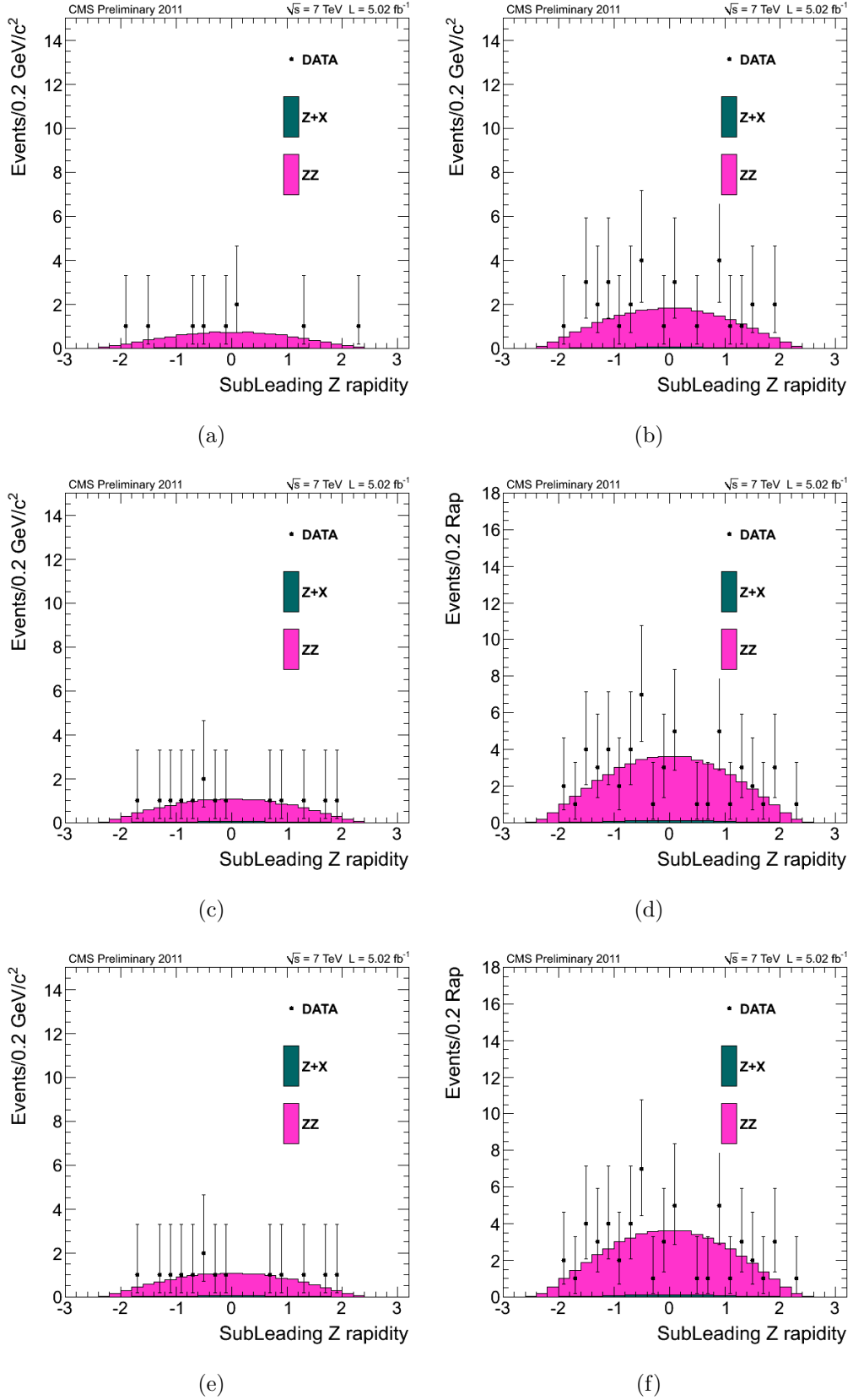


Figure 6.7 – Distribution de la rapidité du Z de plus bas p_T pour les canaux (a) $4e$, (b) $2e2\mu$, (c) 4μ et (d) tous canaux, (e) et (f) sont les versions corrigées de FSR pour 4μ et tous canaux respectivement.

6.2 Recherche de couplages trilinéaires neutres

L'étude des couplages anormaux entre bosons de jauge prend pour base l'analyse ZZ présentée dans la partie précédente. Les couplages anormaux ne faisant pas partie du Modèle Standard, ils sont décrits dans le Lagrangien effectif par deux coefficients f_4^Z et f_5^Z et leur présence se traduirait par un changement dans la forme des distributions de variables telles que la masse invariante quatre leptons ou encore l'impulsion transverse des candidats Z. Au niveau de la simulation de modèles intégrant des couplages anormaux entre trois bosons, il est possible d'introduire des facteurs de forme afin de prévenir la violation d'unitarité à haute énergie par ces couplages anormaux. Les facteurs de forme sont définis par :

$$f_i^Z = \frac{f_{i0}^Z}{\left(1 + \frac{\hat{s}}{\Lambda^2}\right)^n} \quad (6.5)$$

avec :

- $\hat{s}$, le carré de la masse invariante du système di-bosons.
- Λ , le facteur d'échelle indiquant la limite au-dessus de laquelle la présence de nouvelle physique au-delà du Modèle Standard devrait apparaître afin de solutionner les divergences à hautes énergies.
- f_{i0}^Z , l'approximation à basse énergie des couplages anormaux
- n , la puissance du facteur de forme

Cependant étant donné la nature arbitraire de la dépendance en énergie introduite, la collaboration CMS a décidé de ne pas utiliser de facteurs de forme. Par ailleurs les résultats sans facteur de forme sont exacts dans la mesure où l'échelle de la nouvelle physique est très grande devant l'énergie des collisions dans le centre de masse. De plus les résultats obtenus sont alors comparables entre eux.

Nous allons dans un premier temps justifier le choix du générateur Monte-Carlo afin de simuler différents jeux de couples (f_4^Z, f_5^Z) . Nous introduirons ensuite la technique de repondération des distributions ayant pour but de permettre de générer beaucoup plus de points qu'avec un processus de simulations et reconstruction complet de tous les couples (f_4^Z, f_5^Z) . Puis nous discuterons parmi les variables sensibles aux effets des couplages anormaux, le choix de la variable qui sera utilisée pour la recherche de ces couplages. Enfin nous présenterons les résultats obtenus avec les données de 2010 et 2011 correspondant à une luminosité intégrée de 5.02 fb^{-1} .

