



P<sup>q</sup>f ætftg'55/4236"  
N<sub>l</sub> EGP "ó"V"4236/36"

"  
"  
"

Vj<sup>3</sup>ug"

r t<sup>2</sup> ugp<sup>v</sup> g'f gxcpv'

nøWpkxgtuk<sup>2</sup> "Erwfg'Dgtpctf 'N{ qp"3"

! eqng'F qevqtcng'f g'Rj { uks wg'gvf øCuxtqr j { uks wg"

r qwt "nødvgpvkp'f w"

F RN<sup>o</sup> O G'f g'F QE VQTCV"  
Ur<sup>2</sup> ekck<sup>2</sup> "<Rj { uks wg'f gu'r ct vkwgu"

\*ctt<sup>´</sup> v<sup>2</sup> 'f w'9"cqÅv'4228+"  
"  
"  
r ct"  
"  
"

Lwkp'Ej cuugtcv'  
"  
"

### Recherche de leptoquarks de troisième génération dans l'expérience CMS au LHC

Uqwgpwg'rg'7"o ctu"4236"  
f gxcpv'rc'Ego o kuukqp'f øGzco gp"

"  
"

Lxt { "< "	"		
O O"	KO"	Ncmkpgj ""	Rt <sup>2</sup> ulf gpv'f w'lw { "
O O'	O O'	Lcht <sup>2</sup> ""	Tcr r qtvgw "
O O'	E O"	Tq { qp ""	Tcr r qtvgw " "
O o g"	U'	F cxf uqp"	Gzco kpcvgw "
O O'	U O"	O wcp  c"	Gzco kpcvgw "
O O'	R O'	Xgtf lgt "	F ktgevgw'f g'y <sup>3</sup> ug" "



"

"

UNIVERSITE CLAUDE BERNARD LYON 1  
ECOLE DOCTORALE PHAST  
PHYSIQUE des particules

# T H E S E

pour obtenir le titre de

**Docteur en Sciences**

de l'Université Claude Bernard - Lyon 1

**Mention : Physique**

Présentée et soutenue par

Julien CHASSERAT

## Recherche de leptoquarks de troisième génération dans l'expérience CMS au LHC

Thèse dirigée par Patrice VERDIER

préparée à l'Institut de Physique Nucléaire de Lyon

soutenue le 5 Mars 2014

**Jury :**

<i>Rapporteurs :</i>	Michel JAFFRÉ	-
	Christophe ROYON	-
<i>Directeur :</i>	Patrice VERDIER	-
<i>Président :</i>	Imad LAKTINEH	-
<i>Examineurs :</i>	Sacha DAVIDSON	-
	Steve MUANZA	-



A mes parents.

## Remerciements

Voici venu le temps de l'écriture si redoutée des remerciements. C'est une tâche ardue car chaque lecteur qui aura cet ouvrage entre les mains commencera inmanquablement par cette partie (et s'arrêtera bien souvent à celle-ci, je ne suis pas dupe). Tous y chercheront quelque chose de bien particulier. Cela pourra aller d'une simple évocation de leur nom, à la description d'un moment de complicité partagée, ou plus simplement la reconnaissance de l'aide qu'ils m'ont apporté pendant ces 3 merveilleuses années. Dans tous les cas, je dois m'appliquer à ne pas les décevoir. Si malgré tous mes efforts je venais à en oublier certains, je leur demande par avance de bien vouloir être indulgents.

Tout d'abord, merci à Patrice verdier, qui m'a encadré depuis le stage de M1 jusqu'à la fin de la thèse. Sa rencontre a été déterminante dans mon choix de me lancer dans une thèse en physique des particules. Il m'a patiemment initié aux joies des grep, awk et autres commandes occultes et il a sans relâche essayé de me détourner du côté obscur de la cliquouille. Je me dois bien sûr d'évoquer ses connaissances encyclopédiques en physique des particules. Je lui suis très reconnaissant de toutes celles qu'il a réussi à me transmettre pas à pas, toujours de manière passionnée et pédagogique. Il m'a de plus inculqué des méthodes de travail efficaces et rigoureuses qui, j'en suis certain, m'accompagneront tout au long de ma carrière.

Merci aux membres de mon jury de thèse qui ont accepté d'évaluer mon travail avec intelligence et bienveillance. Je veux aussi remercier tout spécialement mes deux rapporteurs messieurs Michel Jaffré et Christophe Royon pour avoir attentivement relu mon manuscrit. Ils ont su faire preuve d'une grande patience et de beaucoup de rigueur pour cette tâche fastidieuse.

Merci à messieurs Bernard Ille et Guy Chanfray successivement directeurs de l'Institut de Physique Nucléaire de Lyon pendant ma thèse. Ils m'ont fourni le meilleur cadre de travail dont j'aurais pu rêver pour obtenir ce doctorat.

De nombreuses personnes de l'Institut méritent aussi d'être citées dans ces remerciements. Tout d'abord Sylvie Florès qui est une véritable deuxième maman pour tous les doctorants de ce laboratoire. Elle sait comme personne déjouer les pièges administratifs de l'école doctorale mais elle fait surtout preuve d'une grande empathie et d'un soutien moral sans faille pour les étudiants en mal de confiance. Je tiens aussi à remercier Corinne Augier qui m'a beaucoup apporté sur la partie enseignement de ma thèse. Elle a su me conseiller et lever mes doutes quand j'ai dû faire face à une classe pour la première fois et elle a continué à me soutenir tout au long de ma thèse avec sa bonne humeur inébranlable. Pour la partie informatique je voudrais dire merci à Guillaume Beaulieu qui m'a énormément aidé quand la mise en place du serveur Madgraph devenait critique et qui a toujours su se mettre en quatre pour régler nos problèmes relationnels avec le tiers 3.

Je voudrais aussi en profiter pour remercier tous les gens de l'expérience CMS qui ont su me former et m'encadrer tout au long de ces 3 ans. Notamment Sylvano Tosi qui a été d'une grande aide pour mon travail avec le logiciel Madgraph, Stéphane Perries pour son avis toujours pertinent que ce soit pour des questions de physique ou

sur le fonctionnement chaotique de lyoserv et enfin Viola Sordini pour sa disponibilité et sa bonne humeur permanente.

J'aimerais aussi dire un énorme merci aux autres thésards qui ont partagé mon bureau pendant ces 3 années. Tout d'abord Yohan, qui était déjà en deuxième année de thèse quand je suis arrivé. Nous avons tout de suite sympathisé, et il m'a beaucoup apporté pendant ces premiers mois, toujours délicats, de ma nouvelle vie de doctorant. Que ce soit pour ses tutoriels très vivant autour du code de CMS, ses relectures de mails pour corriger mon anglais parfois approximatif ou pour toutes ces petites choses qui m'ont fait me sentir immédiatement à l'aise dans ce bureau 116. Il m'a de plus accueilli dans son cercle de proches comme si j'en avais toujours fait partie et plus qu'un simple collègue je pense avoir trouvé un véritable ami. Après le départ de Yohan, j'ai eu la chance d'accueillir dans ce même bureau Sébastien, un tout jeune doctorant qui paraissait bien timide au début. J'ai bien vite compris que je ne pourrai rien lui apprendre en matière de code en particulier et d'informatique en général. Au contraire, c'est lui qui m'a transmis une petite partie de son grand savoir au cours de nos longues discussions à bâtons rompus sur cet univers que je connaissais alors mal. Il m'a de plus donné les clés me permettant d'approfondir ces notions par moi-même tout en me donnant l'envie de découvrir les forums de discussions linux, les fonctions magiques de Latex et les concepts de programmation les plus alambiqués. Merci en particulier pour l'aide inestimable qu'il m'a apporté après une tentative hasardeuse de re-partition de mon disque où j'ai failli perdre l'intégralité des données de mes deux premières années de thèse... (GParted forever). J'ai rarement rencontré quelqu'un qui était autant fait pour ce métier d'enseignant chercheur et j'espère que notre amitié perdurera quand il occupera ce poste de permanent qu'il mérite. Enfin, pendant les derniers mois de ma thèse, nous avons accueilli deux nouvelles doctorantes pétillantes et pleines d'espoir, Elvire et Anne-Laure. Elles ont su apporter beaucoup de fraîcheur et de joie de vivre dans nos routines de geeks. J'espère qu'elles sauront conserver cet état d'esprit tout au long de leur parcours dans le monde de la recherche.

Un grand merci aux différentes générations de doctorants que j'ai pu côtoyer au cours de ma thèse. Les repas du midi à la maison d'hôtes avec cette immense table de joyeux thésards étaient toujours un moment phare de la journée et un régal pour les papilles (sauf les mousses!). Merci à tous les doctorants de CMS avec qui on échangeait souvent des astuces pour se faciliter l'utilisation si ardue de CMSSW. Une mention spéciale pour Maxime qui faisait presque partie de notre bureau tant il traînait autour du 116 pour soutirer des infos à Seb ou plus prosaïquement une tasse de thé. Que ce soit au travail ou en dehors il a toujours été un ami exemplaire et généreux, j'espère que lui aussi saura trouver sa place dans le monde de la recherche. Je remercie aussi les doctorants des autres laboratoires avec qui nous avons su construire un véritable groupe d'amis. Thomas, Dimitri, Lucas et Arnaud, ça dure depuis le M2 et ce n'est pas près de s'arrêter. David, Romain, Guillaume et Binbin à très vite pour une nouvelle session de débats politiques, sportifs ou culinaires autour d'un petit verre.

Merci aussi à tous mes amis éparpillés à travers la France. Je sais que j'ai été

moins présent pendant ces années de thèse mais je compte bien me rattraper dans les années qui viennent. Merci à tous mes potes de Sanguinet qui me chambraient gentiment en me donnant du docteur. Merci à mes amis de prépa notamment Edouard, Mathieu et Christophe qui m'ont regardé m'épanouir dans cette nouvelle discipline. Merci à tous les amis que j'ai conservé de l'époque ENTPE, l'ambiance incroyable qui règne dans ce groupe m'a beaucoup aidé à me ressourcer durant ces weekends si conviviaux organisés aux quatre coins de la France. Une pensée également pour Charles, je lui souhaite de trouver sa voie après les épreuves difficiles que nous avons partagées et qui nous ont beaucoup rapprochés. Merci aussi à tous mes amis de Lyon et notamment la bande des bisontins complètement barrés. Roman, Hélène, Poulet, Méli, Tinouze, Clément et Phoebe (et oui tu es dedans maintenant) vous avez toujours été présents quand j'avais besoin de décompresser autour d'un bon match de foot ou pour une soirée endiablée.

C'est aussi un grand plaisir pour moi de remercier ma famille. Tout d'abord ma mère qui m'a toujours soutenu dans mes choix d'orientation parfois hasardeux. Maman je t'aime, et ce doctorat est en grande partie le tien. Je tiens aussi à exprimer ici une pensée pour mon père qui, je pense, aurait été fier de me voir devenir docteur dans ce monde de la recherche qu'il a tant admiré. Merci aussi à mes frères qui ont dû suivre cette aventure de loin et pour lesquels j'ai été forcément moins disponible. Vous devez être ravis Nu et Lu, grâce à cette thèse je suis officiellement admis au club des geeks.

Je voudrais enfin remercier Stéphanie, même si on ne s'est rencontré que vers la fin de ma thèse, je suis persuadé que je n'aurais pas pu réussir aussi bien sans toi. Tu as par la suite été d'un grand soutien quand il a fallu tourner la page de la recherche académique. Merci pour ta présence, ton affection et pour tes rires...

# Table des matières

<b>Introduction</b>	<b>xi</b>
<b>1 Le modèle standard de la physique des particules</b>	<b>1</b>
1.1 Principes théoriques . . . . .	1
1.1.1 Conventions . . . . .	1
1.1.2 L'invariance de jauge . . . . .	2
1.1.3 Théorie des groupes . . . . .	3
1.1.4 De la mécanique lagrangienne à la Théorie Quantique des champs . . . . .	3
1.1.5 Le Théorème de Noether . . . . .	5
1.1.6 Les symétries . . . . .	5
1.2 Les interactions fondamentales du modèle standard . . . . .	7
1.2.1 L'interaction forte . . . . .	7
1.2.2 L'interaction électrofaible . . . . .	8
1.2.3 Le lagrangien du modèle standard . . . . .	14
1.3 Les problèmes du Modèle Standard . . . . .	14
1.4 Introduction théorique aux leptoquarks . . . . .	16
1.4.1 Un modèle effectif pour la recherche de LQ . . . . .	16
1.4.2 Lagrangien effectif et couplages . . . . .	17
1.4.3 Désintégration des LQ . . . . .	20
1.4.4 État de l'art . . . . .	21
<b>2 Dispositif expérimental</b>	<b>23</b>
2.1 Le Grand Collisionneur de Hadrons (LHC) . . . . .	23
2.1.1 Phénoménologie des collisions protons protons . . . . .	25
2.1.2 Le complexe d'accélération . . . . .	25
2.1.3 Luminosité et nombre d'événements . . . . .	27
2.1.4 Le défi de l'empilement (PU) . . . . .	27
2.1.5 Quelle physique au LHC? . . . . .	29
2.2 L'expérience Compact Muon Solenoid (CMS) . . . . .	31
2.2.1 Le solénoïde . . . . .	33
2.2.2 Le trajectographe . . . . .	34
2.2.3 Le calorimètre électromagnétique . . . . .	36
2.2.4 Le calorimètre hadronique . . . . .	38
2.2.5 Les chambres à muons . . . . .	40
2.3 Le système de déclenchement de CMS . . . . .	44
2.3.1 Le déclenchement de premier niveau L1 . . . . .	46
2.3.2 Le déclenchement de haut niveau HLT . . . . .	47
2.4 Reconstruction d'événements dans CMS . . . . .	48
2.4.1 L'algorithme de Particle Flow . . . . .	48

2.4.2	Identification et reconstruction des leptons . . . . .	50
2.4.3	Les jets . . . . .	52
2.4.4	Énergie transverse manquante . . . . .	54
2.5	Traitement de données et grille de calcul . . . . .	55
2.5.1	CMSSW et le format de données EDM . . . . .	55
2.5.2	La grille de calcul . . . . .	56
<b>3</b>	<b>Simulation Monte-Carlo avec MADGRAPH pour la collaboration CMS</b>	<b>59</b>
3.1	La génération Monte-Carlo en physique des particules . . . . .	60
3.1.1	Intégration numérique par tirage Monte-Carlo . . . . .	60
3.1.2	La méthode Monte-Carlo en physique des particules . . . . .	62
3.2	La chaîne de génération d'événements Monte-carlo . . . . .	63
3.2.1	Les fonctions de densité partonique (PDF) . . . . .	64
3.2.2	L'événement dur . . . . .	66
3.2.3	La gerbe partonique (Parton Shower ou PS) . . . . .	67
3.2.4	Appariement des jets . . . . .	68
3.2.5	Événements sous jacent et interaction partonique multiple . . . . .	69
3.2.6	Hadronisation et désintégration de hadrons . . . . .	70
3.3	Simulation du détecteur . . . . .	71
3.3.1	Simulation rapide . . . . .	71
3.3.2	Simulation complète . . . . .	71
3.4	Présentation du format LHE et des générateurs utilisés . . . . .	72
3.4.1	Le format LHE . . . . .	72
3.4.2	Le générateur MADGRAPH . . . . .	72
3.4.3	Le générateur PYTHIA . . . . .	73
3.4.4	Interface MADGRAPH PYTHIA . . . . .	74
3.5	Génération et validation d'événements Monte-Carlo à grande échelle pour la collaboration CMS . . . . .	76
3.5.1	Production d'événements MC et post processing . . . . .	76
3.5.2	Un exemple de validation d'événements physiques . . . . .	78
3.5.3	Production locale de gridpacks pour les échantillons $t\bar{t}$ systématiques . . . . .	79
3.5.4	Maîtriser l'appariement : détermination de la valeur optimale de $Q_{MS}$ . . . . .	80
3.5.5	Conclusion . . . . .	81
<b>4</b>	<b>Recherche de leptoquarks de 3 ème génération dans l'expérience CMS</b>	<b>83</b>
4.1	Phénoménologie des leptoquarks au LHC . . . . .	84
4.1.1	Modes de production des Leptoquarks au LHC . . . . .	84
4.1.2	Sections efficaces de production des leptoquarks scalaires . . . . .	85
4.1.3	Simulation du signal . . . . .	88
4.2	Simulation MC des événements de fonds du MS . . . . .	88

---

4.3	Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 . . .	90
4.3.1	Lots de données utilisés et facteurs correctifs . . . . .	90
4.3.2	Contrôle des fonds du modèle standard . . . . .	94
4.4	Recherche de Leptoquarks de 3ème génération se désintégrant en top $\tau$	100
4.4.1	État final et canal d'analyse . . . . .	100
4.4.2	Pré-sélection des événements . . . . .	102
4.4.3	Les incertitudes statistiques . . . . .	104
4.4.4	Les incertitudes systématiques . . . . .	105
4.4.5	Optimisation finale . . . . .	107
4.4.6	Masse limite et interprétation des résultats . . . . .	110
4.5	Recherche de Leptoquarks de 3ème génération se désintégrant en top $\nu$	110
4.5.1	État final et canal d'analyse . . . . .	110
4.5.2	Pré-sélection des événements . . . . .	114
4.5.3	Optimisation finale et incertitudes . . . . .	115
4.5.4	Masse limite et interprétation des résultats . . . . .	115
4.6	Recherche de Leptoquarks de 3ème génération se désintégrant en top $\mu$	118
4.6.1	État final et canal d'analyse . . . . .	118
4.6.2	Pré-sélection des événements . . . . .	118
4.6.3	Optimisation finale et incertitudes . . . . .	121
4.6.4	Masse limite et interprétation des résultats . . . . .	125
4.7	Conclusions et perspectives . . . . .	127



# Acronymes

Liste des acronymes les plus fréquents :

<b>ALICE</b>	A Large Ion Collider Experiment
<b>AOD</b>	Analysis Object Data
<b>ATLAS</b>	A Toroidal LHC ApparatuS experiment
<b>CERN</b>	Organisation Européenne pour la Recherche Nucléaire
<b>CMS</b>	Compact Muons Solenoid
<b>CMSSW</b>	CMS Software
<b>ECAL</b>	Electromagnetic Calorimeter
<b>EDM</b>	Event Data Format
<b>FSR</b>	Final State Radiation
<b>GEANT 4</b>	GEometry ANd Tracking : Simulateur du passage de particules à travers la matière par méthode Monte-Carlo
<b>HB</b>	Hadronic Barrel Calorimeter
<b>HCAL</b>	Hadronic CALorimeter
<b>HF</b>	Hadronic Forward Calorimeter
<b>HLT</b>	Higher Level Trigger
<b>HO</b>	Hadronic Outer Calorimeter
<b>ISR</b>	Initial State radiation
<b>L1</b>	Level 1 trigger
<b>LHE</b>	Les Houches events : MC sample standardized format
<b>MC</b>	Monte-Carlo
<b>ME</b>	Matrix Element : type of MC generator
<b>PU</b>	Pile-up
<b>QCD</b>	Quantum Chromo-Dynamics
<b>QED</b>	Quantum Electro-Dynamics
<b>SM</b>	Standard Model
<b>TDR</b>	Technical Design Report
<b>TIB</b>	Tracker Inner Barrel
<b>TID</b>	Tracker Inner Disks
<b>TOB</b>	Tracker Outer Barrel



# Introduction

Notre connaissance des particules élémentaires et de leurs interactions a beaucoup évolué au cours du siècle dernier. Les avancées successives ont été incorporées progressivement dans une théorie élégante et synthétique qu'on appelle Modèle Standard (MS) de la physique des particules. C'est une théorie de jauge basée sur le groupe  $SU(3)_c \times SU(2)_L \times U(1)_Y$  qui décrit trois des quatre interactions fondamentales dans le cadre de la Théorie Quantique des Champs (TQC). Les interactions électromagnétique, faible et forte sont ainsi réunies dans un même formalisme. La gravitation est décrite pour sa part dans le contexte de la relativité générale. L'interaction gravitationnelle est cependant négligeable en dessous de l'échelle de Planck ( $10^{18}$  GeV) ce qui est toujours le cas dans le cadre de la physique des particules. La physique des hautes énergies a depuis permis de renforcer ce modèle et de mesurer ses paramètres avec une grande précision.