6.2.1 Méthodologie

6.2.1.1 Choix et validation du générateur

Comme nous l'avons introduit dans la première partie de ce chapitre, il existe différents générateurs Monte-Carlo pour simuler la physique au premier (LO) ou au second (NLO) ordre ; chacun ayant ses propres qualités et défauts. Les générateurs LO+ n jets ou de second ordre, tels que Madgraph ou Powheg utilisés par CMS pour simuler la physique électrofaible du Modèle Standard, n'implémentent pas les couplages anormaux entre trois bosons Z. Par conséquent notre choix s'est porté sur le seul générateur pouvant simuler des événements de signal aTGC : Sherpa. Ce-dernier est un générateur LO+ n jets et sera

utilisé dans sa version 1.2.2 [35] interfacée avec le programme de simulation et de reconstruction de CMS : CMSSW (cf. la section 2.3). Nous suivrons également le choix fait par la collaboration CMS de ne pas ajouter de facteur de forme en utilisant une valeur infinie pour l'échelle Λ en paramètre d'entrée du générateur.

Le choix du générateur présenté et justifié, nous allons maintenant le valider. Pour ce faire nous comparons les distributions issues des générateurs Sherpa et Powheg pour la production de di-bosons ZZ dans le Modèle Standard. Sherpa est utilisé ici dans une configuration au premier ordre plus 1 jet (LO+n jets, $n \leq 1$).

En observant la figure 6.8, nous remarquons que les distributions de masse invariante

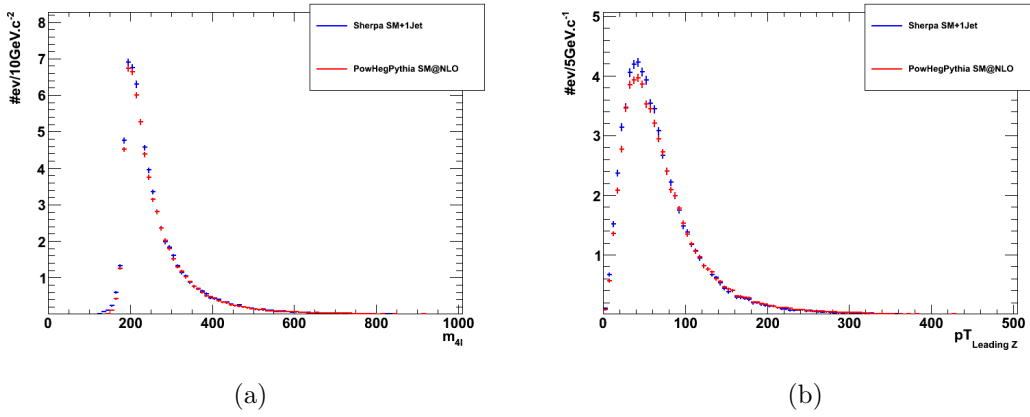


Figure 6.8 – Distributions de la masse invariante 4 leptons (a) et de l'impulsion transverse du Z de plus haut p_T (b) dans le canal $ZZ \rightarrow 4\mu$, sans effet de la reconstruction, pour le générateur Sherpa dans une configuration modèle standard +1 jet (bleu) et pour le générateur Powheg servant de référence (rouge).

du système formée par les quatre muons sont très similaires dans les cas Sherpa+1jet et Powheg. La configuration "+1 jet" du générateur Sherpa s'approchant au mieux de la distribution obtenue par simulation au second ordre avec Powheg et servant de référence, nous ne considérerons plus que ce cas pour la suite de l'étude. Cette compatibilité des distributions nous permet également de valider le choix de Sherpa par sa fiabilité à simuler la production de di-bosons ZZ selon le Modèle Standard.

6.2.1.2 Méthode de repondération des distributions

La mesure des couplages anormaux ou plus précisément la recherche de l'existence de couplages anormaux via l'établissement de limites sur ces-derniers se fait, comme nous le détaillerons plus loin, à l'aide d'une évaluation fréquentiste des niveaux de confiance en prenant deux hypothèses : la première étant l'hypothèse signal caractérisée par une simulation Monte-Carlo d'un modèle aTGC tandis que la deuxième est l'hypothèse du Modèle Standard. Nous évaluons alors le niveau de confiance pour les données observées de satisfaire l'hypothèse signal. Afin de sonder les limites des couplages nous devons dis-

poser d'une grille de points simulés dans le plan (f_4^Z, f_5^Z) .

Deux méthodes se proposent à nous :

- la première consiste en une simulation par Sherpa suivie du processus de reconstruction complet des événements, implémentant l'interaction des particules avec les différents éléments du détecteur CMS ou encore la présence de phénomènes d'empilement. Cette technique est longue, essentiellement à cause de la partie reconstruction très coûteuse en temps de calcul, et par conséquent limite le nombre de couples (f_4^Z, f_5^Z) que nous pourrions considérer.
- la seconde est la méthode dite "de repondération". Avec cette technique, les événements sont seulement simulés au niveau générateur et le rapport de la distribution pour un modèle (f_4^Z, f_5^Z) par la distribution pour le Modèle Standard ($f_4^Z = 0, f_5^Z = 0$) est utilisé pour repondérer les distributions des événements simulés selon le Modèle Standard et complètement reconstruits issus de l'échantillon de référence Powheg. En nous épargnant le long processus de simulation complète et reconstruction nous pouvons alors générer une grille de points (f_4^Z, f_5^Z) beaucoup plus fine. C'est cette technique que j'ai retenue pour la suite de mon étude.

Dans le but d'illustrer et de valider cette technique, j'ai procédé à la simulation, suivie du processus complet de reconstruction des événements, d'un échantillon signal aTGC donné ($f_4^Z = 0.05, f_5^Z = 0.05$). J'ai ensuite comparé la distribution de l'échantillon Modèle Standard Powheg repondérée par le rapport entre modèle aTGC et Modèle Standard avec celle obtenue après reconstruction par Sherpa dans une configuration aTGC.

Le résultat est présenté dans la figure 6.9. Nous pouvons constater un bon accord entre

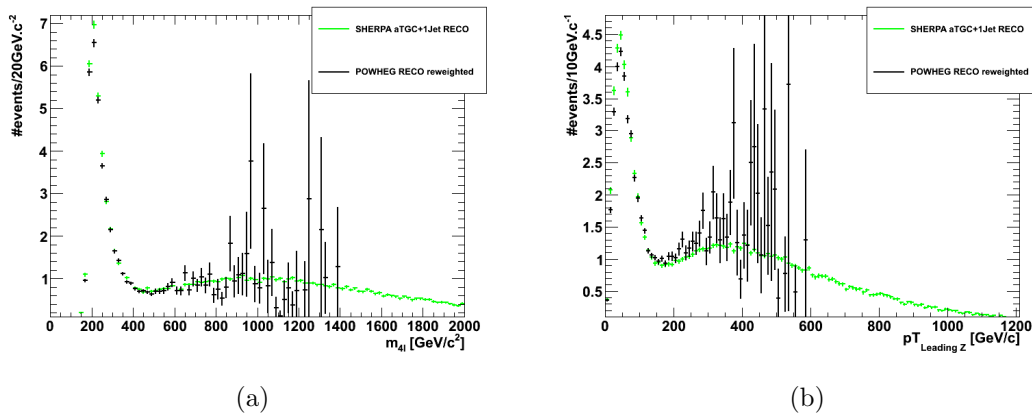


Figure 6.9 – Distribution de la masse invariante 4 leptons (a) et de l'impulsion transverse du Z de plus haut p_T (b) dans le canal $ZZ \rightarrow 4l$, repondération appliquée aux événements modèle standard générés et reconstruits de l'échantillon Powheg (noir) et signal aTGC généré et reconstruit de l'échantillon Sherpa.

la distribution Modèle Standard repondérée et la distribution signal aTGC obtenue en

utilisant la chaîne de simulation-reconstruction complète nous permettant de valider cette méthode.