Le démarrage du LHC en 2009 a ouvert un nouveau chapitre de l'histoire de la physique des particules en permettant l'accès à des énergies jamais atteintes auparavant. Les trois années de fonctionnement intensifs qui ont suivies ont permis d'accumuler des données avec une montée progressive en gamme d'énergie et en luminosité. Aujourd'hui arrêté depuis un an, le LHC et ses expériences sont minutieusement préparés pour relever le défi du redémarrage à 13 TeV prévu en 2015. En attendant, le traitement de ces données a non seulement permis d'améliorer la précision sur la mesure des paramètres du MS, mais surtout d'atteindre un des objectifs majeurs du LHC, la découverte du boson de Higgs<sup>1</sup>. Elles sont aussi exploitées intensivement pour tester des modèles au delà du MS comme la Super Symétrie (SUSY) ou le modèle des leptoquarks. Jusqu'ici, aucun de ces modèles n'a pu être mis en évidence et les expérimentateurs continuent de contraindre ces modèles toujours plus strictement. Espérons que le fonctionnement du LHC à haute énergie nous donnera les pistes nécessaires pour orienter les extensions possibles du MS vers une direction plus précise.

Ce manuscrit, qui résume le travail de mes trois années de thèse effectuées au sein de la collaboration CMS, comporte 4 chapitres. Le premier chapitre introduit le cadre théorique du Modèle Standard et présente le modèle effectif des Leptoquarks que j'ai étudié. Le deuxième chapitre présente le dispositif expérimental composé de l'expérience CMS au sein du collisionneur LHC. Il contient par ailleurs la description du système de déclenchement, de la reconstruction des événements et du schéma de traitement de données. Le troisième chapitre est consacré à mon travail de service effectué durant ma première année de thèse dans le groupe générateur de CMS. Il contient une description globale du fonctionnement des générateurs d'événements Monte-Carlo en physique des particules ainsi que certains exemples représentatifs des tâches qui m'ont été attribuées pendant cette première année. Pour finir, le quatrième et dernier chapitre présente mon travail d'analyse à proprement parler.

---

1. Officiellement le 4 Juillet 2012

Il est divisé en plusieurs sous-parties dont une présentation de la phénoménologie des leptoquarks au LHC, le détail des lots de données et des échantillons MC utilisés et enfin le résumé des trois analyses menées à bien pendant ma thèse. Elles correspondent à trois modes de désintégration possibles des Leptoquarks en top  $\tau$ , top  $\nu$  et top  $\mu$ . Bien que résultant d'une méthodologie commune, elles ont été optimisées séparément et en adéquation avec les spécificités de chaque état final.

# Le modèle standard de la physique des particules

---

## Sommaire

---

<b>1.1 Principes théoriques . . . . .</b>	<b>1</b>
1.1.1 Conventions . . . . .	1
1.1.2 L'invariance de jauge . . . . .	2
1.1.3 Théorie des groupes . . . . .	3
1.1.4 De la mécanique lagrangienne à la Théorie Quantique des champs	3
1.1.5 Le Théorème de Noether . . . . .	5
1.1.6 Les symétries . . . . .	5
<b>1.2 Les interactions fondamentales du modèle standard . . . . .</b>	<b>7</b>
1.2.1 L'interaction forte . . . . .	7
1.2.2 L'interaction électrofaible . . . . .	8
1.2.3 Le lagrangien du modèle standard . . . . .	14
<b>1.3 Les problèmes du Modèle Standard . . . . .</b>	<b>14</b>
<b>1.4 Introduction théorique aux leptosquarks . . . . .</b>	<b>16</b>
1.4.1 Un modèle effectif pour la recherche de LQ . . . . .	16
1.4.2 Lagrangien effectif et couplages . . . . .	17
1.4.3 Désintégration des LQ . . . . .	20
1.4.4 État de l'art . . . . .	21

---

## 1.1 Principes théoriques

### 1.1.1 Conventions

#### Unités naturelles

Toutes les formules sont exprimées en unités naturelles, c'est à dire :

$$\hbar = c = 1 \tag{1.1}$$

Dans ce système d'unités, la longueur, le temps, l'énergie et la masse sont mesurés dans les unités suivantes :

$$[L] = \text{GeV}^{-1} \quad (1.2)$$

$$[T] = \text{GeV}^{-1} \quad (1.3)$$

$$[E] = \text{GeV} \quad (1.4)$$

$$[M] = \text{GeV} \quad (1.5)$$

### Métrie

Tous les calculs sont effectués dans l'espace de Minkowski et le tenseur métrique est choisi comme suivant :

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.6)$$

Nous adopterons par ailleurs la convention de sommation d'Einstein : les indices grecs répétés implique une somme sur les composantes d'un tenseur.

#### 1.1.2 L'invariance de jauge

Pour comprendre ce principe d'invariance qui jouera un rôle prépondérant par la suite, examinons comment il s'applique dans le cadre de la théorie de l'électromagnétisme de Maxwell [1]. Les champs électrique  $\vec{E}$  et magnétique  $\vec{B}$  peuvent être exprimés en termes de potentiels selon les équations (1.7) et (1.8).

$$\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t} \quad (1.7)$$

$$\vec{B} = \vec{\nabla} \wedge \vec{A} \quad (1.8)$$

où  $\vec{A}$  est le potentiel vecteur du champ magnétique et  $V$  le potentiel scalaire électrique.

Ces équations nous apprennent notamment que pour un couple de champs  $(\vec{E}, \vec{B})$ , il y a une infinité de couples de potentiels  $(V, \vec{A})$  correspondants. En effet, si l'on applique les transformations (1.9) aux potentiels, les vecteurs de champs restent inchangés.

$$V \rightarrow V - \frac{\partial f}{\partial t} \quad \vec{A} \rightarrow \vec{A} + \vec{\nabla}f \quad (1.9)$$

En choisissant  $f$  une fonction quelconque de l'espace et du temps. Ces transformations sont appelées transformations de jauge et l'on dit que  $\vec{E}$  et  $\vec{B}$  sont des invariants de jauge.

On peut alors réécrire les équations de Maxwell à l'aide du potentiel quadriverseurs  $A^\mu = (V, \vec{A})$  pour se rapprocher des notations de la physique des particules. La transformation de jauge prend ainsi la forme suivante :

$$A^\mu \rightarrow A^\mu - \partial^\mu f \quad (1.10)$$

avec  $\partial^\mu = (\frac{\partial}{\partial t}, -\vec{\nabla})$  l'opérateur différentiel pour les quadrivecteurs.

Les équations de Maxwell peuvent être réécrites de façon simplifiée à l'aide de cette même notation comme suit :

$$\partial_\mu F^{\mu\nu} = j_{em}^\nu \quad (1.11)$$

avec  $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$  le tenseur électromagnétique, et  $j_{em}^\nu$  le quadricourant électromagnétique. On peut montrer que le tenseur  $F^{\mu\nu}$  est invariant par la transformation de jauge (1.10) et donc que les équations de Maxwell sont des invariants de jauge.

### 1.1.3 Théorie des groupes

En mathématiques, le principe de symétrie est associé à celui de groupe de transformation. Un groupe est un ensemble muni d'une loi de composition interne associative qui comprend un élément neutre et dont chaque élément possède un inverse à l'intérieur du groupe. Pour certains groupes dits abéliens, cette loi de composition est, en sus, commutative. On peut par ailleurs associer à chaque transformation du groupe une matrice  $M$ . L'ensemble des matrices  $M$  correspondantes est appelée une représentation du groupe. Les symétries ayant un rôle central en physique des particules, les physiciens ont naturellement adopté la formulation mathématique de la théorie des groupes. Voilà en résumé les caractéristiques et les notations des groupes les plus souvent utilisés dans notre discipline :

- S : qui signifie « spécial », indique que le déterminant des matrices représentant la transformation vaut 1. Il en découle que les vecteurs colonnes de la matrice sont une base de l'espace sous-jacent ;
- U : pour « unitaire » signifie que l'inverse de la matrice représentant la transformation est aussi sa matrice adjointe ( $A^*A = AA^* = I$ ) ;
- O : signifiant « orthogonale », l'inverse de la matrice représentant la transformation est sa matrice transposée ( $A^tA = AA^t = I$ ) ;
- (N) : exprime la dimension du groupe.

Ainsi, le groupe d'invariance de couleur SU(3) qui régit la ChromoDynamique Quantique de l'interaction forte (QCD), est constitué de l'ensemble des matrices unitaires spéciales de dimension 3. Les groupes de Lie, très utilisés en Théorie Quantique des Champs (TQC), sont des groupes continus dont les opérations de groupes (multiplication et inversion) sont, de plus, différentiables. Pour plus de précisions sur la théorie des groupes, voir [2] et pour une présentation complète des algèbres de Lie, voir [3].

### 1.1.4 De la mécanique lagrangienne à la Théorie Quantique des champs

#### 1.1.4.1 Le mécanique lagrangienne classique et le principe de Hamilton

En mécanique classique, l'un des principes les plus élégant est certainement le principe de moindre action ou principe de Hamilton [4]. Il peut être formulé comme

suit :

« Parmi toutes les trajectoires physiques possibles pour un système dynamique, la trajectoire effectivement empruntée sera celle pour laquelle la fonctionnelle d'action  $S[x(t)]$  est un extremum. »

L'action est elle même définie comme l'intégrale de la différence entre l'énergie cinétique et potentielle (1.12). C'est une fonctionnelle dans le sens où elle associe une fonction (la trajectoire  $x(t)$ ) à un réel (la valeur de l'intégrale).

$$S[x(t)] = \int_{t_1}^{t_2} (T - U) = \int_{t_1}^{t_2} L(t) dt \quad (1.12)$$

La fonctionnelle  $L(\dot{x}, x, t) = T - U$  est appelé le lagrangien du système. En termes mathématiques, le principe de moindre action se traduit donc ainsi :

$$\delta S[x] = \delta \int_{t_1}^{t_2} L(\dot{x}, x, t) = 0 \quad (1.13)$$

En développant le calcul des variations de cette intégrale, on trouve que les extremum de l'action satisfont les équations d'Euler-Lagrange :

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0 \quad (1.14)$$

La puissance de ce formalisme vient du fait que l'on peut trouver l'équation du mouvement très rapidement. Si on prend l'exemple classique de l'oscillateur harmonique à une dimension, le lagrangien prend la forme suivante :

$$L = T - U = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2 \quad (1.15)$$

Et l'équation d'Euler-Lagrange donne directement l'équation du mouvement :

$$kx + m\ddot{x} = 0 \quad (1.16)$$

Il est à noter que toute l'information sur le système est contenu dans le lagrangien.

### 1.1.4.2 Le lagrangien en théorie quantique des champs

Les particules sont maintenant décrites par des champs relativistes (voir 1.2) qui varient selon les trajectoires empruntées et les transitions entre états. Comme c'est une théorie relativiste, les variables de temps et d'espace doivent être traitées sur le même plan, il faut donc réécrire l'action comme suit :

$$S[\Phi] = \int_{t_1}^{t_2} \int_V \mathcal{L} dx dt = \int \mathcal{L} d^4x \quad (1.17)$$

La fonctionnelle  $\mathcal{L}(\partial_\mu \phi(x), \phi(x))$  est, à proprement parler, la densité lagrangienne mais en physique des particules on l'assimile simplement au lagrangien de la théorie. Comme précédemment, on écrit les équations d'Euler Lagrange pour obtenir directement les équations du mouvement :

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} = 0 \quad (1.18)$$

### 1.1.5 Le Théorème de Noether

Ce théorème fut démontré en 1915 par Emmy Noether, et fut qualifié en son temps par Albert Einstein de « monument de la pensée mathématique ». Il s'énonce de la manière suivante :

« A toute transformation infinitésimale qui laisse invariante l'intégrale d'action correspond une grandeur qui se conserve. »

Ainsi, si l'on trouve une symétrie qui laisse invariant notre théorie, elle correspondra automatiquement à une quantité conservée au cours des interactions prévues par ce modèle. Il faut par ailleurs remarquer que la réciproque est également vraie. Dès qu'une quantité physique se conserve, on peut y associer une symétrie des lois naturelles. Ce théorème, très utilisé en mécanique classique et en mécanique des fluides, est d'une importance fondamentale pour la physique des particules. Il va en effet permettre de déduire les conséquences physiques du modèle à partir des symétries de la théorie. On peut par exemple montrer que :

- Homogénéité de l'espace  $\Leftrightarrow$  Conservation de l'impulsion ;
- Isotropie de l'espace  $\Leftrightarrow$  Conservation du moment cinétique ;
- Homogénéité du temps  $\Leftrightarrow$  Conservation de l'énergie totale du système.

### 1.1.6 Les symétries

Les symétries mathématiques jouent un rôle central dans la construction du modèle standard. Comme on le verra, différents groupes de symétries sont associés aux trois interactions fondamentales décrites par le MS. Par l'intermédiaire du théorème de Noether, ces symétries vont non seulement fournir les invariants de la théorie mais elles vont aussi faire apparaître naturellement les bosons qui portent ces interactions fondamentales. L'ensemble des symétries qui laissent une théorie invariante est appelé groupe de symétrie de cette théorie (voir 1.1.3). Les propriétés mathématiques de ces groupes vont directement induire la physique décrite par ces théories.

Il convient de distinguer deux types de symétries fondamentales. Les symétries externes sont des transformations de l'espace temps qui affectent les coordonnées alors que les symétries internes sont des transformations abstraites qui s'appliquent uniquement aux champs.

#### Les symétries externes

Le cadre de la physique des particules actuelle est celui la relativité restreinte (voir par exemple [5]). On se place alors dans l'espace de Minkowski (3 dimensions d'espace et une de temps) et le groupe de symétrie associé est le groupe de Poincaré.

Il contient les transformations suivantes :

- les translations dans le temps et l'espace ;
- les rotations dans l'espace ;
- le renversement du temps T et la parité P ;

— les transformations de Lorentz (qui sont en fait des composés des deux premiers).

On les appelle symétries externes de l'espace. Toute théorie que l'on veut développer dans ce cadre doit rester invariante par toute symétrie externe.

### Les symétries internes

Ces symétries, plus abstraites, agissent directement sur les champs des particules :

$$\psi \rightarrow \psi' = \beta(\psi) \quad (1.19)$$

Cette transformation est une symétrie interne si elle laisse le lagrangien inchangé, soit :

$$\mathcal{L}(\psi', \partial_\mu \psi') = \mathcal{L}(\psi, \partial_\mu \psi) \quad (1.20)$$

On distinguera deux types de symétries internes selon qu'elles dépendent des coordonnées d'espace temps ou non.

Si elles ne dépendent pas de ces coordonnées, on parle d'invariance de jauge globale. Toujours d'après le théorème de Noether, elles vont engendrer la conservation de certaines charges et de courants. Par exemple, dans le cadre de la théorie de l'électrodynamique quantique, la conservation de la charge électrique découle directement de cette invariance de jauge globale. Ces nouvelles symétries permettent de compléter la liste des invariants de la théorie trouvés par l'intermédiaire des symétries externes (énergie, quantité de mouvement et moment angulaire).

Si, en revanche, les paramètres de la transformation dépendent des coordonnées d'espace temps, on parle d'invariance de jauge locale. Quand on essaye de rendre le lagrangien d'une théorie invariant par ce genre de symétrie, on se heurte au problème des dérivées qui entraînent l'ajout de termes interdisant l'invariance. Par analogie avec l'invariance de jauge en électromagnétisme (voir 1.1.2) on règle le problème en introduisant de nouveaux champs (dits de jauge) et en utilisant la dérivée partielle covariante ( $D_\mu = \partial_\mu - igX_\mu$ ) au lieu de la dérivée partielle classique ( $\partial_\mu$ ). Nous verrons plusieurs exemples de ce mécanisme dans la partie 1.2 quand nous détaillerons les différents lagrangiens associés aux interactions. Cette fois ci, les propriétés mathématiques de ces transformations vont avoir des conséquences encore plus grandes sur le contenu physique de la théorie. En effet, elles font émerger naturellement les bosons et permettent de décrire leur couplage avec les autres particules. Ces symétries internes jouent donc un rôle prépondérant dans la construction du modèle standard (MS) car elles nous donnent les clés de la description des interactions entre particules.

**Les brisures de symétries** Pour finir cette partie sur les symétries, il convient de rappeler les mécanismes de brisure de symétrie. Ces mécanismes auront en effet des conséquences importantes pour le MS.

Il existe deux grands types de brisures de symétrie. D'une part, les brisures explicites qui font disparaître l'invariance du lagrangien après transformation (comme la brisure de la symétrie globale SU(2)). D'autre part, les brisures spontanées qui

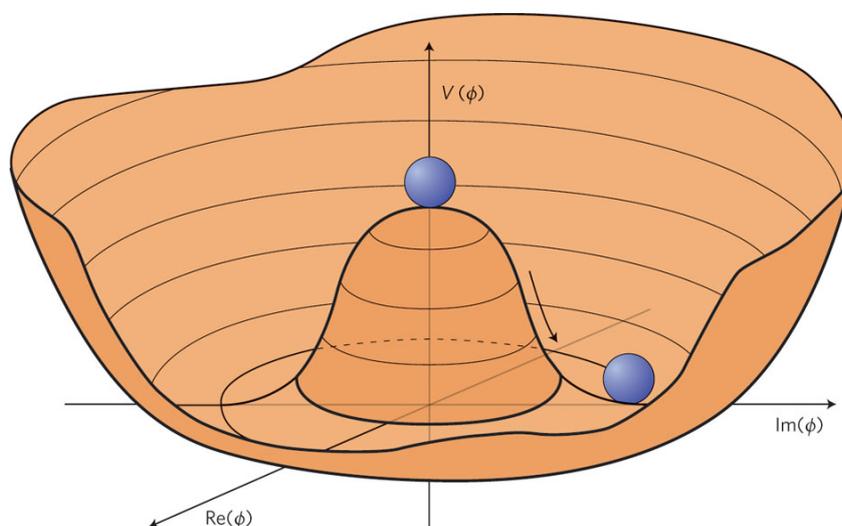


Fig. 1.1 – Le potentiel du champs de Higgs.

laissent le lagrangien invariant après transformation bien que l'état fondamental ne soit pas symétrique. On rencontrera une telle brisure dans le chapitre 1.2.2.4 qui décrit le mécanisme de Higgs. On donne souvent comme exemple celui du « chapeau mexicain » dont l'état instable, la pointe du chapeau, correspond à un état symétrique. Alors que l'état fondamental qui correspond au puits de potentiel perd sa symétrie autour de l'axe  $z$  (voir 1.1).

## 1.2 Les interactions fondamentales du modèle standard

Pour savoir quel groupe de symétrie est associé à chaque interaction, il faut s'intéresser au nombre de degrés de liberté de chaque théorie. En effet, le nombre de générateurs du groupe correspond au nombre de bosons de jauge de la théorie. L'électromagnétisme est porté par le seul photon, elle ne comporte donc qu'un seul degré de liberté ce qui justifie le choix du groupe de symétrie  $U(1)$ . Pour l'interaction faible, il y a 3 bosons de jauge ( $W^+$ ,  $W^-$  et  $Z$ ) ce qui nécessite un groupe de symétrie à 3 degrés de liberté comme  $SU(2)$ . Ces deux interactions ne sont cependant pas orthogonales, elles ne peuvent être décrites de façon précise que dans un modèle couplé. Voilà pourquoi on parle le plus souvent d'interaction électrofaible avec pour groupe de symétrie  $SU(2) \times U(1)$ . Enfin, l'interaction forte qui comporte 8 gluons de jauge, est décrite à l'intérieur de la symétrie  $SU(3)$ .

### 1.2.1 L'interaction forte

La Chromodynamique Quantique est une théorie de jauge qui formalise la conservation de la charge de couleur dans les réactions régies par l'interaction forte. Le groupe de symétrie représentatif irréductible de la théorie est  $SU(3)$ . Contrairement au groupe représentatif de la QED  $U(1)$ , il n'est pas commutatif

ce qui a une implication directe sur les interactions permises (l'auto couplage des gluons par exemple). Il admet pour générateurs les 8 matrices de Gell-Mann  $\lambda_{1-8}$  et l'algèbre de Lie correspondante est définie par les relations suivantes :

$$[\lambda_i, \lambda_j] = f_{ijk} \lambda_k \quad (1.21)$$

avec  $f_{ijk}$  les constantes de structure réelles du groupe. Le lecteur pourra trouver une large bibliographie traitant de la QCD qui dépasse le cadre de ces travaux, voir notamment [6] ou [7].

## 1.2.2 L'interaction électrofaible

Nous allons maintenant présenter les principales caractéristiques du modèle électrofaible. Pour de plus amples développements sur le sujet, on peut retenir parmi tant d'autres [8] et [9].

### 1.2.2.1 L'électrodynamique quantique

L'électromagnétisme ne s'applique qu'aux particules chargées électriquement. La quantité conservée à chaque interaction est la charge électrique  $Q$  (On note donc le groupe de symétrie sous-jacent  $U(1)_Q$ ). Le lagrangien de base de la QED qui correspond à l'équation de Dirac est le suivant :

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (1.22)$$

La symétrie de jauge locale agit sur les fermions comme dans l'équation 1.23 et doit laisser le lagrangien de la théorie invariant.

$$\psi(x) \rightarrow e^{i\beta(x)}\psi(x) \quad (1.23)$$

Pour conserver l'invariance du lagrangien sous cette transformation, on doit introduire le champ de jauge  $A_\mu(x)$  dont l'interaction avec les particules est caractérisée par le couplage  $e$ . Il se transforme comme suit :  $A_\mu(x) \rightarrow A_\mu(x) + \frac{1}{e}\partial_\mu\beta(x)$ . La dérivée partielle  $\partial_\mu$  est, elle, remplacée par la dérivée covariante  $D_\mu = \partial_\mu - ieA_\mu$ . Enfin, il faut rajouter au lagrangien un terme qui prend en compte la propagation libre du champ de jauge dans le vide. Ce terme peut être exprimé  $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$  avec  $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$  le tenseur électromagnétique qui s'identifie à celui utilisé en relativité. Ce tenseur est d'ailleurs invariant par symétrie de jauge locale ce qui garantit l'invariance du lagrangien de la QED dans son ensemble (1.24)

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (1.24)$$

Le champ de jauge  $A_\mu$  est ainsi la représentation du boson de jauge de l'électromagnétisme, le photon. Demander l'invariance du lagrangien sous la symétrie de jauge locale a fait apparaître naturellement une nouvelle particule de spin 1 qui est le vecteur de l'interaction électromagnétique. La propagation de ce boson dans

le vide est donné par le terme  $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ . L'interaction du photon avec les autres particules émerge de la dérivée covariante et elle est caractérisée par le couplage  $e$ . Il apparaîtra pour chaque vertex d'interaction dans un calcul d'élément de matrice pour un diagramme donné. Cette constante  $e$  n'est autre que la constante de structure fine  $e=1/137$ . Elle fut initialement introduite en physique atomique pour caractériser les sauts d'énergies des électrons du nuage électronique de l'atome par interaction avec un photon. A chaque fois qu'on rajoute une boucle dans un diagramme de Feynman de QED, on multiplie l'élément de matrice par  $e$ . On peut alors réaliser un développement en série entière par ordre de  $e$  de cet élément de matrice. La valeur relativement faible de ce couplage fait qu'on peut utiliser la théorie des perturbations pour traiter l'électrodynamique. Selon la précision voulue, on peut négliger les termes de plus haut degré en  $e$  et obtenir un résultat approché fiable avec une quantité raisonnable de calculs.

### 1.2.2.2 L'interaction faible

Le premier exemple d'interaction faible découvert fut celui de la désintégration  $\beta$ . Elle se produisait avec un temps anormalement long et fut donc associée à une force plus faible que celles que l'on connaissait déjà (selon le principe d'incertitude d'Heisenberg  $\Delta t \Delta E \geq \hbar$ ). Elle indiquait ainsi l'existence d'un boson chargé massif qui se couplait à l'électron ainsi qu'au neutrino. Ces interactions semblaient par ailleurs, impliquer uniquement des particules d'hélicité gauche. Il fallut attendre 1957, pour que C.S Wu démontre dans un article devenu célèbre [10] que l'interaction faible se comportait différemment avec les deux états de chiralité des champs fermioniques. Ce fut la fondation de la théorie de la violation de parité. Si  $\psi(x)$  est un champ spinoriel quelconque, ses états de chiralité gauche et droit sont donnés par :

$$\psi_L(x) = \frac{1 - \gamma^5}{2}\psi(x) \text{ et } \psi_R(x) = \frac{1 + \gamma^5}{2}\psi(x) \quad (1.25)$$

Cette expérience montre par ailleurs que la violation de parité est maximale, c'est à dire que les courants chargés de l'interaction faible se couplent uniquement aux états gauches des fermions (ou aux états droits des anti fermions). Chaque champ fermionique peut ainsi être écrit selon la décomposition précédente  $\psi = \psi_R + \psi_L$ . Les composantes droites ne sont pas sensibles à l'interaction faible et sont représentées par un singulet. Les composantes gauches sont regroupées en doublets. A chaque génération de fermions correspond donc un doublet et deux singulets. Pour la première génération , on obtient par exemple :

$$L = \begin{pmatrix} \nu \\ e \end{pmatrix}, e_R, Q = \begin{pmatrix} u \\ d \end{pmatrix}, u_R, d_R \quad (1.26)$$

Pour rappeler que cette interaction ne s'applique qu'aux champs gauches (Left), on note le groupe de symétrie de la théorie  $SU(2)_L$ . Il faut par ailleurs noter que les neutrinos interagissent uniquement par interaction faible et sont de masse nulle dans

le MS. Comme l'interaction faible agit seulement sur les états gauches, il ne peut exister de neutrinos droits.

Cette symétrie  $SU(2)_L$  n'était cependant pas satisfaisante. Le troisième générateur  $\sigma_3$  ne correspondait à aucun boson connu (Le Z n'avait pas encore été découvert). Il ne pouvait pas s'agir du photon car la matrice  $\sigma_3$  impliquait un couplage du photon aux neutrinos qui n'avait jamais été observé. C'est ainsi que naquit la théorie électrofaible basée sur le groupe  $SU(2)_L \times U(1)_Y$  qui permet de construire de façon réaliste les photons et les bosons Z comme des combinaison de champs de SU(2) et U(1). On introduit de plus l'hypercharge, nouveau nombre quantique, combinaison linéaire de la charge électrique Q et de la troisième composante de l'isospin  $I_3$  :

$$Y = 2(Q - I_3) \tag{1.27}$$

L'hypercharge des composants du doublet  $\begin{pmatrix} \nu \\ e \end{pmatrix}$  est ainsi -1 alors que celle du doublet  $\begin{pmatrix} u \\ d \end{pmatrix}$  est  $\frac{1}{3}$ . Il faut par ailleurs ajouter que les états propres de masse sont différents des états propres de l'interaction faible. On introduit donc la matrice CKM (Cabbibbo-Kobayashi-Maskawa) qui contient toute l'information nécessaire pour passer des uns aux autres. Les coefficients de cette matrice ont été déterminés expérimentalement et affinés grâce à de nombreux groupes de recherche au cours des 50 dernières années. Ils quantifient la probabilité d'un changement de saveur lors de la désintégration d'une particule et autorisent la violation de CP [11].

### 1.2.2.3 Le lagrangien électrofaible sans masses

Nous appellerons  $B_\mu$  le champ de jauge correspondant à la symétrie  $U(1)_Y$ . Comme dans l'exemple précédent de la QED, pour garder le lagrangien invariant par la transformation locale  $\psi(x) \rightarrow e^{i\beta(x)}\psi(x)$ , on introduit ce champ  $B_\mu$  avec le couplage  $g'$  qui se transforme comme suit :  $B_\mu(x) \rightarrow B_\mu(x) + \frac{1}{g'}\partial_\mu\beta(x)$ . De même, on définit les 3 champs de jauge  $W_\mu^i$  correspondants au groupe  $SU(2)_L$ . La symétrie locale associée à ce même groupe ne s'applique qu'aux composantes gauches :

$$\psi_L(x) \rightarrow e^{i\frac{\sigma_i}{2}\beta_i(x)}\psi_L(x) \tag{1.28}$$

Les trois nouveaux champs sont associés au couplage g et se transforment comme suit :

$$W_\mu^i \rightarrow W_\mu^i + \frac{1}{g}\partial_\mu\beta^i + \varepsilon^{ijk}W_{\mu j}\beta_k \tag{1.29}$$

Enfin, la dérivée covariante se présente comme telle :

$$D_\mu = \partial_\mu - ig'\frac{Y}{2}B_\mu(x) - ig\frac{\sigma_i}{2}W_\mu^i(x) \tag{1.30}$$

On définit les tenseurs correspondants aux 4 nouveaux champs de jauge similairement au tenseur électromagnétique :

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \text{ et } W_{\mu\nu} = W_{\mu\nu}^i \frac{\sigma_i}{2} \tag{1.31}$$

avec :  $W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\varepsilon^{ijk}W_{\mu j}W_{\nu k}$

Si on ne prend pas en compte les masses, le lagrangien électrofaible avant brisure de symétrie se réduit donc à l'équation 1.32.

$$\mathcal{L} = \bar{\psi}i\gamma^\mu D_\mu\psi - \frac{1}{4}(B^{\mu\nu}B_{\mu\nu} + W_{\mu\nu}^iW_i^{\mu\nu}) \quad (1.32)$$

Il ne s'agit cependant pas du lagrangien complet de la théorie électrofaible. En effet il présente deux problèmes liés aux masses des particules qu'il prédit. Tout d'abord, il manque un terme qui décrit la masse des nouveaux bosons de jauge massifs. Le terme générique pour la masse d'un boson est de la forme  $-\frac{1}{2}M_V^2 V_\mu V^\mu$ . Or ce terme ne serait pas invariant par une symétrie locale de  $SU(2)$  qui ne change que les états gauches des particules. Il faut donc trouver un moyen d'incorporer un terme de masse pour les bosons faibles dans le lagrangien. Ensuite, il manque aussi le terme de masse pour les fermions. Là encore, le terme de masse classiquement utilisé pour les fermions  $-m\psi\bar{\psi}$  n'est pas invariant sous les symétries du groupe  $SU(2)_L$ . En effet, si on décompose le spineur selon ces deux états de chiralité ( $\psi = \psi_R + \psi_L$ ), on obtient pour, l'électron par exemple :

$$-m_e\bar{e}e = \bar{e}_R e_L + \bar{e}_L e_R \quad (1.33)$$

qui n'est pas invariant par une transformation de  $SU(2)_L$  car  $e_L$  appartient à un doublet de  $SU(2)_L$  alors que  $e_R$  est un singulet.

On doit donc trouver un moyen d'inclure la masse des bosons faibles et des fermions dans ce lagrangien tout en préservant la symétrie  $SU(2)_L$ .

#### 1.2.2.4 La brisure électrofaible et le secteur de Higgs

En 1964, Peter Higgs mit au point un formalisme, dans lequel les particules acquièrent leurs masse par interaction avec un champ présent dans le vide appelé champ de Higgs. En appliquant ce formalisme à la physique des particules et au MS, on peut expliquer comment les fermions et certains bosons de jauge peuvent acquérir une masse alors que les photons ou les neutrinos non. La masse n'est plus considérée comme une propriété intrinsèque des particules (comme la charge ou le spin) mais elle est fixée par l'interaction avec le champ de Higgs. Le fameux boson de Higgs est en fait la particule qui porte le champ de Higgs.

#### Le mécanisme de Brout Englert Higgs

Tout commence avec la solution au problème des masses du modèle électrofaible exposé ci dessus. Brout, Englert et Higgs ont séparément et presque simultanément découvert que le fait d'ajouter à la théorie un doublet scalaire de champs complexe  $\Phi = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$  et d'hypercharge  $Y=1$  va faire apparaître les termes de masse attendus. Le potentiel du champ scalaire s'écrit alors à une constante près :

$$V_{CS}(\Phi) = \mu^2|\Phi|^2 + \lambda|\Phi|^4 \quad (1.34)$$

avec  $\lambda > 0$  et  $\mu^2 < 0$  deux constantes.

Ce potentiel est minimal pour les valeurs du champ telles que :  $|\Phi|^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}$ . Ces valeurs correspondent au puits de potentiel représenté figure 1.1. La forme générale du potentiel est symétrique par rotation autour de l'axe Z, cependant, dès qu'une valeur du champ est fixée, la symétrie est brisée spontanément (voir partie 1.1.6). Le potentiel étant défini à une constante près, on peut le réécrire de la manière suivante :

$$V_{BS}(\Phi) = \lambda(|\Phi|^2 - \frac{1}{2}v^2)^2 \quad (1.35)$$

Comme on peut choisir arbitrairement l'état du vide, on peut le considérer comme  $\langle \Phi \rangle = \frac{1}{\sqrt{2}}\Phi = \begin{pmatrix} 0 \\ v \end{pmatrix}$ . En choisissant la jauge appropriée, on peut maintenant annuler les composantes du vide  $\phi_{1,2,4}$  et tout exprimer selon l'axe portant  $\phi_3$ . Finalement, on obtient la perturbation et la valeur moyenne dans le vide selon la même direction, ce qui donne le champ suivant :

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (1.36)$$

En injectant cette nouvelle expression du champ à l'intérieur du potentiel 1.34 on fait ressortir naturellement la masse du champ h ainsi que ses couplages :

$$V_{BS}(h) = \lambda(v^2h^2 + vh^3 + \frac{h^4}{4}) = \frac{m_h^2}{2}h^2 - \frac{ig_h^3}{3!}h^3 - \frac{ig_h^4}{4!}h^4 \quad (1.37)$$

Le champ h(x) décrit ainsi une nouvelle particule appelé boson de Higgs dont la masse est  $m_h = \sqrt{2\lambda v^2} = \sqrt{2}\mu$ .

### La masse des bosons de l'interaction faible

Le lagrangien correspondant au champ scalaire  $\Phi$  s'écrit maintenant :

$$\mathcal{L} = |D_\mu\Phi|^2 - \lambda(|\phi|^2 - \frac{1}{2}v^2)^2 - \frac{1}{4}(B_{\mu\nu}B^{\mu\nu} + W_{\mu\nu}^iW_i^{\mu\nu}) \quad (1.38)$$

La valeur moyenne du vide  $\frac{v}{\sqrt{2}}$  génère la masse des bosons de l'interaction faible par l'intermédiaire de la dérivée covariante :

$$|D_\mu\phi|^2 = |(\partial_\mu - ig'\frac{Y}{2}B_\mu(x) - ig\frac{\sigma_i}{2}W_\mu^i(x))|^2 \quad (1.39)$$

En développant cette expression et en sommant sur les matrices de Pauli, on obtient les 4 champs de jauge pour les bosons électrofaibles :

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2) \\ Z_\mu &= \frac{1}{\sqrt{g'^2 + g^2}}(gW_\mu^3 - g'B_\mu) \\ A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}}(g'W_\mu^3 + gB_\mu) \end{aligned}$$

Ces potentiels correspondent exactement aux termes de masse des trois champs associés aux bosons massifs. Pour  $W_\mu^\pm$  on trouve  $M_W = \frac{gv}{2}$  et pour  $Z_\mu$ ,  $M_Z = \frac{v}{2}\sqrt{g^2 + g'^2}$ . Le champs  $A_\mu$  n'est associé à aucune masse et permet une fois pour toute de l'identifier au photon.