6.2.1.3 Choix de la variable discriminante

Parmi les différentes observables cinématiques dont les distributions sont sensibles aux couplages anormaux, je ne retiens pour cette étude que la masse invariante m_{4l} et l'impulsion transverse du Z de plus haut p_T : $p_T^{LeadingZ}$. Avec la figure 6.10, nous remarquons que l'effet des couplages anormaux se manifeste dans les queues de distribution que ce soit à haute masse dans le spectre de masse invariante quatre leptons ou bien à haut p_T concernant la distribution de l'impulsion transverse $p_T^{LeadingZ}$. L'objectif ici est de déterminer le choix final de la variable discriminante qui sera utilisée pour la recherche de couplage anormaux. Ce choix reposera sur le critère de sensibilité aux couplages anormaux mais également d'accord entre la distribution Modèle Standard repondérée et celle du signal aTGC.

La sensibilité aux couplages anormaux est évaluée à l'aide d'un test de signifiacnce formée

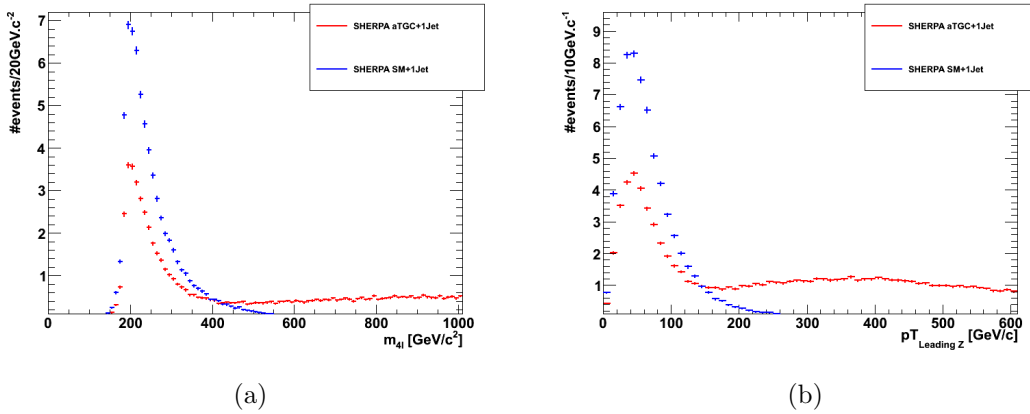


Figure 6.10 – Distribution de la masse invariante 4 leptons (a) et de l'impulsion transverse du Z de plus haut p_T (b) dans le canal $ZZ \rightarrow 4l$, modèle standard Sherpa en bleu et signal aTGC Sherpa en rouge.

par un rapport des vraisemblances pour l'hypothèse signal aTGC ($hyp1$) et l'hypothèse Modèle Standard ($hyp0$). Le rapport des vraisemblances pour des distributions suivant une statistique de Poisson s'exprime par :

$$Q = \left(\frac{N_{hyp1}}{N_{hyp0}} \right)^{N_i} \cdot \frac{e^{-N_{hyp1}}}{e^{-N_{hyp0}}} \quad (6.6)$$

Avec

- N_i , le nombre d'événements dans le $i^{ème}$ bin de la distribution considérée (variable m_{4l} ou $p_T^{LeadingZ}$) pour l'hypothèse aTGC, i.e. l'hypothèse pour laquelle nous cherchons à évaluer la sensibilité de cette variable.
- N_{hyp1} , le nombre total d'événements attendus selon l'hypothèse signal aTGC.

– N_{hyp0} , le nombre total d'événements attendus selon l'hypothèse Modèle Standard. La signification est finalement donnée par la formule :

$$significance = \sqrt{2\ln(Q)} \quad (6.7)$$

En observant la figure 6.11 représentant la signification pour les variables m_{4l} et $p_T^{LeadingZ}$ en fonction d'une coupure sur ces variables, nous constatons qu'elles sont comparables, avec une sensibilité cependant légèrement plus importante dans le cas $p_T^{LeadingZ}$. Notre second

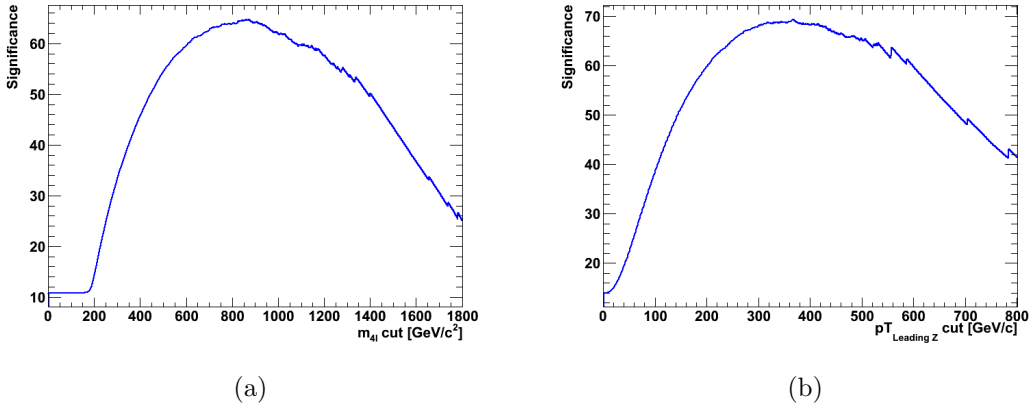


Figure 6.11 – Signification pour la masse invariante 4 leptons (a) et l'impulsion transverse du Z de plus haut p_T (b), en fonction de la coupure.

critère de sélection reposant sur l'accord entre les distributions pour le Modèle Standard repondérées issues de l'échantillon Powheg et pour le signal aTGC généré par Sherpa, déjà présenté dans la figure 6.9, est quantifié par la figure 6.12. Nous pouvons voir sur cette figure que la prédiction Powheg repondérée est en très bon accord avec la distribution signal aTGC obtenue avec Sherpa pour la variable m_{4l} tandis que la distribution $p_T^{LeadingZ}$ présente une différence de forme entre simulation au premier ordre (LO) et au second (NLO).