On peut alors réécrire la dérivée covariante sous une forme plus explicite pour faire apparaître les masses des bosons :

$$|D_\mu\phi|^2 = \frac{1}{2}\partial_\mu h\partial^\mu h + \left(\frac{1}{2}g(v+h)\right)^2 W_\mu^+ W^{\mu-} + \frac{1}{2}\left(\frac{v+h}{2}\sqrt{g'^2 + g^2}\right)^2 Z^\mu Z_\mu \quad (1.40)$$

Pour simplifier les notations, on introduit l'angle de Weinberg  $\theta_W$  défini de façon univoque par :

$$\cos(\theta_W) = \frac{g}{\sqrt{g'^2 + g^2}} \text{ et } \sin(\theta_W) = \frac{g'}{\sqrt{g'^2 + g^2}} \quad (1.41)$$

Il en résulte de nouvelles expressions pour les champs des bosons massifs :

$$\begin{aligned} Z_\mu &= \cos(\theta_W)W_\mu^3 - \sin(\theta_W)B_\mu \\ A_\mu &= \cos(\theta_W)B_\mu - \sin(\theta_W)W_\mu^3 \\ \frac{M_W}{M_Z} &= \cos(\theta_W) \end{aligned}$$

Il faut noter que les dernières mesures expérimentales  $M_Z = 91,1876 \pm 0.0021$  GeV et  $M_W = 80,385 \pm 0.015$  GeV [12] sont en excellent accord avec les prédictions du modèle standard.

### La masse des fermions

Nous avons déjà vu plus haut que le terme de masse classique  $-m\bar{f}f$  ne peut être utilisé car il briserait l'invariance locale  $SU(2)_Y$ . Le champ de Higgs va en fait solutionner le problème de manière naturelle grâce à l'introduction du couplage de Yukawa :

$$\begin{aligned} \mathcal{L}_Y &= -\lambda_f(\bar{f}_R\Phi^+ f_L + \bar{f}_L\Phi f_R) \\ \mathcal{L}_Y &= -\frac{\lambda_f v}{\sqrt{2}}\bar{f}f - \frac{\lambda_f}{\sqrt{2}}\bar{f}fh \end{aligned}$$

Ces nouveaux couplages font apparaître naturellement la masse des fermions ainsi que les termes d'interaction Higgs/fermions. Le couplage qui correspond à ces interactions est proportionnel à la masse du fermion. On retrouve le résultat annoncé dans le préambule de cette section, les particules de matière acquièrent leur masse par interaction avec le champ de Higgs qui a une valeur non nulle dans le vide.

	Génération			Nombres quantiques					
	I	II	III	$U(1)_Y$	$SU(2)_L$	$SU(3)_c$	Q	I	$I_3$
$L_L^i$	$\begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}$	$\begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}$	-1/2	2	1	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	1/2	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
$e_R^i$	$e_R$	$\mu_R$	$\tau_R$	-1	1	1	-1	0	0
$Q_L^i$	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$	1/6	2	3	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	1/2	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$
$u_R^i$	$u_R$	$c_r$	$t_R$	2/3	1	3	+2/3	0	0
$d_R^i$	$d_R$	$s_r$	$b_R$	-1/3	1	3	-1/3	0	0

TABLE 1.1 – Les champs fermioniques du Modèle standard et les nombres quantiques associés. I et  $I_3$  représentent respectivement l'isospin faible et sa troisième composante ; Q est la charge électrique.

### 1.2.3 Le lagrangien du modèle standard

Quand on rajoute la chromodynamique quantique et donc le groupe  $SU(3)_c$ , on obtient par les mêmes mécanismes que précédemment un lagrangien invariant sous une transformation de  $SU(3)_c \times SU(2)_L \times U(1)_Y$ . La dérivée covariante devient :

$$D_\mu = \partial_\mu - ig' \frac{Y}{2} B_\mu(x) - ig \frac{\sigma_i}{2} W_\mu^i(x) - ig_c \frac{\lambda_i}{2} G_\mu^i(x) \quad (1.42)$$

Le terme cinétique gagne aussi un terme provenant de la QCD :

$$\mathcal{L}_g = -\frac{1}{4} (B_{\mu\nu} B^{\mu\nu} + W_{\mu\nu}^i W_i^{\mu\nu} + G_{\mu\nu}^i G_i^{\mu\nu}) \quad (1.43)$$

Et finalement, on peut écrire le lagrangien complet du modèle standard :

$$\mathcal{L}_{SM} = \mathcal{L}_f + \mathcal{L}_Y + \mathcal{L}_g + \mathcal{L}_H \quad (1.44)$$

avec les notations précédemment employées,  $\mathcal{L}_f$  est le terme de propagation fermionique,  $\mathcal{L}_Y$  est le terme de Yukawa qui donne leur masse aux fermions,  $\mathcal{L}_g$  le terme cinétique et enfin  $\mathcal{L}_H$  le terme de Higgs qui fournit le champ de Higgs ainsi que la masse des bosons de l'interaction faible.

Pour finir avec cette présentation du modèle standard, le tableau 1.1 récapitule l'ensemble des champs fermioniques et leurs nombres quantiques associés.

## 1.3 Les problèmes du Modèle Standard

En dépit des nombreux succès du modèle standard, plusieurs raisons nous poussent à croire qu'il ne s'agit que d'une théorie effective valable uniquement dans un certain intervalle d'énergie. Ainsi, la nouvelle fenêtre d'énergie exploitable par le LHC devrait permettre de révéler dans un avenir proche de nouvelles structures plus complexes que celles prédites par le MS. Dans cette partie nous dresserons une liste non exhaustive des raisons qui laissent supposer que le MS est incomplet.

### Les paramètres libre du MS

Tout d'abord, il contient 19 paramètres internes non prédits qu'il faut évaluer à partir des expériences. Il s'agit des masses des 9 fermions, des 3 angles de mélanges de la matrice CKM, de la phase de violation de CP, des 3 couplages de jauge, des 2 paramètres du secteur de Higgs et enfin l'angle du vide de la QCD. Idéalement, on voudrait mettre au point une théorie qui prévoit directement les valeurs exactes de ces paramètres sans qu'ils soient des paramètres d'entrée du modèle.

### L'oscillation des neutrinos

Ensuite, diverses expériences sur les neutrinos solaires et atmosphériques (voir [13] pour une revue récente) ont démontré l'oscillation de saveur entre les neutrinos. Ces changements de saveur nécessitent l'existence de neutrinos droits qui n'existent pas dans le formalisme du MS (voir partie 1.2.2) et impliquent que les neutrinos soient massifs (alors que le MS les prédit sans masse).

### Le traitement de la gravitation

La théorie de la gravité, la relativité générale n'est pas renormalisable. Cela provient du fait que la constante de Newton qui caractérise les interactions n'est pas adimensionnelle comme c'est le cas pour les interactions du modèle standard. Les effets de la gravité quantique deviennent non négligeables à partir de l'échelle de Planck (de l'ordre de  $10^{19}$  GeV). Ce qui montre à coup sûr que le modèle standard n'est plus valable au delà de cette échelle. Le MS est donc une théorie effective valable uniquement pour un intervalle d'énergies donné.

### Le problème de hiérarchie des masse

Ceci n'est pas un problème à proprement parlé mais plutôt une insatisfaction intellectuelle. En effet il n'existe pas d'explication théorique de la structure en famille des fermions ni de leur nombre. Les disparités dans les masses et dans les couplages des particules restent aussi des paramètres d'entrée du MS et on peut espérer l'existence d'une structure plus globale qui permettrait d'expliquer cette architecture si singulière.

### Le problème de naturalité

Pour calculer précisément la masse du Higgs, il faut inclure les contributions des boucles de fermions qui sont quadratiquement divergentes. Si on développe la masse du Higgs comme suit :  $m_H^2 = m_0^2 - \delta m_H^2$  avec  $m_0$  la masse nue. Il faut ajuster  $m_0$  à plusieurs décimales près pour obtenir des masses de Higgs acceptables ce qui ne paraît pas naturel du tout.

**L'unification des interactions à haute énergie** Comme précédemment, ce problème n'est pas vraiment une preuve de l'invalidité du Modèle standard mais plutôt une insatisfaction formelle pour les théoriciens. Si l'on trace l'évolution des constantes de couplage en fonction de l'échelle d'énergie on s'aperçoit que les courbes

tendent à se rapprocher sans jamais vraiment converger. Pour continuer le travail d'unification engagé depuis plus d'un siècle, il paraîtrait plaisant intellectuellement de voir ces couplages converger à haute énergie. Cela tendrait à prouver que toutes les interactions sont issues d'une seule et même force. De ce point de vue, le modèle standard apparaît comme une impasse sur la voie recherche de la grande unification.

Pour conclure cette partie, il faut rappeler que le modèle standard, malgré tous ses succès, reste une théorie effective. Les progrès réalisés par le LHC et son proche redémarrage permettront sans doute d'explorer les limites du domaine de validité du MS. On peut ainsi espérer que les disparités entre l'expérience et la théorie se fassent de plus en plus nombreuses. Ainsi elles aiguilleront petit à petit les théoriciens vers une théorie plus générale qui permettra de régler une partie des problèmes précédemment cités.

## 1.4 Introduction théorique aux leptoquarks

Comme nous venons de le voir, certaines divergences avec les prédictions du modèle standard mènent à penser qu'il existe un modèle plus général qui engloberait le modèle standard. Parmi tous ces candidats à l'extension du modèle standard, une grande partie prévoit l'existence de bosons qui portent à la fois un nombre leptonique et un nombre baryonique ainsi qu'une charge électrique. Ces nouveaux bosons, triplet de couleur, sont appelés leptoquarks (LQ) et seraient médiateurs d'une interaction entre leptons et quarks [14]. On peut alors construire un modèle effectif le plus général possible pour rechercher des leptoquarks génériques prédits par différentes théories. En s'appuyant sur les résultats expérimentaux de changement de saveur par courant neutre, la violation du nombre leptonique ou de la désintégration du proton on a supposé historiquement que les leptoquarks ne se couplaient pas avec les particules appartenant à d'autres familles. Ceci mena naturellement à la supposition qu'il existait 3 familles de leptoquarks répondant aux 3 familles de quarks et de leptons.

### 1.4.1 Un modèle effectif pour la recherche de LQ

Les LQ sont prédits par de nombreux modèles au delà du modèle standard (SUSY, Technicouleur, supercordes...). Pour simplifier leur recherche, nous allons nous placer dans un modèle effectif en nous basant sur quelques hypothèses de base. Ce modèle a été développé par Buchmüller-Rückl-Wyler (BRW) dans les années 80 [15]. Il s'agit d'un modèle inclusif le plus général possible afin d'être indépendant du modèle prédisant les LQ. La théorie BRW part des hypothèses suivantes sur les interactions entre LQ :

1. Elles conservent les nombres leptonique  $L$  et baryonique  $B$  séparément.  
En effet dans le cas où ils ne le sont pas, la limite inférieure sur les masses de LQ est très importante ( $> 10^{15}$  GeV). Ce chiffre provient des limites expérimentales sur la stabilité du proton ( $> 10^{40}$  s). Si ces nombres ne sont

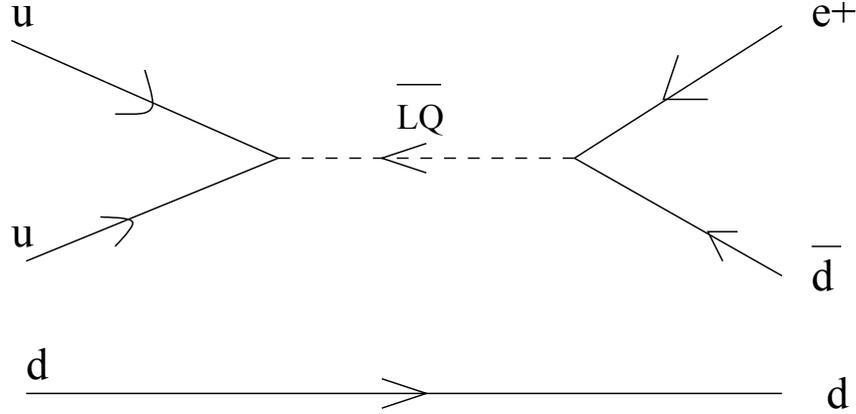


Fig. 1.2 – Une désintégration possible du proton par l'intermédiaire d'un LQ.

pas conservés, le proton pourrait se désintégrer par l'intermédiaire d'un LQ selon le diagramme 1.2.

2. Elles sont invariantes par les symétries du groupe de jauge du MS :  $U(1)_Y \times SU(2)_L \times SU(3)_c$ .
3. Elles sont renormalisables.
4. Elles prévoient un couplage des LQ uniquement avec les fermions et bosons de jauge du MS.

On caractérise les leptoquarks par leur nombre fermionique  $F$  (équation 1.45). Il sera fonction du nombre de fermions et d'anti-fermions qui les composent.

$$F = 3B + L \quad (1.45)$$

$$B = \frac{1}{3}(n_q - n_{\bar{q}}) \quad (1.46)$$

$$L = n_l - n_{\bar{l}} \quad (1.47)$$

Les seules possibilités pour  $|F|$  sont 0 ou 2. Les LQ de nombre fermionique  $F=2$  se couplent uniquement à des paires fermion fermion (type  $lq$ ). Ceux pour lesquels  $F=0$  se couplent à des paires fermion anti-fermion (type  $l\bar{q}$ ). On postule de plus que le spin des leptoquarks peut uniquement prendre les valeurs 0 (leptoquark scalaire) ou 1 (leptoquark vecteur). Le tableau 1.2 présente les différentes possibilités pour les leptoquarks dans le cadre de ces hypothèses en fonction de leur nombre fermionique ( $F=0$  ou 2), de leur spin ( $S=0$  ou 1), de leur isospin faible ( $I=0,1$  ou  $\frac{1}{2}$ ) et enfin de leur charge électrique ( $Q=-\frac{4}{3}, -\frac{1}{3}, \frac{2}{3}$  ou  $\frac{5}{3}$ ).

#### 1.4.2 Lagrangien effectif et couplages

Le lagrangien le plus général pour le secteur des leptoquarks satisfaisant les hypothèses ci dessus devient [15] :

Type	F	$U(1)_Y$	$SU(2)_L$	$SU(3)_c$	$I_3$	$Q_{em}$	$\lambda_L^{lq}$	$\lambda_R^{lq}$	$\lambda_L^{\nu q}$	Couplages
$\tilde{S}_1$	-2	1/3	1	$\bar{3}$	0	+1/3	$g_{1L}$	$g_{1R}$	$-g_{1L}$	$\bar{q}_L^c \ell_L$ ou $\bar{u}_R^c e_R$
$\tilde{S}_1$	-2	4/3	1	$\bar{3}$	0	+4/3	0	$\tilde{g}_{1R}$	0	$\bar{d}_R^c e_R$
$\tilde{S}_3$	-2	1/3	3	$\bar{3}$	+1	+4/3	$-\sqrt{2}g_{3L}$	0	0	$\bar{q}_L^c \ell_L$
					0	+1/3	$-g_{3L}$	0	$-g_{3L}$	
					-1	-2/3	0	0	$\sqrt{2}g_{3L}$	
$V_{2\mu}$	-2	5/6	2	$\bar{3}$	+1/2	+4/3	$g_{2L}$	$g_{2R}$	0	$\bar{q}_L^c \gamma^\mu e_R$ ou $\bar{d}_R^c \gamma^\mu \ell_L$
					-1/2	+1/3	0	$g_{2R}$	$g_{2L}$	
$\tilde{V}_{2\mu}$	-2	-1/6	2	$\bar{3}$	+1/2	+4/3	$\tilde{g}_{2L}$	0	0	$\bar{u}_R^c \gamma^\mu \ell_L$
					-1/2	+1/3	0	0	$\tilde{g}_{2L}$	
$R_2$	0	7/6	2	3	+1/2	+5/3	$h_{2L}$	$h_{2R}$	0	$\bar{q}_L e_R$ ou $\bar{u}_R \ell_L$
					-1/2	+2/3	0	$-h_{2R}$	$h_{2L}$	
$\tilde{R}_2$	0	1/6	2	3	+1/2	+2/3	$\tilde{h}_{2L}$	0	0	$\bar{d}_R^c \ell_L$
					-1/2	-1/3	0	0	$\tilde{h}_{2L}$	
$U_{1\mu}$	0	2/3	1	3	0	+2/3	$h_{1L}$	$h_{1R}$	$h_{1L}$	$\bar{q}_L \gamma^\mu \ell_L$ ou $\bar{d}_R \gamma^\mu e_R$
$\tilde{U}_{1\mu}$	0	5/3	1	3	0	+5/3	0	$\tilde{h}_{1R}$	0	$\bar{u}_R \gamma^\mu e_R$
					+1	+5/3	$\sqrt{2}h_{3L}$	0	0	
$\tilde{U}_{3\mu}$	0	2/3	3	3	0	+2/3	$-h_{3L}$	0	$h_{3L}$	$\bar{q}_L \gamma^\mu \ell_L$
					-1	-1/3	0	0	$\sqrt{2}h_{3L}$	

TABLE 1.2 – Les différentes possibilités de leptosquarks et leur nombre quantiques associés. Les colonnes  $U(1)_Y$ ,  $SU(2)_L$  et  $SU(3)_c$  représentent respectivement les représentations d'hypercharge, d'isospin et de la QCD.

$$\mathcal{L}_{LQ} = \mathcal{L}_{|F|=0}^f + \mathcal{L}_{|F|=2}^f + \mathcal{L}^V \quad (1.48)$$

avec  $\mathcal{L}_{|F|=0,2}^f$  les lagrangiens décrivant l'interaction avec les fermions du MS et  $\mathcal{L}^V$  représentant celle avec les bosons de jauge du MS. Par exemple, les termes caractérisant l'interaction des LQ scalaires avec des quarks et leptons sont :

$$\begin{aligned} \mathcal{L}_{|F|=0}^f &= (h_{2L}\bar{u}_R\ell_L + h_{2R}\bar{q}_L i\sigma_2 e_R)R_2 + \tilde{h}_{2L}\bar{d}_R\ell_L\tilde{R}_2 \\ &\quad + (h_{1L}\bar{q}_L\gamma^\mu\ell_L + h_{1R}\bar{d}_R\gamma^\mu e_R)U_{1\mu} + \tilde{h}_{1R}\bar{u}_R\gamma^\mu e_R\tilde{U}_{1\mu} \\ &\quad + h_{3L}\bar{q}_L\vec{\sigma}\gamma^\mu\ell_L\vec{U}_{3\mu} + h.c \\ \mathcal{L}_{|F|=2}^f &= (g_{1L}\bar{q}_L^c i\sigma_2\ell_L + g_{1R}\bar{u}_R^c e_R)S_1 + \tilde{g}_{1R}\bar{d}_R^c e_R\tilde{S}_1 + g_{3L}\bar{q}_L^c i\sigma_2\vec{\sigma}\ell_L\vec{S}_3 \\ &\quad + (g_{2L}\bar{d}_R^c\gamma^\mu\ell_L + g_{2R}\bar{q}_L^c\gamma^\mu e_R)V_{2\mu} + \tilde{g}_{2L}\bar{u}_R^c\gamma^\mu\ell_L\tilde{V}_{2\mu} + h.c. \end{aligned}$$

avec  $\sigma_i$  les matrices de Pauli,  $q_L$  et  $\ell_L$  les doublets d'interaction de  $SU(2)_L$  et  $e_R$ ,  $d_R$ ,  $u_R$  les singulets droit correspondants pour les leptons, les quarks de type down et de type up.