Finalement pour des sensibilités aux effets aTGC relativement similaires entre m_{4l} et $p_T^{LeadingZ}$, les prédictions aTGC Sherpa présentent un accord bien meilleur avec Powheg repondéré dans le cas m_{4l} et nous choisissons donc la masse invariante m_{4l} comme variable discriminante pour la suite de cette étude.

6.2.2 Interprétation statistique, modèles simulés et résultats

Nous allons maintenant aborder la recherche d'existence de couplages anormaux de bosons triples ZZZ. Une existence éventuelle se manifeste au travers des déviations par rapport au Modèle Standard dans les distributions cinématiques comme cela a été illustré dans la figure 6.10. Nous avons préalablement justifié la procédure d'analyse via la méthode de repondération des distributions, le choix du générateur Sherpa, ainsi que celui de la variable discriminante m_{4l} .

Nous allons donc, dans un premier temps, présenter la technique, dite du CL_s , servant

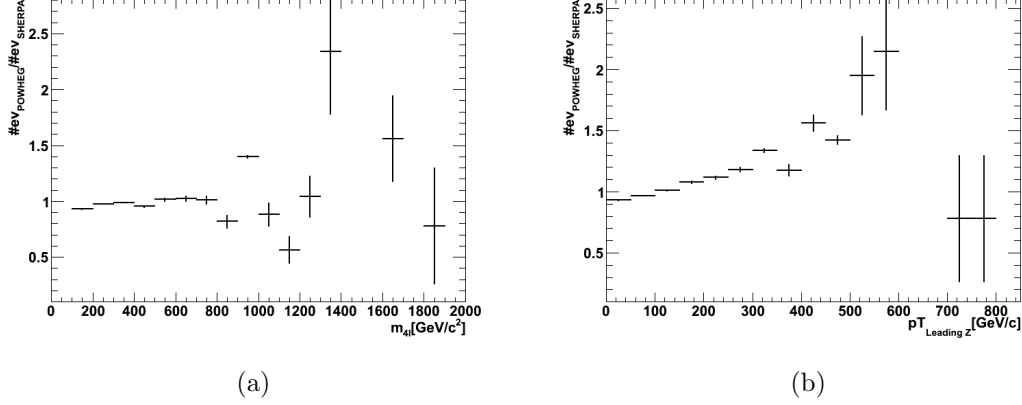


Figure 6.12 – Rapport des distributions masse invariante 4 leptons (a) et impulsion transverse du Z de plus haut p_T (b)

à estimer et établir dans notre cas une limite sur l’existence de ces couplages. Puis nous détaillerons les simulations faites et exposerons les résultats obtenus que nous discuterons notamment au travers d’une comparaison avec les résultats publics de la collaboration CMS ainsi que celui de la collaboration ATLAS.

6.2.2.1 Méthode CL_s

La procédure utilisée pour tester la compatibilité des mesures avec l’hypothèse de l’existence de couplages trilineaires anormaux emploie la méthode fréquentiste [36] et la forme d’une distribution, dans notre cas la distribution de la masse invariante du système 4 leptons. Nous devons ainsi estimer le niveau de confiance que nous portons à notre résultat d’être compatible avec :

- l’hypothèse **signal**, dans notre cas l’hypothèse aTGC. Dans le cadre de la méthode CL_s cette hypothèse est usuellement notée $s + b$. Étant donné qu’il s’agit ici non pas d’un signal noyé dans un bruit de fond mais plutôt de deux modèles différents nous lui préférons la notation hyp_1 .
- l’hypothèse **absence de signal**, dans notre cas l’hypothèse Modèle Standard seul. Dans le cadre de la méthode CL_s cette hypothèse est usuellement notée b . Nous la noterons hyp_0 .

Nous avons déterminé en partie 6.2.1.3 la variable discriminante à savoir la masse reconstituée des quatre leptons. Nous calculons ensuite le rapport de vraisemblance $Q = \frac{\mathcal{L}(hyp_\alpha)}{\mathcal{L}(hyp_\beta)}$ ou plus précisément la grandeur $-2\ln(Q)$ définie de manière générale à partir d’un histogramme de masse quatre leptons de pas i pour un observé obs et deux hypothèses quelconques hyp_α et hyp_β par :

$$-2\ln(Q) = 2(N_{hyp_\alpha} - N_{hyp_\beta}) - 2 \sum_i N_i^{obs} \cdot \ln \left(\frac{N_{hyp_\alpha, i}}{N_{hyp_\beta, i}} \right) \quad (6.8)$$

Nous construisons ensuite la fonction densité de probabilité (ou p.d.f. pour probability density function) du rapport de vraisemblance pour un ensemble d'expériences aléatoires pour les hypothèses Modèle Standard (hyp_0) et aTGC (hyp_1). Puis nous évaluons la grandeur $-2\ln(Q)$ en utilisant les nombres N_i^{obs} observés dans les données. Les niveaux de confiance relatifs aux hypothèses testées peuvent alors être déterminés et finalement le niveau de confiance est $CL = 1 - CL_s$ avec :

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{CL_{hyp_1}}{CL_{hyp_0}} \quad (6.9)$$

Nous allons maintenant illustrer plus en détail ces étapes de calcul. Considérons tout d'abord l'hypothèse hyp_0 , Modèle Standard. Le rapport de vraisemblance dans ce cas est défini par

$$-2\ln(Q) = 2(N_{hyp_1} - N_{hyp_0}) - 2 \sum_i^{N_{alea_0}} N_{hyp_0,i}^{obs} \cdot \ln \left(\frac{N_{hyp_1,i}}{N_{hyp_0,i}} \right) \quad (6.10)$$

avec $N_{hyp_0,i}^{obs}$, le nombre d'entrées du bin i d'un histogramme construit en tirant aléatoirement N_{alea_0} entrées selon l'allure de l'histogramme de l'hypothèse hyp_0 ayant un total d' N_{hyp_0} entrées. Notons que le nombre N_{alea_0} est également tiré aléatoirement selon une loi de Poisson d'espérance N_{hyp_0} . Nous élaborons ensuite la fonction densité de probabilité de $-2\ln(Q)$ dans le cas de l'hypothèse hyp_0 en procédant à un grand nombre de tirages aléatoires. Nous obtenons la distribution bleue sur la figure 6.13.