Si maintenant on considère uniquement les termes caractérisant l'interaction des LQ scalaires avec des quarks et leptons de troisième génération, le lagrangien effectif devient :

$$\begin{aligned} \mathcal{L}_{eff}^f &= (h_{2L}\bar{t}_R\ell_L + h_{2R}\bar{q}_L i\sigma_2 e_R)R_2 + \tilde{h}_{2L}\bar{b}_R\ell_L\tilde{R}_2 \\ &\quad + (g_{1L}\bar{q}_L^c i\sigma_2\ell_L + g_{1R}\bar{t}_R^c\tau_R)S_1 + \tilde{g}_{1R}\bar{b}_R^c\tau_R\tilde{S}_1 + g_{3L}\bar{q}_L^c i\sigma_2\vec{\sigma}\ell_L\vec{S}_3 \end{aligned}$$

avec  $q_L$  et  $\ell_L$  les doublets d'interaction de  $SU(2)_L$  et  $t_R$ ,  $b_R$ ,  $\tau_R$  les singulets droits correspondants de 3ème génération.

Les leptoquarks interagissent aussi avec les bosons de jauge du modèle standard. Dans le cas de l'interaction des LQ avec les gluons, les lagrangiens pour les cas scalaire (1.49) et vecteur (1.50) sont donnés par :

$$\mathcal{L}_S^g = \sum_{\text{scalaires}} \left[ (D_{ij}^\mu \psi^j)^\dagger (D_\mu^{ik} \psi_k) - m_S^2 \psi^{i\dagger} \psi_i \right] \quad (1.49)$$

$$\begin{aligned} \mathcal{L}_V^g &= \sum_{\text{vecteurs}} \left[ -\frac{1}{2} V_{\mu\nu}^{i\dagger} V_i^{\mu\nu} + M_V^2 \phi_\mu^{i\dagger} \psi_i^\mu \right. \\ &\quad \left. - i g_s \left( (1 - k_G) \psi_\mu^{i\dagger} t_{ij}^a \psi_\nu^j \mathcal{G}_a^{\mu\nu} + \frac{\lambda_G}{M_V^2} V_{\sigma\mu}^{i\dagger} t_{ij}^a V_\nu^{j\mu} \mathcal{G}_a^{\nu\sigma} \right) \right] \quad (1.50) \end{aligned}$$

avec  $D_{ij}^\mu = \partial_\mu \delta^{ij} - i g_s t_{ij}^a G_\mu^a$  la dérivée covariante,  $V_i^{\mu\nu} = D_\mu^{ij} \psi_{\nu k} - D_\nu^{ik} \psi_{\mu k}$  le champ des leptoquarks,  $\mathcal{G}_a^{\mu\nu} = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_{\mu b} G_{\nu c}$  le tenseur de force des gluons et  $m_S$  ( $m_V$ ) la masse des LQ scalaires (vecteurs). Enfin,  $k_G$  et  $\lambda_G$  représentent les couplages anormaux de la théorie qui peuvent apparaître aux vertex  $gLQ_V LQ_V$  et  $ggLQ_V LQ_V$ . Ils sont supposés réels et sont reliés aux moments magnétique anormal  $\mu_V$  et quadrupolaire électrique  $q_V$  par les relations suivantes :

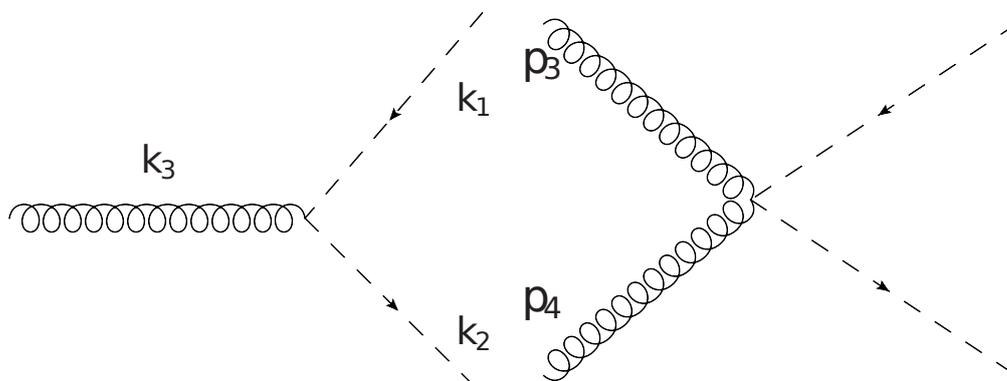


Fig. 1.3 – Vertex d’interaction gluon avec LQ  $\bar{L}Q$  (à gauche) gluon gluon avec LQ  $LQ$  (à droite).

$$\mu_V = \frac{g_s}{2M_V}(2 - k_G + \lambda_G)$$

$$q_V = -\frac{g_s}{M_V^2}(1 - k_G - \lambda_G)$$

Le lagrangien pour les LQ vecteurs dépend donc du modèle par l’intermédiaire de ces couplages anormaux. Deux modèles sont habituellement envisagés, celui de Yang-Mills avec  $k_G = 0$   $\lambda_G = 0$  et celui de couplage vectoriel minimum avec  $k_G = 1$   $\lambda_G = 0$  [16]. Dans les deux cas scalaire et vecteur, ce lagrangien permet des vertex d’interactions à 3 corps LQ-LQ-g et à 4 corps LQ-LQ-g-g 1.3. Il est à noter que les leptoquarks se couplent également aux autres bosons de jauge du MS mais ces interactions peuvent être négligées en collisionneur hadronique. On rajoute souvent comme hypothèse que les LQ se couplent uniquement aux leptons et quarks de la même famille mais ce n’est pas obligatoire. La théorie exige seulement qu’un type de LQ ne se couple qu’à une seule famille de lepton et une seule de quarks (mais celles si peuvent être différentes). Par exemple un LQ peut se coupler avec des quarks de 3ème génération et des leptons appartenant à la première.

### 1.4.3 Désintégration des LQ

On définit la quantité  $\beta$  comme le rapport de branchement du LQ en un lepton chargé et un quark ( $\beta = BR(LQ \rightarrow \ell q)$ ). La quantité  $(1 - \beta)$  est alors associée au rapport de branchement en un neutrino et un quark ( $(1 - \beta) = BR(LQ \rightarrow \nu q)$ ).

Dans le cas d’une production de LQ par paires, les rapports de branchements des différents canaux sont donc les suivants :

$$BR(LQ\bar{L}Q \rightarrow \ell\ell q\bar{q}) = \beta^2$$

$$BR(LQ\bar{L}Q \rightarrow \nu\ell q\bar{q}) = 2\beta(1 - \beta)$$

$$BR(LQ\bar{L}Q \rightarrow \nu\nu q\bar{q}) = (1 - \beta)^2$$

#### 1.4.4 État de l'art

En principe, le meilleur endroit pour chercher des LQ scalaires reste les collisionneurs leptons protons car ils peuvent directement être produits et étudiés par l'intermédiaire du couplage lepton/quark des LQ (couplage de Yukawa). Cependant, pour de faibles couplages (ou de grande masses de LQ) les collisionneurs hadroniques semblent plus appropriés. En effet les leptoquarks scalaires peuvent être produits par interaction forte même à bas couplage de Yukawa. Les limites sur les leptoquarks peuvent être obtenus de façon directe ou indirecte. Les limites directes sont obtenues à partir des section efficaces de production dans les collisionneurs. Les limites indirectes sont calculées grâce aux interactions à 4 fermions induites par des leptoquarks dans des expériences à basse énergie [17]. Des recherches directes de leptoquarks ont été menées dans différents collisionneurs comme le LEP, HERA et le TeVatron et le LHC, aucune preuve de leur existence n'a pu à ce jour être mise en évidence. Les meilleures limites sur la production de LQ scalaires de première et deuxième génération ont été déterminées par les expériences CMS [18] et ATLAS [19]. En ce qui concerne les LQ de 3ème génération, on peut trouver un résumé des études déjà publiées dans [20]. Avant le LHC, la limite la plus forte sur la masse des LQ de troisième génération était de 210 GeV. Elle a été fournie par la collaboration D0 en 2008 [21]. Des publications récentes ont étudié les LQ de troisième génération se désintégrant à 100% en un quark b et un  $\tau$  pour le run 2011 à 7 TeV du LHC. L'expérience CMS a permis d'exclure ces LQ jusqu'à 525 GeV [22] et ATLAS jusqu'à 534 GeV [23]. Le canal de désintégration  $LQ_3 \rightarrow b\nu_\tau$  a aussi été étudié dans CMS [24] ce qui a conduit à une limite d'exclusion de 450 GeV. Jusqu'à aujourd'hui, aucune étude n'avait été menée au LHC pour étudier les LQ de troisième génération se désintégrant en un quark top et un lepton. En effet, à cause de la paire  $t\bar{t}$  produite par la désintégration de ces LQ, le signal peut difficilement être mis en évidence car très proche du fond du MS  $pp \rightarrow t\bar{t}$ . Avec la montée en énergie du LHC en 2013, une fenêtre d'étude de ce mode de désintégration s'est ouverte. Voilà pourquoi j'ai choisi de mener à bien cette étude pour la partie analyse de ma thèse. Selon le type de lepton produit avec le quark top, des ajustements sont nécessaires pour séparer au mieux le signal du bruit de fond. Ces choix sont explicités en détails dans la dernière partie de cette thèse (partie 4).



# Dispositif expérimental

---

## Sommaire

---

<b>2.1</b>	<b>Le Grand Collisionneur de Hadrons (LHC)</b>	<b>23</b>
2.1.1	Phénoménologie des collisions protons protons	25
2.1.2	Le complexe d'accélération	25
2.1.3	Luminosité et nombre d'événements	27
2.1.4	Le défi de l'empilement (PU)	27
2.1.5	Quelle physique au LHC?	29
<b>2.2</b>	<b>L'expérience Compact Muon Solenoid (CMS)</b>	<b>31</b>
2.2.1	Le solénoïde	33
2.2.2	Le trajectographe	34
2.2.3	Le calorimètre électromagnétique	36
2.2.4	Le calorimètre hadronique	38
2.2.5	Les chambres à muons	40
<b>2.3</b>	<b>Le système de déclenchement de CMS</b>	<b>44</b>
2.3.1	Le déclenchement de premier niveau L1	46
2.3.2	Le déclenchement de haut niveau HLT	47
<b>2.4</b>	<b>Reconstruction d'événements dans CMS</b>	<b>48</b>
2.4.1	L'algorithme de Particle Flow	48
2.4.2	Identification et reconstruction des leptons	50
2.4.3	Les jets	52
2.4.4	Énergie transverse manquante	54
<b>2.5</b>	<b>Traitement de données et grille de calcul</b>	<b>55</b>
2.5.1	CMSSW et le format de données EDM	55
2.5.2	La grille de calcul	56

---

## 2.1 Le Grand Collisionneur de Hadrons (LHC)

Le Grand Collisionneur de Hadrons, ou LHC pour Large Hadron Collider, est le plus grand et le plus puissant accélérateur de particules jamais construit. Les deux anneaux supraconducteurs de ce synchrotron géant sont installés dans un tunnel de 26,7 km de circonférence, anciennement construit pour le LEP (Large Electron Positron collider) sous le site du CERN (Organisation Européenne pour la Recherche Nucléaire). Il est enterré à une moyenne de 100 m, mais sa profondeur varie de 40 m sous la chaîne du Jura à 170 m sous le lac Léman (figure 2.1). Le LHC est

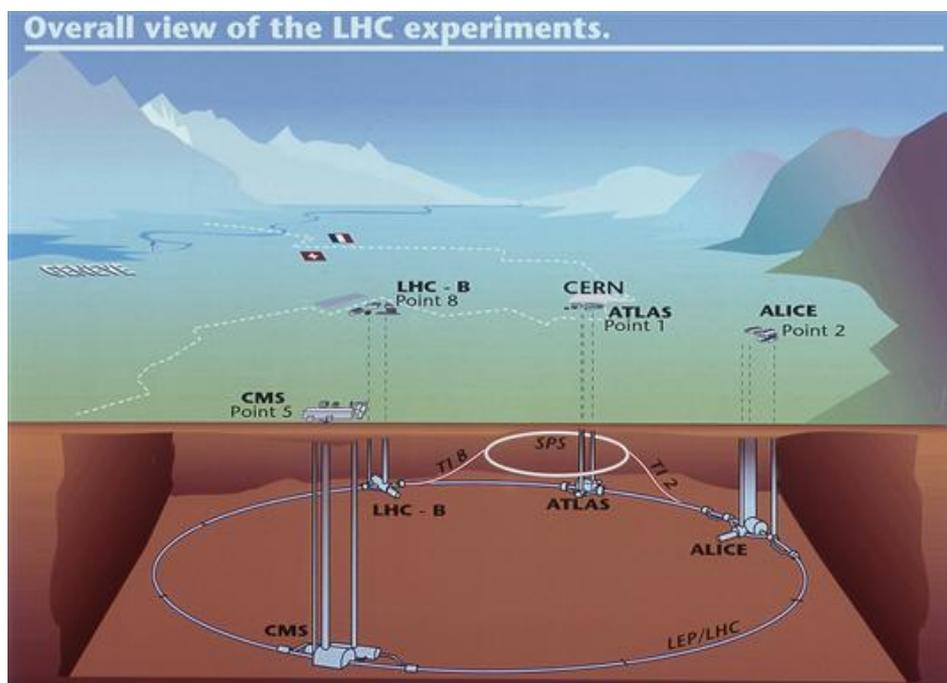


Fig. 2.1 – Vue en coupe du LHC ©CERN.

constitué d'environ 9500 aimants. Les 1232 dipôles sont amenés à une température de 1.9 K en injectant de l'hélium super-fluide dans le circuit de refroidissement. A cette température, les dipôles sont dans un état supraconducteur et ils peuvent produire un champ magnétique nominal de 8,33 T nécessaire pour courber la trajectoire des deux faisceaux de 7 TeV. Les faisceaux se croisent dans 4 zones qui accueillent les 4 expériences principales de physique, ALICE [25], ATLAS [26], CMS [27] et LHCb [28]. Le LHC permet d'atteindre des énergies dans le Centre de Masse (CM) jusqu'ici inégalées. Depuis l'arrêt du TeVatron en 2011, il est d'ailleurs le seul accélérateur géant encore en fonctionnement. Le LHC permet plusieurs types de collisions différentes dont les collisions protons-protons (p-p), proton-plomb (p-Pb) et plomb-plomb (Pb-Pb). Le premier type de collision correspond au mode de fonctionnement principal du LHC et sera le seul décrit dans cette partie.

En septembre 2008, les ingénieurs et les chercheurs du CERN ont réussi à faire circuler les premiers paquets de protons à l'intérieur du LHC offrant ainsi aux détecteurs leurs premières données avec faisceau. Seulement une semaine avant les premières collisions prévues, un incident technique a forcé l'arrêt du LHC. Il fallut ensuite attendre un an d'enquêtes et de vérifications avant de pouvoir relancer l'aventure LHC. Finalement, le 23 novembre 2009, les premières collisions pp furent enregistrées. S'ensuivirent 3 années de fonctionnement intensif avec une montée progressive de la luminosité et un passage de 7 à 8 TeV en 2012. En mars 2013, le LHC fut mis en pause afin de préparer les collisions à 13 TeV prévues pour mars 2015.

### 2.1.1 Phénoménologie des collisions protons protons

Il y a deux grandes classes d'événements en collision pp qui dépendent de l'énergie transférée entre les partons qui interagissent. En effet, à ces énergies, ce sont les partons (quarks et gluons) portant une fraction de l'énergie des protons incidents qui interagissent physiquement. C'est là que se situe la principale différence avec les collisionneurs leptoniques, l'énergie dans le centre de masse de chaque collision n'est pas connue. Cela permet de balayer des gammes en énergie plus importantes et cela fait des collisionneurs hadroniques les instruments idéaux pour la découverte de nouvelles particules dont on ne connaît pas la masse (et plus généralement pour l'étude de nouvelle physique).

Les collisions périphériques entre deux protons donnent lieu à un faible transfert d'impulsion entre les partons. Les particules qui résultent de ces collisions ont donc un faible moment transverse ( $p_t$ ) et sont produites presque dans l'axe des faisceaux. On dit que ce sont des événements de biais minimum (minimum bias events). Statistiquement ils représentent la majorité des collisions pp. Mais du point de vue physique, ce sont les moins pertinents pour la recherche de nouvelle physique.

Les collisions centrales, elles, occasionnent un grand transfert d'impulsion entre les partons (« hard scattering »). Ici, les particules sont produites avec des angles importants par rapport à l'axe des faisceaux (grand  $p_t$ ) et des particules très massives peuvent être créées. Ces événements, très intéressants du point de vue physique, sont cependant très rares. A 8 TeV, la section efficace pp totale est de 101,7 mb<sup>1</sup> et la section efficace pour l'ensemble des processus inélastiques est de 74,7 mb [29]. Une partie du travail d'analyse de données consiste donc à sélectionner ces événements durs parmi des milliards d'événements de biais minimum (voir système de déclenchement dans la partie 2.3).

### 2.1.2 Le complexe d'accélération

La chaîne d'injection des protons du LHC [30] est présentée sur la figure 2.2 ; on décrira ici uniquement le schéma de remplissage pour les collisions protons protons :

1. Les protons sont obtenus à partir d'hydrogène pur en bouteille injecté dans le « Duoplasmatron Proton Ion Source ». L'hydrogène y est ensuite chauffé et ionisé sous forme de plasma. Les électrons sont alors séparés des protons par un puissant champ électrique. Les protons sont enfin regroupés par des électro-aimants pour former les premières ébauches de paquets ;
2. Le Linac 2 (1979), un accélérateur linéaire, récupère les paquets de protons et les amène jusqu'à une énergie cinétique de 50 MeV (31,4% $c$ ) avec une fréquence de 1 Hz ;
3. Le Booster (PSB 1972) est un synchrotron de 157 m de circonférence composé de 4 anneaux superposés de 25 m de rayon chacun. Les protons sont accélérés

---

1. Un barn est défini comme  $10^{-28}$  m<sup>2</sup>

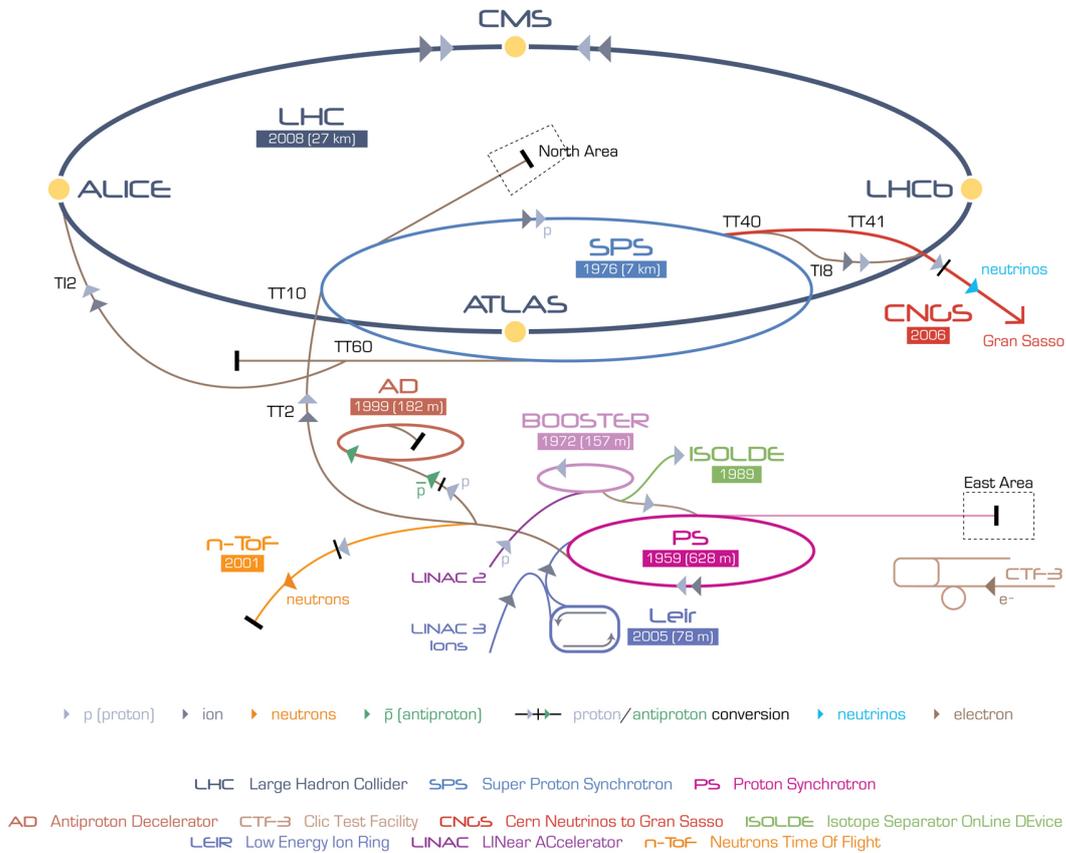


Fig. 2.2 – Schéma du complexe d'accélérateurs du CERN ©CERN.

dans chaque anneau jusqu'à 1,4 GeV (91,6% $c$ ). Puis le contenu de chaque anneau est injecté séquentiellement dans le PS ;

4. Le Synchrotron à Protons (PS 1959) amène les protons jusqu'à une énergie de 25 GeV (99,93% $c$ ) ;
5. Le Super Synchrotron à Protons (SPS 1976) qui a permis la découverte des bosons Z et W en 1983 a une circonférence de 6,9 km. Il est rempli à partir de plusieurs cycles de fonctionnement du PS. Il permet d'accélérer les protons jusqu'à 450 GeV (99,9998% $c$ ) et les injecte finalement dans le LHC.

Il faut 12 cycles du SPS pour remplir le LHC entièrement. Remplir le SPS nécessite de son côté 3 à 4 cycles du PS. Le temps de remplissage du SPS étant de 21,6 s et celui du PS de 3,6 s. Il faut environ 4 minutes pour remplir le LHC de protons. Ensuite, le fait d'élever l'énergie de ces protons de 450 GeV à 4 TeV prend 25 minutes. Le même temps est nécessaire pour descendre l'intensité des aimants après l'arrêt des faisceaux. En prenant en compte les 10 minutes nécessaires à la vérification programmée des systèmes principaux on obtient au final un temps minimum entre deux remplissages d'environ 70 minutes (dans le cas où le faisceau stable serait interrompu immédiatement après avoir été obtenu).

Caractéristiques	Conditions			
	nominales	2010	2011	2012
Énergie par faisceau (TeV)	7	3,5	3,5	4
$\mathcal{L}_{inst}$ max [ $cm^{-2}s^{-1}$ ]	$1 \cdot 10^{34}$	$2,1 \cdot 10^{32}$	$3,7 \cdot 10^{33}$	$7,7 \cdot 10^{33}$
$\mathcal{L}_{int}$	$100 \text{ fb}^{-1}$	$36 \text{ pb}^{-1}$	$4,98 \text{ fb}^{-1}$	$19,7 \text{ fb}^{-1}$
Séparation des paquets (ns)	25	> 150	75/50	50
Nombre de paquets max	2808	< 368	1092	1380
Nb de p max par paquets $\cdot 10^{11}$	1,15	1,2	1,45	1,7
$\beta^*$ (m)	0,55	2,0/3,5	1,5/1,0	0,6

TABLE 2.1 – Configuration du LHC pour les premières années de fonctionnement.

### 2.1.3 Luminosité et nombre d'événements

Le nombre  $N$  d'interactions par seconde pour un processus de section efficace  $\sigma$  est donné par la relation (2.1), avec  $\mathcal{L}$  la luminosité instantanée (en  $cm^{-2}s^{-1}$ ).

$$N = \sigma \times \mathcal{L} \quad (2.1)$$

La luminosité instantanée dépend de la configuration et de la qualité des faisceaux (2.2) :

$$\mathcal{L}_{inst} = \frac{\gamma \cdot f \cdot n_p \cdot N_p^2}{4 \cdot \pi \cdot \varepsilon_n \cdot \beta^*} = \frac{f \cdot n_p \cdot N_p^2}{4 \pi \cdot \sigma_x \cdot \sigma_y} \quad (2.2)$$

où  $\gamma$  est le facteur de Lorentz,  $f$  la fréquence de révolution des paquets,  $n_p$  le nombre de paquets,  $N_p$  le nombre de protons par paquets,  $\varepsilon_n$  l'émittance transverse normalisée,  $\beta^*$  la fonction d'amplitude au point d'interaction et  $\sigma_{x,y}$  les tailles transverses des faisceaux au point d'interaction. La luminosité intégrée est alors définie comme la luminosité accumulée au bout d'un temps donné ( $\mathcal{L} = \int \mathcal{L}_{inst} dt$ ). C'est elle qui caractérise la statistique disponible pour les analyses de données.

### 2.1.4 Le défi de l'empilement (PU)

Plusieurs interactions pp peuvent se produire au cours du même croisement de paquets (figure 2.4). Ces interactions multiples (aussi nommé empilement ou Pile Up) sont en fait des collisions inélastiques qui se superposent à l'événement dur. Le niveau d'empilement dépend de la configuration du LHC. Il est fonction de la fréquence des collisions  $f_c$ , du nombre de protons par paquet, du profil des faisceaux, etc. Le nombre moyen d'interactions est donné par  $\mu = \sigma_{pp} \times \mathcal{L} / f_c$  avec  $f_c = 40 \text{ MHz}$ <sup>2</sup> à luminosité nominale, le PU devrait être de l'ordre de  $\mu \simeq 25$  lors de la reprise du LHC en 2015. Par comparaison, il était en moyenne de 1 en 2010, de 4,5 en 2011 et enfin de 15 pour les années 2012/2013.

<sup>2</sup> 31,7 MHz nominaux plus les temps morts

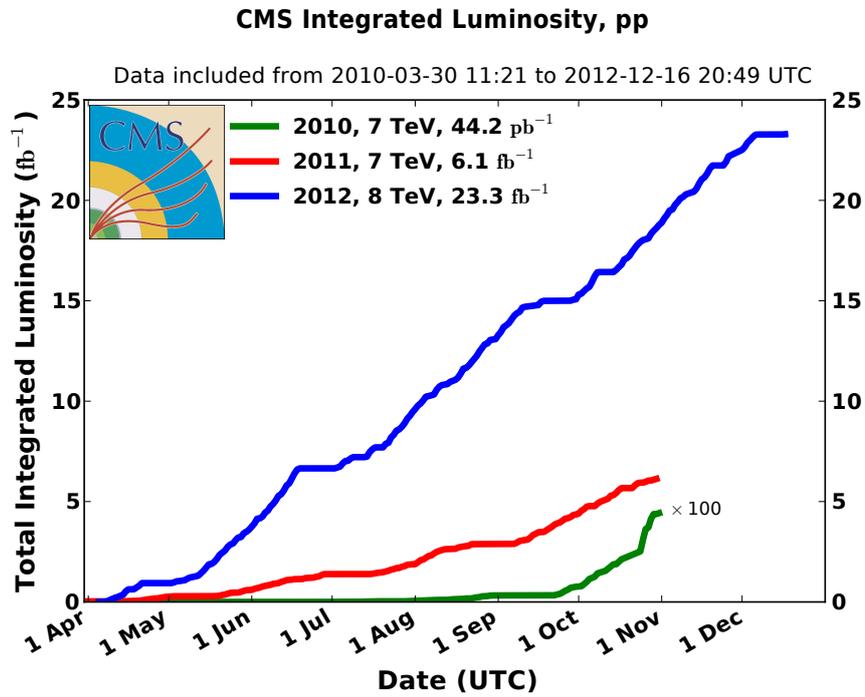


Fig. 2.3 – Luminosité intégrée enregistrée par l'expérience CMS au cours des trois premières années de fonctionnement [31].

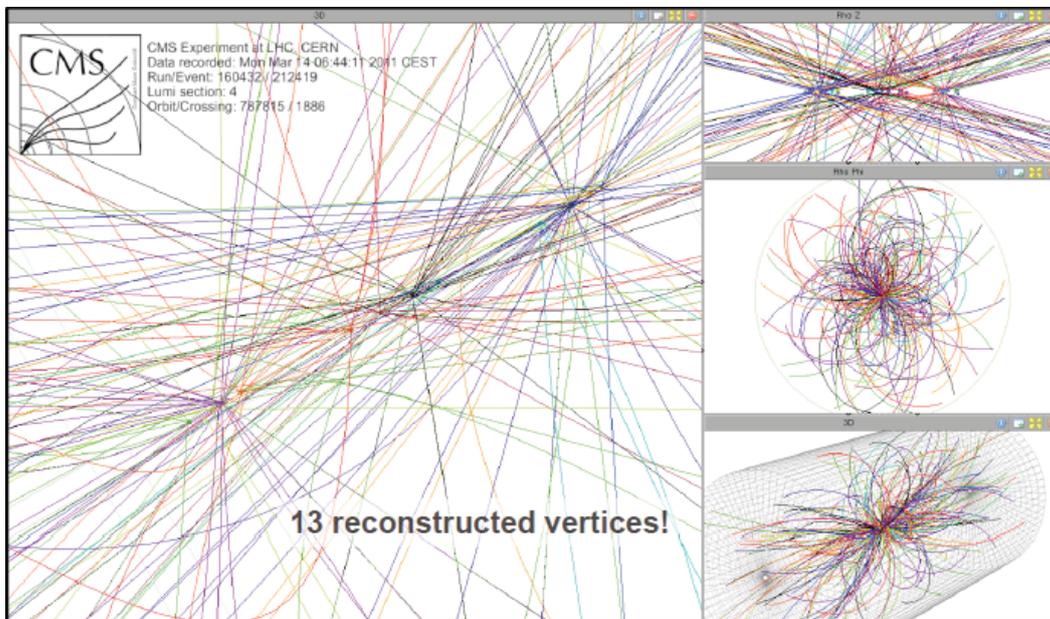


Fig. 2.4 – Un exemple d'événement à haute multiplicité dans CMS avec les traces associées à chaque vertex.

On fait la distinction entre l'empilement produit en même temps que la collision principale (PU « in-time ») et celui qui provient des croisements suivants ou précédents (PU out « off-time »). Le premier a des effets non négligeables sur la plupart des analyses physiques car il provoque une augmentation générale de l'activité dans le détecteur. Il peut ainsi donner lieu à une mauvaise identification des particules provenant de la collision dure. Même si l'ensemble des détecteurs choisis dans CMS ont un temps de réponse rapide et un faible temps mort, les croisements de paquets sont si fréquents que des signaux électriques provenant des collisions précédentes peuvent se superposer aux signaux de la collision enregistrée. En modifiant la forme des signaux électriques sortant des cartes d'acquisition des détecteurs, ce PU « off-time » va dégrader les mesures des différentes observables. Plus le temps séparant les paquets est faible plus ce phénomène devient problématique.

### 2.1.5 Quelle physique au LHC ?

L'organisation circulaire du LHC permet de disposer de plusieurs points de croisement de faisceaux. Il est donc possible d'utiliser le faisceau pour plusieurs expériences distinctes. Au LHC, on en retrouve 7 différentes le long de l'anneau : ATLAS, ALICE, CMS, LHCb, LHCf, MoEDAL, et enfin TOTEM.

- ATLAS (A Toroidal LHC ApparatuS) [26] et CMS (Compact Muon Solenoid) [27] sont deux détecteurs généralistes qui permettent d'étudier un large spectre d'observables en physique des particules. Les principaux axes de recherche sont la quête et l'étude du boson de Higgs, les mesures de précision des paramètres du modèle standard, la recherche de nouvelle physique ainsi que la mise en évidence de candidats matière noire. Ces deux expériences mènent par ailleurs un programme de recherche sur les collisions d'ions lourds. Le 4 Juillet 2012, ces deux expériences ont d'ailleurs conjointement annoncé la découverte d'un nouveau boson dont les propriétés sont très proches de celle du boson de Higgs du MS.
- ALICE (A Large Ion Collider Experiment) [25] est principalement dédiée à l'étude des collisions d'ions lourds. Elle permet notamment d'étudier l'état déconfiné de la matière nucléaire : le plasma de quark et gluon. Des mesures sur les collisions pp et p-Pb sont également étudiées afin de comprendre et d'étalonner le détecteur.
- LHCb (Large Hadron Collider beauty)[28, 32] permet d'étudier plus en détail l'asymétrie matière anti-matière grâce aux propriétés du quark b et de la violation de symétrie CP (Charge-Parité). La collaboration LHCb a d'ailleurs récemment publié des résultats d'une grande importance pour la cohérence du modèle standard sur la désintégration des mésons B [33, 34]. L'observation de ces désintégrations extrêmement rares pour la première fois a confirmé les prédictions du modèle standard et a permis de contraindre certains modèles d'extension.

Collisions	$E_{CM}$ [TeV]	Année	Luminosité intégrée totale délivrée							
			ALICE		ATLAS		CMS		LHCb	
p-p	7,00	2010	×		48,1	$pb^{-1}$	44,2	$pb^{-1}$	×	
	7,00	2011	4,84	$pb^{-1}$	5,3	$fb^{-1}$	5,3	$fb^{-1}$	1,2	$fb^{-1}$
	8,00	2012	9,96	$pb^{-1}$	23,3	$fb^{-1}$	23,2	$fb^{-1}$	2,2	$fb^{-1}$
Pb-Pb	2,76	2010	10,0	$\mu b^{-1}$	9,69	$\mu b^{-1}$	10,00	$\mu b^{-1}$	0,0	$\mu b^{-1}$
	2,76	2011	144	$\mu b^{-1}$	168	$\mu b^{-1}$	148,7	$\mu b^{-1}$	0,0	$\mu b^{-1}$

TABLE 2.2 – Luminosité délivrée par le LHC aux 4 expériences principales pendant les 3 premières années de fonctionnement en collisions p-p et Pb-Pb.

- LHCf (Large Hadron Collider Forward)[35] étudie les particules créées à très petits angles pour simuler les rayons cosmiques en laboratoire. Ces résultats sont très utiles pour la calibration des détecteurs des expériences géantes chargées de l'étude de ces rayons en haute atmosphère. LHCf est située à 140 m de chaque côté du point de collision d'ATLAS.
- TOTEM (TOTal Elastic and diffractive cross section Measurement)[36, 37] étudie la physique à très petits angles de déviation qui est inaccessible aux grandes expériences généralistes. Cette expérience est installée de part et d'autres du point d'interaction de l'expérience CMS<sup>3</sup> (Point 5). Elle permet entre autre de mesurer les sections efficaces des collisions élastiques et de mesurer plus précisément la luminosité délivrée par le LHC.
- MoEDAL (Monopole and Exotics Detector at the LHC) [38] a pour mission de rechercher directement le monopôle magnétique ainsi que de nouvelles particules stables et fortement ionisantes prédites par des théories au-delà du modèle standard. Installé autour du même point d'interaction que LHCb, MoEDAL est semblable à un appareil photo géant qui traque les signes révélateurs de nouvelle physique grâce à des détecteurs de traces nucléaires.

Les expériences ALICE et LHCb ont été conçues pour enregistrer des luminosités instantanées plus faibles que CMS et ATLAS. Ainsi, pour leur permettre d'acquérir des données de collisions lors des schémas de remplissage à haute luminosité, on joue sur l'alignement des faisceaux afin d'obtenir localement une luminosité moindre. Toutes les expériences du LHC ne disposent donc pas de la même luminosité au final, le tableau récapitulatif 2.2 résume ces différences pour chaque année de fonctionnement.

3. Sur des distances allant de 7,5m jusqu'à 220m.

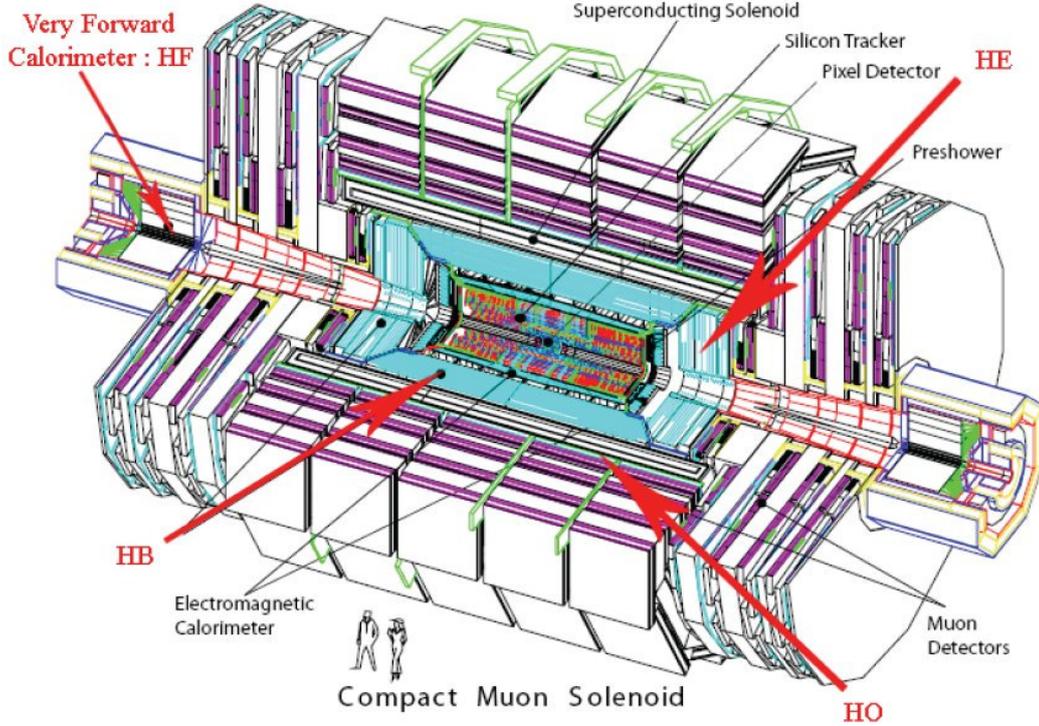


Fig. 2.5 – Représentation de l'expérience CMS.