Nous procédons de même pour le cas de l'hypothèse aTGC hyp_1 . Le rapport de vraisemblance s'exprime dans ce cas par :

$$-2\ln(Q) = 2(N_{hyp_1} - N_{hyp_0}) - 2 \sum_i^{N_{alea_1}} N_{hyp_1,i}^{obs} \cdot \ln \left(\frac{N_{hyp_1,i}}{N_{hyp_0,i}} \right) \quad (6.11)$$

avec $N_{hyp_1,i}^{obs}$, le nombre d'entrées du bin i d'un histogramme construit en tirant aléatoirement N_{alea_1} entrées selon l'allure de l'histogramme de l'hypothèse hyp_1 ayant un total d' N_{hyp_1} entrées. Ici aussi le nombre N_{alea_1} est tiré aléatoirement selon une loi de Poisson d'espérance N_{hyp_1} . La fonction densité de probabilité de $-2\ln(Q)$ dans le cas de l'hypothèse hyp_1 est représentée par la distribution rouge sur la figure 6.13.

Enfin nous calculons $-2\ln(Q)$ mesurée avec les données CMS :

$$-2\ln(Q) = 2(N_{hyp_1} - N_{hyp_0}) - 2 \sum_i^{N_{tot\data}} N_{data,i}^{obs} \cdot \ln \left(\frac{N_{hyp_1,i}}{N_{hyp_0,i}} \right) \quad (6.12)$$

Cela nous donne un seul point de mesure représenté par le trait noir sur la figure 6.13. L'intégrale de la p.d.f. $_{hyp_0}$ bleue à gauche de la valeur observée évalue la probabilité p_{hyp_0} que l'hypothèse Modèle Standard fluctue et fournisse un attendu compatible avec la présence de couplages anormaux. Le niveau de confiance sur la compatibilité de la mesure avec l'hypothèse Modèle Standard est donnée par $CL_{hyp_0} = 1 - p_{hyp_0}$. L'intégrale de la p.d.f. $_{hyp_1}$ à droite de l'observé nous donne la probabilité p_{hyp_1} que l'hypothèse hyp_1 de présence de couplage anormaux soit vérifiée ou en d'autres termes le niveau de confiance $CL_{hyp_1} = p_{hyp_0}$ quant à la mesure d'être compatible avec l'hypothèse hyp_1 de présence de couplages anormaux.

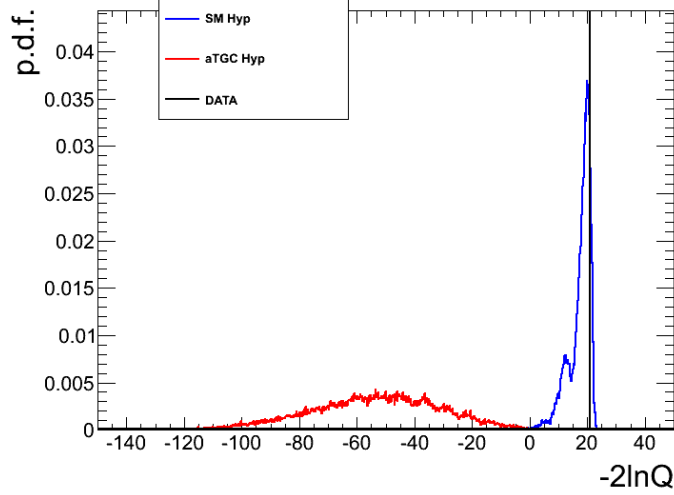


Figure 6.13 – Exemple de p.d.f obtenues pour les hypothèses hyp_0 (bleu), hyp_1 (rouge) et pour les données (noir). L’hypothèse hyp_1 a été simulée ici avec $f_4^Z=0.015$ et $f_5^Z=0.015$

6.2.2.2 Limites d’exclusion

Des mesures de l’auto-couplage entre bosons neutres et en particulier pour les bosons Z ont été effectuées précédemment dans les expériences au LEP, au Tevatron et au LHC. Les résultats obtenus jusqu’à maintenant sont en accord avec le Modèle Standard et les limites les plus contraignantes à ce jour $0.07 < f_4^Z < 0.07$ et $0.07 < f_5^Z < 0.07$ ont été extraites par la collaboration ATLAS [37]. Le travail présenté dans cette section a été réalisé à l’intérieur du groupe en charge de la publication en cours de finalisation qui s’est mis en place dans CMS à partir de mars 2012.

Pour procéder à la mesure des couplages anormaux ZZZ et en accord avec la méthodologie exposée dans la partie 6.2.1, je procède tout d’abord à la simulation d’échantillons paramétrés selon une grille de différentes valeurs de couplage allant de -0.025 à 0.025 par pas de 0.01 . Les sections efficaces, en fb, des échantillons générés sont regroupées dans la table 6.8.

$f_4^Z \setminus f_5^Z$	-0.025	-0.015	-0.005	0.005	0.015	0.025
-0.025	15.3418	14.72	14.38	14.37	14.69	15.29
-0.015	14.68	14.34	13.70	13.71	14.05	14.76
-0.005	14.34	13.70	13.37	13.39	13.69	14.39
0.005	14.41	13.71	13.40	13.35	13.69	14.34
0.015	14.68	14.02	13.70	13.72	14.04	14.69
0.025	15.29	14.67	14.40	14.35	14.73	15.29

Table 6.8 – Tableau des sections efficaces de production [fb] pour les différentes valeurs f_4^Z et f_5^Z simulées.

$f_4^Z \setminus f_5^Z$	-0.025	-0.015	-0.005	0.005	0.015	0.025
-0.025	100	100	100	100	100	100
-0.015	100	100	99.8	99.7	100	100
-0.005	100	99.8	76.2	65.1	99.5	100
0.005	100	99.7	71.5	68.9	99.7	100
0.015	100	100	99.9	99.6	100	100
0.025	100	100	100	100	100	100

Table 6.9 – Tableau des CL pour différentes valeurs f_4^Z et f_5^Z simulées après correction FSR.

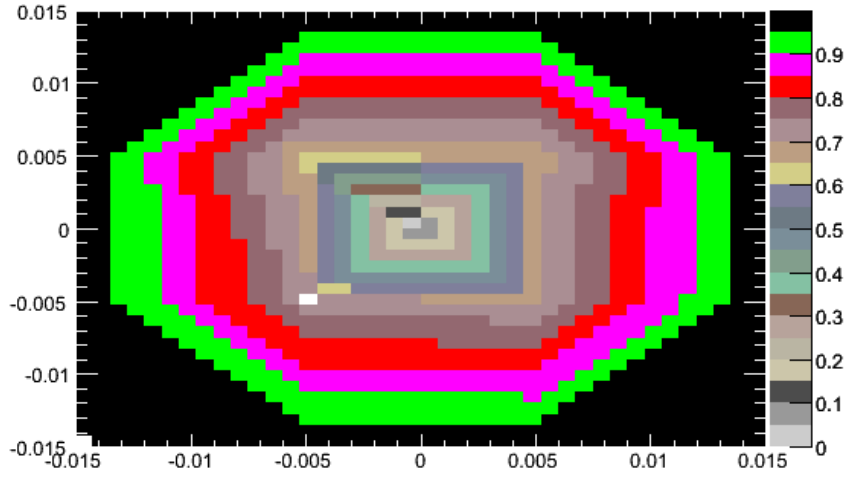


Figure 6.14 – Contour CL approximé par une interpolation de Delaunay.