## 2.2 L'expérience Compact Muon Solenoid (CMS)

CMS est un détecteur généraliste optimisé pour l'étude de la physique à l'échelle du TeV. La figure 2.5 montre le détecteur CMS dans son ensemble avec sa longueur de 28,7 m, son rayon de 7,5 m et son poids total de 14000 t. CMS utilise un système de coordonnées dont l'origine se situe au centre du détecteur au point d'interaction P5 comme indiqué en figure 2.6. L'axe Z est tangent à la direction des faisceaux. L'axe X pointe vers le centre de l'anneau du LHC. L'axe Y, perpendiculaire aux deux autres, pointe vers le haut. L'angle azimutal  $\phi \in [-\pi, \pi]$  est mesuré à partir de l'axe X dans le plan transverse. L'angle  $\theta$  est pour sa part défini à partir de l'axe Z. En physique des particules, on lui préfère souvent la pseudorapidité  $\eta \in [-\infty, \infty]$  définie par l'équation (2.3).

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] = \frac{1}{2} \ln\left(\frac{|\vec{p}| + p_L}{|\vec{p}| - p_L}\right) \quad (2.3)$$

La rapidité d'une particule dont le quadrivecteur énergie impulsion est  $p^\mu = (E, p_x, p_y, p_z)$  est rappelé dans l'équation (2.4), où  $p_L$  est la projection du moment sur l'axe du boost selon l'axe Z.

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

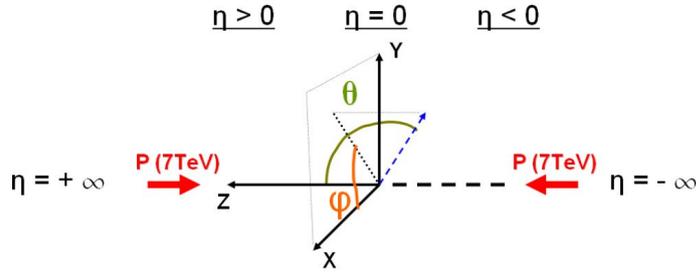


Fig. 2.6 – Le système de coordonnées de l'expérience CMS.

A la limite ultra-relativiste ( $|\vec{p}| \gg m$ ), la pseudo-rapacité tend vers la rapidité qui est invariant de Lorentz. Ainsi pour les particules ultra-relativistes (ce qui est tout le temps le cas au LHC)  $\eta$  est un invariant de Lorentz.

On définit l'impulsion d'une particule dans l'équation (2.5) avec  $m$  la masse de la particule et  $v$  sa vitesse.

$$p = \gamma m v \text{ avec } \gamma = \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}} \quad (2.5)$$

Le moment transverse  $p_t$  est la projection du moment de la particule dans le plan transverse  $XY$ . Il est défini comme :

$$p_t = \sqrt{p_x^2 + p_y^2} \quad (2.6)$$

L'énergie transverse est définie comme :

$$E_t = E \times \sin \theta = \frac{E}{\cosh \eta} \quad (2.7)$$

Pour une particule sans masse, l'énergie transverse et l'impulsion transverse sont égales. Aux énergies considérées, on considérera que c'est également le cas pour les électrons et les muons. Enfin, la distance  $\Delta R$  entre deux particules dans le plan  $(\eta, \phi)$ , définie dans l'équation (2.8), est aussi un invariant de Lorentz pour les boost selon l'axe  $Z$ .

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} \quad (2.8)$$

L'une des observables qui a dicté la conception du détecteur est la mesure précise du  $p_t$  des muons<sup>4</sup>. Pour remplir cet objectif, il est nécessaire de plonger le détecteur dans un très fort champ magnétique. Ainsi CMS comporte en son sein un solénoïde supraconducteur de 3.8 T détaillé dans la section 2.2.1. L'intense flux magnétique généré par le solénoïde est contenu par la culasse en fer, dite de retour de champs.

La précision sur la mesure du  $p_t$  des muons est aussi limitée par la mesure des traces dans le centre du détecteur. Elle est réalisée par le détecteur à pixels ainsi que par le trajectographe à piste de silicium décrits en section 2.2.2.

4. De l'ordre du GeV.

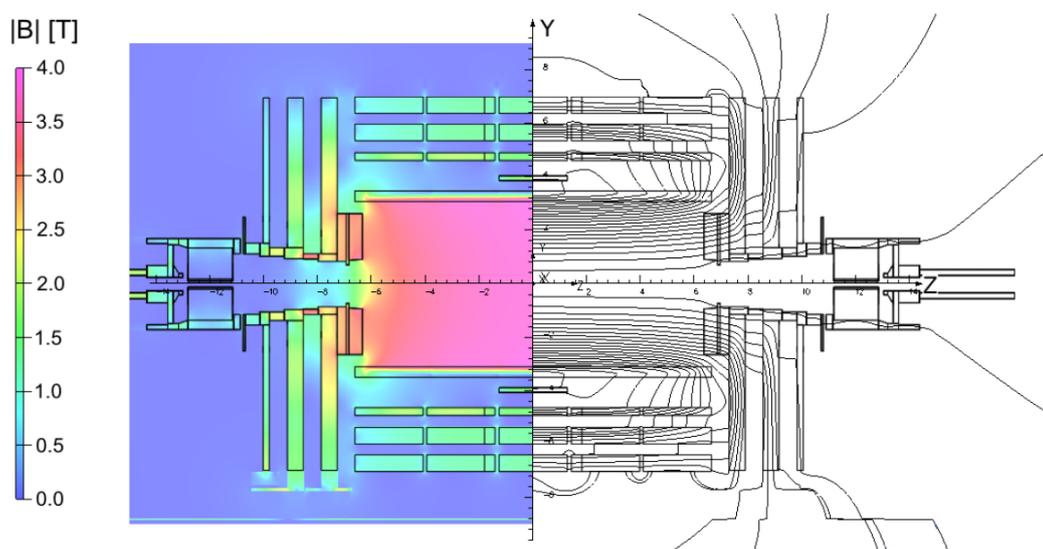


Fig. 2.7 – Vue en coupe du détecteur CMS avec l'intensité et les lignes de champ du champ magnétique.

Pour la recherche de nouvelle physique notamment, on a besoin de connaître avec une grande précision l'énergie transverse manquante ( $E_T$ ). On recherche donc la meilleure herméticité possible du détecteur en réduisant au maximum la couverture angulaire des zones non appareillées. Le calorimètre équipant le tonneau et les bouchons permettent de couvrir jusqu'à  $|\eta| < 3,0$ . Il est constitué du calorimètre électromagnétique (ECAL) et du calorimètre hadronique (HCAL) détaillés respectivement dans les sections 2.2.3 et 2.2.4.

Une description plus complète du détecteur CMS est donnée dans l'article de référence [27] et dans les rapports techniques de conception [39, 40]. Le principe de fonctionnement des différents sous-détecteurs est expliqué en détail dans le document [41].

### 2.2.1 Le solénoïde

Le solénoïde supraconducteur de CMS [42] a été conçu pour fournir un champ axial uniforme de 4 T avec un courant nominal de 19,14 kA dans un cylindre de 6 m de diamètre et de 12,5 m de long. Cela correspond à une énergie stockée à pleine puissance de 2,6 GJ. Le flux magnétique est contenu par une culasse en fer de 10000 t constituée de 5 roues et 2 bouchons. Cette culasse représente, à elle seule, les trois quarts de la masse totale de CMS. Le solénoïde est composé de 4 couches bobinées d'un câble supraconducteur en NbTi renforcé mécaniquement par un alliage d'aluminium. Il est contenu dans un cryostat pesant 220 t qui permet de le maintenir à sa température de fonctionnement de 1,9 K grâce à la circulation de l'hélium superfluide.

La collaboration CMS a décidé de faire fonctionner le solénoïde à 3,8 T au moins pour les premières années de mise en route. Le solénoïde de CMS est suffisamment

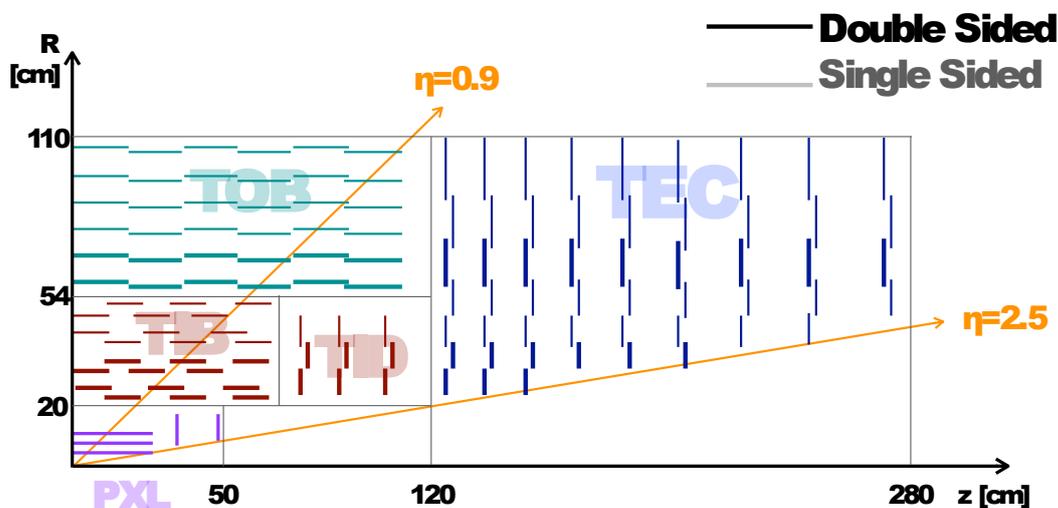


Fig. 2.8 – Coupe longitudinale du trajectographe de l'expérience CMS.

grand pour accueillir en son sein le trajectographe et les deux calorimètres. C'est de là que l'expérience tire son qualificatif de compact. Il est à noter que la technologie retenue par ATLAS est complètement différente puisqu'elle est basée sur un aimant toroïdal ce qui a comme conséquence parmi d'autres une taille 2 fois plus importante. On peut voir sur la figure 2.7 l'intensité et les lignes de flux du champ magnétique autour du solénoïde en fonctionnement. Grâce à sa taille et sa robustesse, le solénoïde supporte une grande partie de la structure de CMS.

### 2.2.2 Le trajectographe

Le trajectographe de CMS représenté en figure 2.8 est un composant cylindrique de 5,5 m de long pour un rayon de 1,1 m. Il permet de reconstruire les trajectoires des particules chargées (traces) à partir des points d'impact (hits) mesurés dans les différentes couches du détecteur. Ces hits correspondent au signal électrique renvoyé quand une particule est détectée par une composante du détecteur. Le trajectographe a par ailleurs pour mission d'identifier le vertex d'interaction principal ainsi que les vertex secondaires. Il couvre la région angulaire  $0 < |\eta| < 2,5$ . Il est composé de 2 détecteurs principaux. Le détecteur silicium à pixels au plus proche du point d'interaction et le détecteur à micropistes de silicium qui l'entoure. Le détecteur silicium à pixels est composé de 3 couches de détection dans le tonneau (situées en moyenne à  $r=4,4$  cm,  $r=7,3$  cm et  $r=10,2$  cm) et quatre disques de détection disposés de chaque côté (à  $z = \pm 34,5$  cm et  $z = \pm 46,5$  cm). Il comporte 66 millions de pixels couvrant chacun une surface de  $100 \times 150 \mu\text{m}^2$ . L'ensemble du détecteur à pixels est organisé en 1440 modules et couvre une surface d'environ  $1\text{m}^2$ . Ce sous-détecteur permet d'obtenir 3 points de haute précision sur les trajectoires de chaque particule chargée. Sa résolution spatiale est de  $10 \mu\text{m}$  dans le plan  $r-\phi$  et de  $20 \mu\text{m}$  dans la direction de l'axe Z.

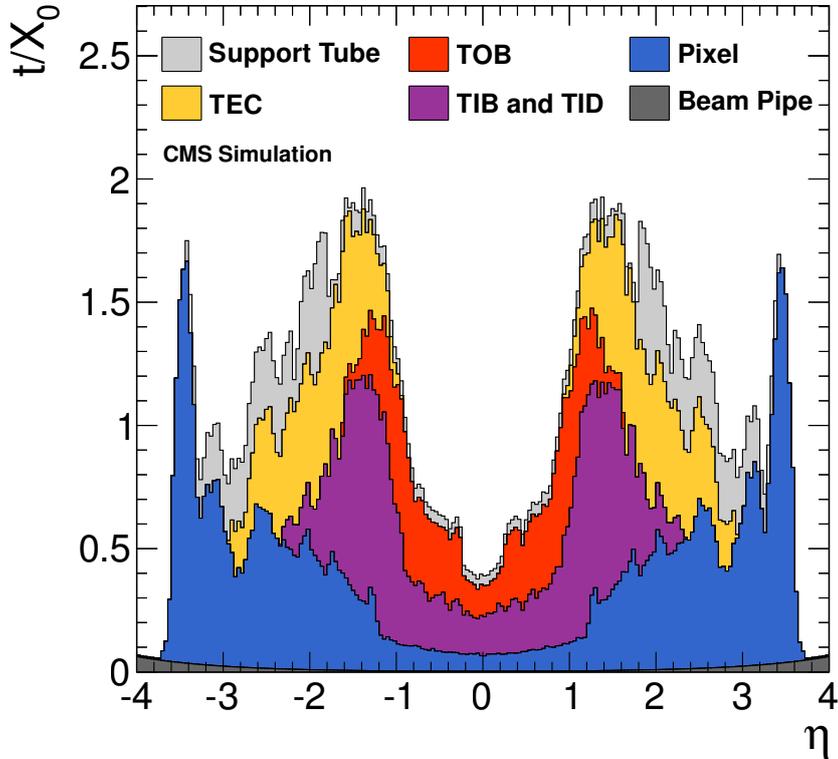


Fig. 2.9 – Quantité de matière dans le trajectographe, exprimée en longueurs de radiation en fonction de  $\eta$  [43].

Le détecteur à micropistes est, quant à lui, constitué des sous éléments suivants :

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

Le TIB comprend 4 couches situées à 25,5, 33,9, 41,8 et 49,8 cm de l'axe du faisceau. Les deux premières présentent deux faces et les autres une seule. Il est complété par le TID dans les parties bouchon avec 3 disques de part et d'autres (à  $|z|=70, 80$  et  $90$  cm). Le TOB, plus long que le TIB, se constitue de 6 couches distinctes sises à 60,8, 69,2, 78,8, 86,8, 96,5 et 108 cm de l'axe des faisceaux. Finalement, le TEC vient fermer latéralement le trajectographe avec des disques disposés entre  $z = \pm 124$  cmet  $z = \pm 280$  cm. Chaque TEC est formé de 9 disques de rayon compris entre 22 et 113,5 cm.

Le trajectographe est par ailleurs constitué de matériaux non actifs qui perturbent les mesures à cause de leur interaction avec les particules détectées. Il s'agit des structures mécaniques qui assurent la solidité du détecteur, des circuits de

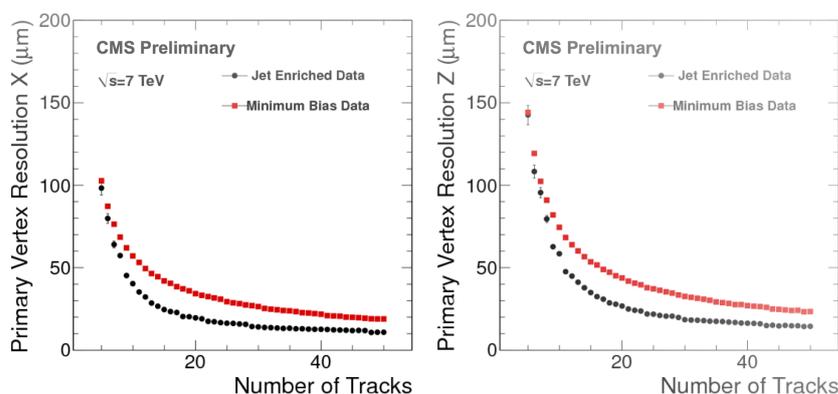


Fig. 2.10 – Résolution du vertex primaire dans le plan transverse (à gauche) et le long de l’axe des faisceaux (à droite) en fonction du nombre de traces rattachées au vertex [44].

refroidissement et de l’ensemble de l’électronique nécessaire au bon fonctionnement des capteurs. L’épaisseur de ces matériaux est illustrée dans la figure 2.9 en fonction de  $\eta$  pour les différents sous parties du trajectographe. Ces épaisseurs sont exprimées en termes de longueur de radiations,  $X_0$  étant la longueur moyenne au bout de laquelle l’énergie d’un électron est réduite d’un facteur  $1/e$  par bremsstrahlung.

La résolution sur les vertex d’interaction primaires dans le trajectographe est présentée sur la figure 2.10. Elle a été estimée par la méthode dite de séparation où l’ensemble des traces est divisé en deux groupes. La résolution est alors proportionnelle à la distance entre les deux vertex déterminés avec les deux sous ensemble de traces [44]

### 2.2.3 Le calorimètre électromagnétique

Le calorimètre électromagnétique (ECAL pour Electromagnetic CALorimeter) a été conçu pour mesurer précisément l’énergie des photons et des électrons (et donc indirectement, celle du boson de Higgs). Il a été pensé pour l’étude de la désintégration  $H \rightarrow \gamma\gamma$  et c’est le seul sous-détecteur à fournir une information sur les photons. C’est un scintillateur hermétique et homogène constitué de cristaux de tungstate de plomb ( $PbWO_4$ ). Il est séparé en une partie tonneau EB ( $|\eta| < 1,479$ ) et une partie bouchon EE ( $1,479 < |\eta| < 3$ ). Des photo-détecteurs sont fixés à l’extrémité de chaque cristal pour collecter et amplifier la lumière produite par scintillation par la gerbe électromagnétique.