J'ai ensuite calculé les niveaux de confiance selon la méthode CLs (cf section 6.2.2.1) en procédant à 50000 pseudo-expériences pour chacune des valeurs de couple (f_4^Z, f_5^Z) simulées. Un affichage graphique via *root* des résultats obtenus nous fournit, à l'aide d'une interpolation de Delaunay, un contour grossier de la limite à 95% CL. Le contour obtenu, visible sur la figure 6.14, associé à la table de valeur d'exclusion 6.9 nous permet de déduire que la limite 2D se situe pour des valeurs de $|f_4^Z| \leq 0.0135$ et $|f_5^Z| \leq 0.0135$. Par ailleurs nous formulons l'hypothèse que la limite est symétrique : $f_{4,5}^Z \rightarrow -f_{4,5}^Z$. En effet les limites sont mesurées à partir de taux de production di-bosons dépendant de manière quadratique des couplages f_4^Z et f_5^Z . Cette hypothèse, ainsi que la première limite observée, me permettent de procéder à un maillage encore plus fin en me limitant aux valeurs positives inférieures ou égales à 0.015 de f_4^Z par exemple. L'ajustement par une fonction parabolique, présenté dans la figure 6.15, des points de niveau de confiance obtenus, et regroupés dans la table 6.10, nous permet d'extraire la limite à 95(%) de CL pour le couplage f_4^Z : $|f_4^Z| < 0.0116$. Nous remarquons également que la correction FSR n'a qu'un effet très limité, compte tenu qu'elle ne concerne qu'environ 1% des événements. Nous attendons un effet du même ordre sur les limites. Les limites à une dimension

extraites à 95% de niveau de confiance sont, sans et avec correction FSR :

$$- 0.00115 < f_4^Z < 0.00115 \quad (\text{sans correction FSR}) \quad (6.13)$$

$$- 0.00116 < f_4^Z < 0.00116 \quad (\text{avec correction FSR}) \quad (6.14)$$

f_4^Z	CL (%)	CL _{FSR} (%)
0.005	50.85	50.95
0.007	73.93	72.98
0.010	92.73	92.33
0.011	94.19	94.27
0.012	95.91	95.78
0.013	97.97	98.00
0.014	99.12	99.24
0.015	99.50	99.46

Table 6.10 – Tableau des CL pour différentes valeurs f_4^Z simulées.

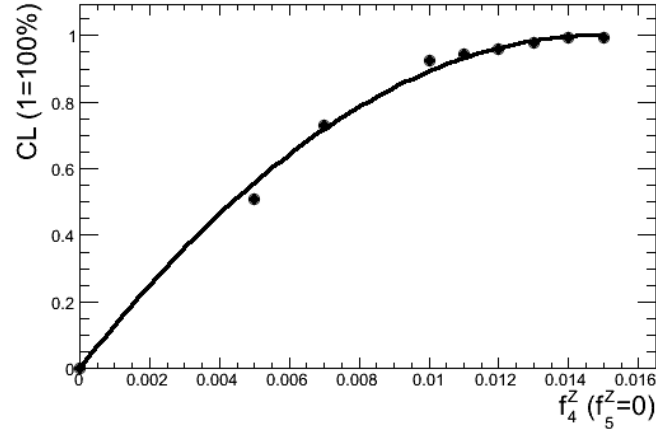


Figure 6.15 – Courbe d'ajustement des niveaux de confiance CL pour les différents points f_4^Z simulés avec correction FSR.

En procédant de même pour f_5^Z , j'obtiens :

$$- 0.00119 < f_5^Z < 0.00119 \quad (\text{avec correction FSR}) \quad (6.15)$$

Nous allons maintenant comparer ces limites avec les autres résultats obtenus au LHC, que ce soit le résultat officiel de la collaboration CMS ou celui d'ATLAS, comprendre les différences et conclure. La table 6.11 présente les résultats actuels et publics des collaborations ATLAS et CMS, ainsi que le résultat obtenu dans cette thèse. La collaboration ATLAS ayant publié deux résultats, un pour un facteur de forme avec $\Lambda = 2$ TeV et un sans facteur de forme (Λ infini), le résultat mentionné ici est celui sans facteur de forme de façon à pouvoir comparer avec les résultats de CMS.

Expérience	f_4^Z	f_5^Z	Publication
ATLAS	$[-0.07 ; +0.07]$	$[-0.07 ; +0.07]$	[37]
CMS	$[-0.011 ; +0.012]$	$[-0.012 ; +0.012]$	en cours
CMS	$[-0.0116 ; +0.0116]$	$[-0.0119 ; +0.0119]$	mon résultat

Table 6.11 – *Tableau comparatif des limites sur les couplages anormaux ZZZ pour CMS et ATLAS.*

En confrontant mon résultat à celui obtenu par la collaboration ATLAS, une source de différence réside dans la luminosité intégrée beaucoup plus faible utilisée pour les limites ATLAS ($\mathcal{L}_{\text{int}} = 1 \text{ fb}^{-1}$). Le résultat d'ATLAS est significativement moins contraignant que notre résultat, ce qui peut s'expliquer essentiellement par deux effets : cette moindre statistique d'une part mais aussi par le fait que le nombre total d'événements est considéré pour extraire les limites, et non la forme du spectre. Comme dans mon cas, ATLAS a procédé à une extraction des limites en fixant un des couplages à zéro. Cependant d'autres différences intrinsèques au détecteur et aux algorithmes ATLAS doivent être signalées :

- la couverture angulaire pour les muons est plus grande, $|\eta| < 2.5$, tandis que pour les électrons elle est légèrement plus petite $|\eta| < 2.47$.
- les leptons sont sélectionnés s'ils vérifient $p_T > 15 \text{ GeV.c}^{-1}$.
- les efficacités de reconstruction, identification et isolation respectives sont également différentes.
- la technique de simulation et reconstruction complète est utilisée.

En comparant mon résultat avec le résultat officiel de CMS, nous pouvons conclure à un très bon accord : les limites officielles sont $-0.011 < f_4^Z < 0.012$ et $-0.012 < f_5^Z < 0.012$ tandis que j'obtiens les valeurs très légèrement plus contraignantes de $-0.0116 < f_4^Z < 0.0116$ et de $-0.0119 < f_5^Z < 0.0119$. Les sources de la différence entre les deux sont :

- la version officielle utilise la méthode de simulation et reconstruction complète des événements avec couplages anormaux. Ceci implique la simulation de moins de points aTGC.
- le bruit de fond, très faible, est soustrait aux données dans l'analyse officielle.
- les limites officielles prennent en compte les systématiques
- aucune correction des pertes dues aux émissions de photons FSR n'est appliquée pour l'extraction des limites officielles.

Cependant le bruit de fond soustrait dans la version officielle est très faible et les systématiques incluses dans l'analyse officielle le sont également. De plus comme nous l'avons déjà mentionné, la correction FSR n'a qu'un effet très limité car elle ne concerne qu'environ 1% des événements. L'effet attendu sur l'extraction des limites est donc également faible. Enfin le résultat public de CMS repose sur un maillage plus grossier du plan (f_4^Z, f_5^Z) et une interpolation quadratique. Mes résultats présentés dans la figure 6.15, montrent que cette approximation est pleinement justifiée ce qui permet de comprendre que les limites obtenues sont très proches.

Conclusion

Le LHC a recueilli, en à peine trois ans, suffisamment de données pour sonder des domaines d'énergie jamais atteints et permettre des découvertes majeures dans le domaine de la physique des particules. Les données acquises ont également servies à mieux comprendre le détecteur CMS et le bon fonctionnement des algorithmes intervenant en amont des analyses car servant à reconstruire les objets physiques comme les électrons par exemple. Les différentes études menées dans cette thèse s'inscrivent complètement dans les défis et opportunités engendrés par le caractère exceptionnel des débuts d'une expérimentation de l'ampleur de CMS.

L'intégration de la combinaison des deux algorithmes de reconstruction des électrons a été réalisée et ses performances mesurées avant les premières collisions : les gains apportés par cette combinaison par rapport à l'algorithme originel sont notables à basse impulsion transverse ainsi que dans les régions difficiles du détecteur telles que les zones de transition tonneau-bouchons.

Pour faciliter la certification des électrons, un programme de contrôle a été développé, testé et mis en place dans l'interface de certification de CMS : le *DQMOffline*. Conjointement les distributions caractéristiques des électrons reconstruits ont été contrôlées avec l'attendu Monte-Carlo pour les premières données acquises de l'année 2010. Par ailleurs les efficacités de reconstruction des électrons ont été mesurées pour les années 2010 et 2011. Différentes techniques ont été étudiées ainsi que différentes sources de systématiques. Les efficacités ont été mesurées également selon différentes régions en pseudo-rapacité d'une part, et selon différentes fenêtres d'impulsion transverse d'autre part. Les valeurs obtenues ont servi de base aux analyses présentées lors des conférences des années 2010 et 2011.

L'effet des radiations dans l'état final (FSR) a été étudié, notamment sur le spectre de masse 4 leptons. Des algorithmes de récupération des photons de FSR ont été étudiés, testés et mis en place. L'algorithme retenu corrige 4.2% des événements sélectionnés par l'analyse ZZ tout en garantissant une grande pureté (95.6%) indépendante du phénomène d'empilement.

L'analyse $ZZ \rightarrow 4l$ et ses performances ont été décrites. Les distributions différentielles des Z du système di-boson dans la production $pp \rightarrow ZZ + X$ à $\sqrt{s}=7$ TeV sont pour la première fois montrées dans cette thèse. Enfin la recherche de couplage trilineaires neutres a été effectuée permettant l'extraction des limites les plus contraignantes à ce jour sur les valeurs de ces couplages : $-0.00116 < f_4^Z < 0.00116$ et $-0.00119 < f_5^Z < 0.00119$. Ces limites sont en accord avec les résultats de l'analyse officielle de CMS en cours de publication et confirment le Modèle Standard.

Bibliographie

- [1] M.A. Tavel. Milestones in mathematical physics : Noether's theorem. *Transp.Theory Statist.Phys.*, 1 :183–185, 1971.
- [2] S. Weinberg. *Phys.Rev.Lett.* (19) :1264, 1967.
- [3] S.L. Glashow. *Nucl.Phys.* (B-22) :579, 1961.
- [4] A. Salam, J.C. Ward. *Phys. Lett.* 168(13), 1964.
- [5] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13 :508–509, Oct 1964.
- [6] F. Englert, G.S. Guralnik *et al.* *Phys.Rev.Lett.* (13) :585, 1964.
- [7] M. Peskin and D. Schroeder. *An Introduction to Quantum Field Theory*. Westview, 1995.
- [8] K. Nakamura *et al.* (Particle Data Group). Review of particle physics. *J. Phys.*, G37 :075021, 2010.
- [9] Serguei Chatrchyan *et al.* Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716 :30–61, 2012.
- [10] J. Alcaraz. On the experimental effects of the off-shell structure in anomalous neutral triple gauge vertices. *Phys.Rev.*, D65 :075020, 2002.
- [11] K. Hagiwara, R.d. Peccei, D. Zeppenfeld. Probing the weak boson sector in $e^+e^- \rightarrow W^+W^-$. *Nuclear Physics*, B282(253-307), 1987.
- [12] S. Schael *et al.* Measurement of Z-pair production in e^+e^- collisions and constraints on anomalous neutral gauge couplings. *JHEP*, 0904 :124, 2009.
- [13] J. Abdallah *et al.* Study of triple-gauge-boson couplings ZZZ, ZZgamma and Zgamma gamma LEP. *Eur.Phys.J.*, C51 :525–542, 2007.
- [14] M. Acciarri *et al.* Measurement of triple gauge boson couplings of the W boson at LEP. *Phys.Lett.*, B467 :171–184, 1999.
- [15] G. Abbiendi *et al.* Study of Z pair production and anomalous couplings in e^+e^- collisions at $\sqrt{s}^*(1/2)$ between 190-GeV and 209-GeV. *Eur.Phys.J.*, C32 :303–322, 2003.
- [16] J. Alcaraz *et al.* A Combination of preliminary electroweak measurements and constraints on the standard model. 2006.
- [17] Aidan Robson. Diboson Physics at the Tevatron. *EPJ Web Conf.*, 28 :06001, 2012.
- [18] V.M. Abazov *et al.* Search for ZZ and $Z\gamma^*$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV and limits on anomalous ZZZ and $ZZ\gamma^*$ couplings. *Phys.Rev.Lett.*, 100 :131801, 2008.

- [19] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3 :S08004, 2008.
- [20] <http://root.cern.ch/drupal/>.
- [21] C. Charlot. La reconstruction des électrons dans CMS. <http://cmsfrance.in2p3.fr/HabilitationCCharlot.pdf>, janvier 2007.
- [22] E. Meschi *et al.* Electron Reconstruction in the CMS Electromagnetic Calorimeter. *CMS AN*, 2001/034, 2001.
- [23] M. Anderson. A Review of clustering algorithms and energy corrections in the Electromagnetic Calorimeter. <http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/barfuss/NoteClustering/clustering.pdf?revision=1.1>.
- [24] H. Bethe and W. Heitler. On the Stopping of fast particles and on the creation of positive electrons. *Proc.Roy.Soc.Lond.*, A146 :83–112, 1934.
- [25] Clémentine Broutin. *Electron Measurements and Search for Higgs Bosons in Multi-Lepton Channels with the CMS Experiment at LHC*. PhD thesis, Ecole Polytechnique, 2011.
- [26] N. Adam *et al.* Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7\text{TeV}$. *CMS AN*, 2010/116, 2010.
- [27] <http://root.cern.ch/drupal/content/roofit>.
- [28] Stefan Dittmaier. A General approach to photon radiation off fermions. *Nucl.Phys.*, B565 :69–122, 2000.
- [29] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605 :026, 2006.
- [30] Johan Alwall, Pavel Demin, Simon de Visscher, Rikkert Frederix, Michel Herquet, et al. MadGraph/MadEvent v4 : The New Web Generation. *JHEP*, 0709 :028, 2007.
- [31] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations : the POWHEG method. *JHEP*, 0711 :070, 2007.
- [32] T. Binoth, N. Kauer, and P. Mertsch. Gluon-induced QCD corrections to $pp \rightarrow ZZ \rightarrow l\bar{l}l'\bar{l}'$. page 142, 2008.
- [33] CMS Collaboration. Observation of z to 4l production in pp collisions at 7 tev. *CMS PAS*, 2012/009, 2012.
- [34] CMS Collaboration. Search for the higgs boson in the decay channel $h \rightarrow zz^{(*)} \rightarrow 4l$. *CMS AN*, 2011/387, 2011.
- [35] <http://sherpa.hepforge.org/doc/SHERPA-MC-1.2.2.html>.
- [36] A.L. Read. Presentation of search result : the cls technique. *J. Phys. G, Nucl. Part. Phys.* 28 2693, 2002.
- [37] Georges Aad et al. Measurement of the ZZ production cross section and limits on anomalous neutral triple gauge couplings in proton-proton collisions at $\sqrt{s} = 7\text{TeV}$ with the ATLAS detector. *Phys.Rev.Lett.*, 108 :041804, 2012.

RÉSUMÉ

La thèse en physique des particules que j'ai effectuée au sein de l'équipe CMS (Compact Muon Solenoid) du Laboratoire Leprince-Ringuet de l'École Polytechnique porte sur l'analyse de la production de di-bosons $pp \rightarrow ZZ+X$ et la recherche de couplages anormaux et ce, dans le cadre de l'expérience CMS et des données accumulées à l'aide du grand collisionneur de hadrons du CERN, le LHC.

Elle débute avec le démarrage du LHC en fin 2009 et l'étude et l'optimisation de la reconstruction des électrons via la combinaison de deux algorithmes afin d'améliorer les performances à bas moment transverse. Les critères de sélection des événements caractéristiques des désintégrations de bosons Z, assurant une sélection pure d'électrons, ont été étudiés. Les efficacités de reconstruction des électrons pour les données 2010 et 2011 ainsi que leurs erreurs statistiques et systématiques ont été mesurées et ces mesures ont ensuite été utilisées par la collaboration CMS pour les analyses 2010 et 2011.

Une étude des caractéristiques des radiations dans l'état final (Final State Radiation) affectant les leptons issus de la désintégration du Z a été menée. Un algorithme en vue de récupérer les photons FSR et reconstruire pleinement la masse du Z a été développé et optimisé. Il corrige 4.2% des cas $ZZ \rightarrow 4\mu$ tout en assurant une pureté de sélection des photons FSR de 95.6%.

L'analyse physique développée dans cette thèse porte sur la recherche et l'extraction de limites sur les couplages trilineaires neutres. L'existence de ces couplages est interdite dans le formalisme du Modèle Standard mais peut cependant être introduite dans un Lagrangien effectif par l'ajout de deux termes de couplages f_4^Z et f_5^Z dans le cas du couplage ZZZ, en vue de la recherche d'une nouvelle physique. La présence de ces couplages anormaux se traduit généralement par une augmentation de la section efficace de production des di-bosons ZZ ainsi que par les déformations des distributions cinématiques des Z ou encore de la masse des 4 leptons par rapport à l'attendu Modèle Standard. Dans le cadre de l'analyse du canal $pp \rightarrow ZZ \rightarrow 4l$, une sélection pure des événements a été établie et les distributions différentielles du moment transverse et de la rapidité des bosons Z sont également présentées. Une recherche des couplages anormaux est effectuée en utilisant une méthode de repondération des distributions simulées et des limites sont extraites par la technique CL_s . Les limites obtenues avec un niveau de confiance de 95% sont : $-0.0116 < f_4^Z < 0.0116$ et $-0.0119 < f_5^Z < 0.0119$.

ABSTRACT

I did my PhD in Particle Physics with the CMS (Compact Muon Solenoid) team from Laboratoire Leprince-Ringuet at Ecole Polytechnique. The subjects of my study are the $pp \rightarrow ZZ+X$ production and the search for anomalous triple gauge couplings (ATGCs) using the data collected by the CMS detector at LHC at $\sqrt{s}=7$ TeV.

This work began with the first collisions at the LHC in the end of 2009 with the study and optimisation of electron reconstruction algorithms. Selection criteria of pure electrons samples from Z decays have been studied. The electron reconstruction efficiency has been measured on 2010 and 2011 data and the results have been used by the CMS collaboration for the 2010 and 2011 analyses.

Then I dealt with the analysis of $ZZ \rightarrow 4l$ channel and the search for anomalous triple gauge couplings, as possible signature of new physics. A recovery algorithm to correct the 4-leptons mass from Final State Radiations loss has been developed : 4.2% of $ZZ \rightarrow 4\mu$ events are corrected with an FSR selection purity of 95.6%. Then search and limits extraction of ATGCs have been done. Neutral trilinear couplings are forbidden in Standard Model but can be introduced using an effective Lagrangian with new couplings terms f_4^Z and f_5^Z in the ZZZ case. ATGCs generally generate an increase of the ZZ cross section production and some distortions in the ZZ mass spectrum or Z kinematical distributions. Thus a pure selection of $ZZ \rightarrow 4l$ events has been established and kinematical distributions have been studied. I have performed a search for ATGCs thanks to a reponderation method. Using the shape of the invariant mass of the 4-leptons system, limits on ZZZ ATGCs are set at 95% confidence level : $-0.0116 < f_4^Z < 0.0116$ and $-0.0119 < f_5^Z < 0.0119$.