Les pions sont des particules à faible durée de vie se désintégrant selon  $\pi^0 \rightarrow \gamma\gamma$ . S’ils sont produits dans la région des bouchons, l’angle entre les deux photons est si petit que la granularité du ECAL n’est pas suffisante pour les différencier. Un détecteur à pied de gerbe (PS pour preshower) a donc été ajouté devant les bouchons du ECAL pour pouvoir distinguer ces pions à grand  $|\eta|$  des photons uniques.

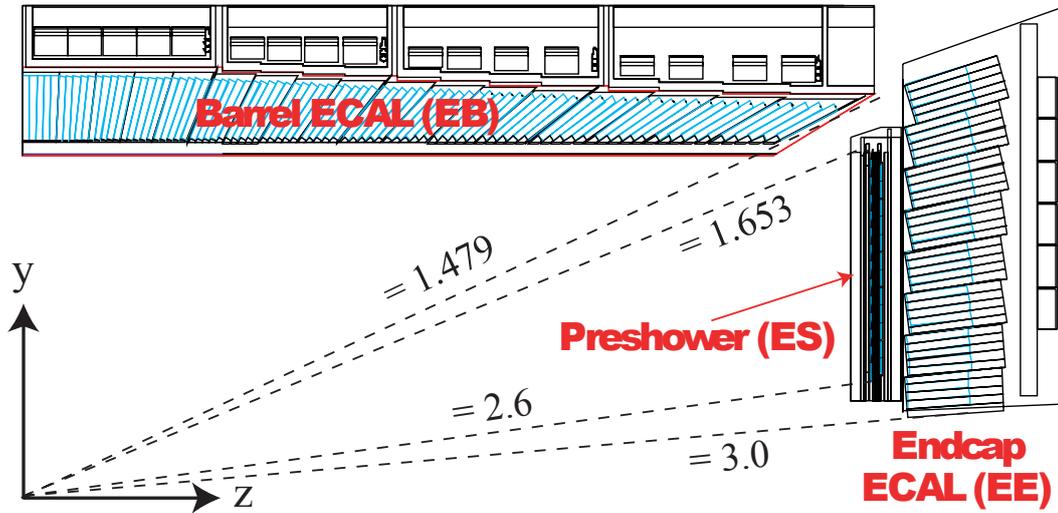


Fig. 2.11 – Vue longitudinale des parties tonneau, bouchon ainsi que du détecteur de pied de gerbes constituant le calorimètre électromagnétique de CMS.

**Les cristaux du ECAL** Pour comprendre le choix du tungstate de plomb, il faut faire l'inventaire des contraintes relatives au ECAL. Le matériau doit d'abord être compact pour minimiser la taille du calorimètre (il doit tenir à l'intérieur du solénoïde avec le HCAL). Grâce à sa haute densité ( $8,28 \text{ SI kg/m}^3$ ) et sa faible longueur de radiation ( $0,89 \text{ cm}$ ), le  $\text{PbWO}_4$  permet d'absorber l'énergie des photons et des électrons dans une longueur raisonnable ( $25,8X_0$  dans le tonneau et  $24,7X_0$  dans les bouchons). Il faut, par ailleurs, que le calorimètre électromagnétique ait une granularité suffisante. Ceci est rendu possible par le faible rayon de Molière du tungstate de plomb ( $2,2 \text{ cm}$ ) qui caractérise le développement latéral de la gerbe. Ainsi l'utilisation de cristaux de surface  $2,2\text{cm} \times 2,2\text{cm}$  permet une bonne séparation des gerbes électromagnétiques. Enfin le temps de réponse de ces cristaux est en adéquation avec les conditions de fonctionnement nominales du LHC. On estime que  $80 \%$  de la lumière émise par les cristaux est captée en  $25 \text{ ns}$ <sup>5</sup>. Dans le EB, les cristaux sont regroupés en sous-modules  $2 \times 5$  (en  $\phi \times \eta$ ). Ces sous-modules sont ensuite eux même regroupés en 36 super-modules qui couvrent chacun un  $\Delta\phi$  de  $20^\circ$ . Dans le EE, les matrices  $5 \times 5$  de cristaux sont appelés super-cristaux. Chaque bouchon est ainsi divisé en deux parties contenant 138 super-cristaux chacun. Les cristaux sont montés de façon projective (voir figure 2.11) pour éviter des zones non actives alignées avec la trajectoire d'une particule.

**Les photo détecteurs** Ils doivent être rapides, résistants aux radiations et fonctionnels dans un champ magnétique de  $3,8 \text{ T}$ . Pour s'adapter au différentes conditions de radiation, deux modèles différents sont utilisés dans le tonneau et dans les bouchons. Des photodiodes à avalanches (APD) sont montées sur chaque cristal dans le tonneau. Alors que des phototriodes à vide (VPT) équipent ceux des

5.  $25 \text{ ns}$  est la séparation nominale entre 2 croisements (50 en 2012)

bouchons.

**Le détecteur de pied gerbe (PS)** C'est un détecteur à échantillonnage de 20 cm d'épaisseur placé de part et d'autre du trajectographe juste devant les bouchons du ECAL. En terme de pseudorapidité, il couvre la région  $1.653 < |\eta| < 2.6$  (voir figure 2.11). Sa partie absorbeur est constituée de plomb et permet d'initier la gerbe électromagnétique créée par les particules incidentes. Des capteurs de silicium sont intercalés derrière l'absorbeur pour mesurer l'énergie déposée et l'étalement transverse des gerbes. La granularité de ces capteurs de silicium est suffisante pour pouvoir distinguer deux photons provenant de la désintégration d'un  $\pi_0$  très à l'avant. Le PS présente une granularité plus fine que le ECAL avec des pistes de détecteurs de 2 mm de large contre 3 mm pour le ECAL. On estime que seulement 5% de l'énergie d'un photon ou d'un électron est déposée dans le PS.

### Performances

Le mécanisme de scintillation n'est pas affecté par l'augmentation des radiations. Cependant, les cristaux se teintent au fur et à mesure de l'irradiation ce qui réduit leur transparence. Afin de prendre en compte cet effet dans nos mesures, un système de surveillance par laser a été mis en place. On injecte une lumière laser avec une fréquence proche du pic de scintillations du  $PbWO_4$ . La figure 2.12 (à gauche) montre la réponse moyennée au laser pour différentes régions en  $\eta$  des cristaux au cours du temps. Les changements dans la réponse au cours de 2012 sont de quelques pour-cents pour la partie la plus centrale mais atteignent les 30 % pour  $|\eta| \simeq 2.5$  (la partie la plus proche des faisceaux). Ce système de surveillance fournit les coefficients correctifs permettant d'ajuster la réponse du ECAL au cours du temps. Il est intéressant d'observer que durant les périodes sans faisceau, les cristaux récupèrent partiellement leur transparence ce qui permet d'espérer de bonnes conditions de fonctionnement du ECAL pour la reprise du LHC en 2015.

La figure de droite 2.12 montre que la résolution en énergie est affectée par la quantité de matière devant le ECAL. On note aussi une dégradation des performances à chaque interface entre les modules du ECAL (lignes verticales sur la figure à droite). Les calibrations les plus récentes qui utilisent l'ensemble des données 2012 permettent d'améliorer significativement cette résolution.

#### 2.2.4 Le calorimètre hadronique

Le calorimètre hadronique (HCAL pour Hadronic CALorimeter) permet de mesurer l'énergie des hadrons neutres et chargés. Sa partie tonneau est constituée de plaques de laiton de 5 cm d'épaisseur. Le laiton a été retenu car il n'est pas sensible au champ magnétique généré par le solénoïde (il est paramagnétique). L'utilisation du laiton engendre ainsi moins de déformation du détecteur sous l'influence du champ magnétique. De plus, il perturbe très peu la trajectoire des muons qui le traversent (la mesure de leur impulsion n'est donc pas biaisée). Des feuilles de scintillateur de  $4 \cdot 10^{-3}m$  sont intercalées entre les couches de laiton et leur lumière

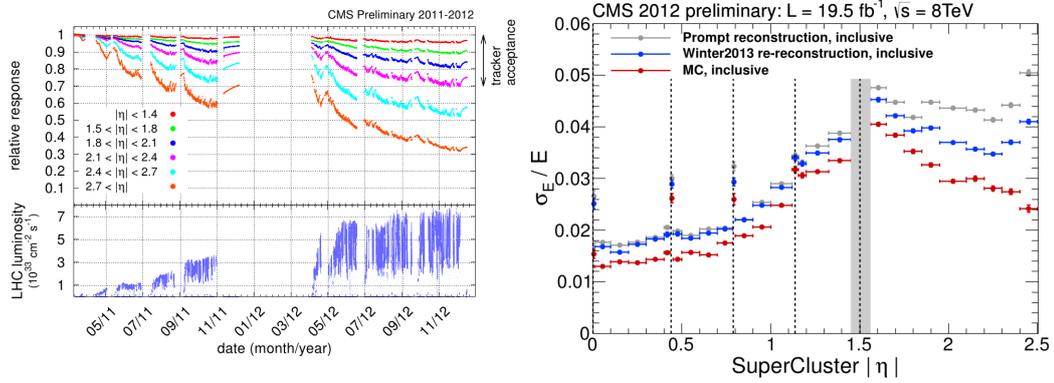


Fig. 2.12 – à gauche : Réponse relative du ECAL aux rayons laser (447nm en 2012) moyenné sur tous les cristaux pour les périodes de prises de données 2011 et 2012 à droite : Résolution en énergie relative du ECAL par super clusters en  $\eta$  obtenue pour des événements  $Z \rightarrow ee$  [45].

est récoltée par fibre optique. La technologie retenue pour les photo-détecteurs est ici celle des photodiodes hybrides. La partie tonneau se sépare entre le HB (Hadron Barrel calorimeter) situé à l'intérieur du solénoïde et le HO (Hadron Outer calorimeter) à l'extérieur. En effet la partie interne (HB) n'est pas suffisamment profonde pour absorber la totalité de la gerbe hadronique. La bobine sert ensuite d'absorbeur avant que le HO ne puisse mesurer les gerbes tardives ou longues. Dans les bouchons HE, l'organisation reste inchangée mais l'épaisseur des plaques de laiton est cette fois de 8 cm. Ils couvrent la zone angulaire :  $1,48 < |\eta| < 3,0$ . Et leurs épaisseurs de 1,8 m leur permettent de contenir l'ensemble de la gerbe hadronique. Pour augmenter l'acceptance jusqu'à  $|\eta| < 5$ , deux calorimètres « à l'avant » (HF pour Hadron Forward calorimeter) ont été installés de part et d'autre du point d'interaction à une distance  $z = \pm 11,1$ m. Ils ont un rayon de 1,4 m pour une longueur de 1,65 m.

La résolution relative en énergie du calorimètre hadronique peut être paramétrée comme suit :

$$\frac{\sigma_E}{E} = \frac{s}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (2.9)$$

où  $s$  est le terme stochastique,  $b$  le terme de bruit et  $c$  le terme constant.  $s$  prend en compte les fluctuations statistiques lors des dépôts d'énergie et de la photo multiplication dans les PD.  $b$  tient compte du bruit gaussien de la chaîne d'acquisition électronique et du bruit dû à l'empilement. Enfin,  $c$  caractérise la qualité du cristal et de la calibration. La résolution a été testée sur faisceau de pions avant le démarrage du LHC [46, 47] et a permis d'établir les formules suivantes :

$$\text{Dans le tonneau : } \frac{\sigma_E}{E} = \frac{1,15}{\sqrt{E}} \oplus 0,055 \quad (2.10)$$

$$\text{Dans les bouchons : } \frac{\sigma_E}{E} = \frac{1,19}{\sqrt{E}} \oplus \frac{0,38}{E} \oplus 0,04 \quad (2.11)$$

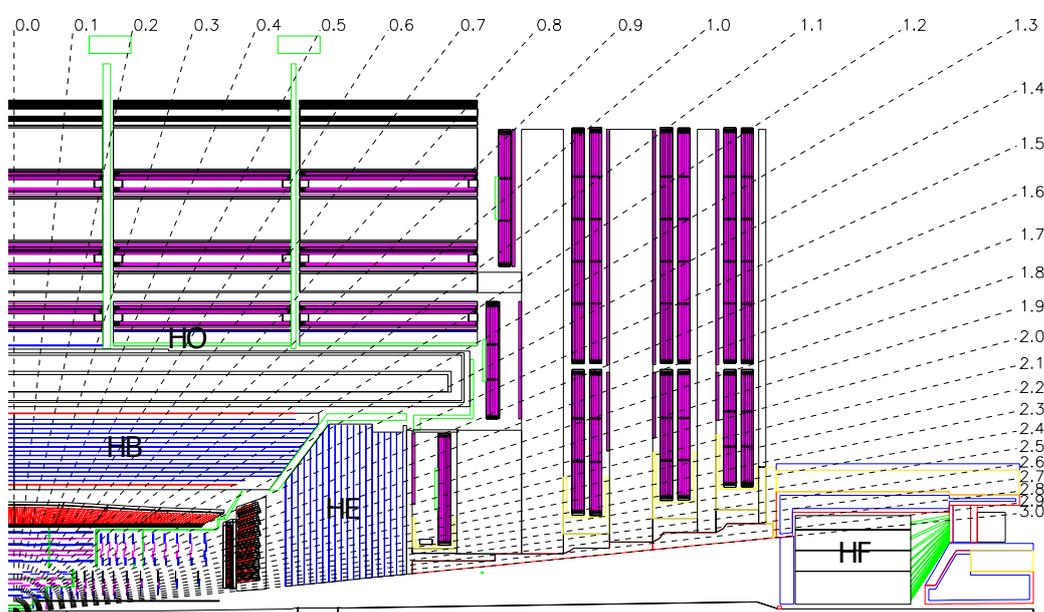


Fig. 2.13 – Vue longitudinale du calorimètre Hadronique dans son ensemble, HB, HE, HO et HF.

Après les premières collisions à 7 TeV en 2010 la réponse du HCAL dans ses différents modules a été évaluée [48]. Les résultats de cette étude sont résumés sur la figure 2.14.

### 2.2.5 Les chambres à muons

Les chambres à muons du détecteur CMS sont présentées sur la coupe longitudinale 2.15 et sur la coupe transverse 2.16. Elles ont 3 fonctions principales : identifier les muons, mesurer précisément leurs impulsions et enfin, assurer une partie du déclenchement (voir partie 2.3). Le système de détection des muons est constitué de 2 types différents de détecteurs de particules déployés dans le tonneau et les bouchons en dehors du solénoïde. Il s'agit de deux détecteurs à gaz, les tubes à dérive (DT pour Drift Tube) et les chambres à pistes cathodiques (CSC pour Cathode Strip Chamber) et d'un détecteur à plaque résistives (RPC). Le spectromètre à muons constitue la partie la plus externe du détecteur CMS, les muons étant les seules particules à pouvoir atteindre cette profondeur. En effet, ils sont beaucoup moins sensibles au rayonnement de freinage que les autres particules chargées et ils n'interagissent pas fortement. L'éloignement des points de mesure entre le trajectographe et les chambres à muons permet en outre d'obtenir une meilleure précision sur la mesure de l'impulsion des muons.

#### Les tubes à dérive

Dans la partie tonneau, le champ magnétique est uniforme et le taux de muons est assez faible pour permettre l'utilisation des tubes à dérive (figure 2.18). Les

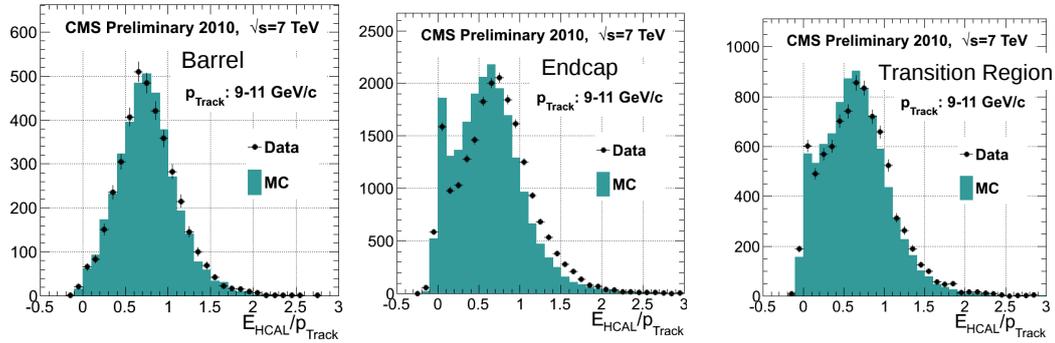


Fig. 2.14 – Réponse du HCAL pour des traces de 9 à 11 GeV pour 3 régions en  $\eta$ . A gauche  $|\eta| < 1,1$  le tonneau ; au milieu  $1,7 < |\eta| < 2,2$  les bouchons ; A droite  $1,1 < |\eta| < 1,7$  la région de transition.

chambres à dérive ont été conçues comme des améliorations des chambres à fils proportionnelles. En effet, en utilisant le temps de dérive des électrons vers l'anode on améliore la résolution spatiale ce qui permet d'augmenter l'espace entre les fils et de diminuer les besoins en électronique. On a ainsi pu équiper un gros volume avec cette technologie tout en gardant un coût raisonnable. Les tubes à dérives sont constitués d'enceintes en aluminium, renfermant un mélange gazeux Argon-Dioxyde de carbone (85%/15%). En leur centre, on retrouve un fil d'acier inoxydable plaqué or de 50  $\mu\text{m}$  de diamètre : l'anode. Lorsque un muon traverse l'enceinte, il ionise le mélange gazeux et les charges électriques ainsi produites sont collectées par l'électrode au centre des tubes après avoir dérivé avec une durée maximale de 400 ns. Comme on peut le voir sur la figure 2.18, les tubes à dérive sont disposés dans le tonneau selon 4 couches concentriques de stations. La figure 2.16 permet de comprendre l'arrangement de ces stations segmentées en 12 secteurs qui couvrent chacun un  $\Delta\phi$  d'environ  $30^\circ$ . On peut également voir qu'elles sont intercalées entre les couches de fer de la culasse de retour de champ.

### Les chambres à pistes cathodiques (CSC)

Dans la région du bouchon par contre, le champ magnétique est puissant et fortement inhomogène et le taux de muons est très important ce qui empêche l'usage des tubes à dérives. La solution technique choisie, les CSC, sont des chambres à fil proportionnelles. Elles sont de forme trapézoïdale et couvrent des angles  $\Delta\phi = 10^\circ$  ou  $\Delta\phi = 20^\circ$ . Les 468 CSC mises bout à bout permettent de couvrir toute la surface des bouchons. Le fil qui déclenche donne la position en  $r$  et on obtient la mesure en  $\phi$  grâce à l'interpolation de la quantité de charge induite sur la cathode.

**Les chambres à plaques résistives** Pour compléter le système de détection dédié aux muons, on a besoin de détecteurs avec une bonne réponse temporelle pour fournir des informations au système de déclenchement (voir partie 2.3). En effet, on a besoin de savoir très rapidement si on doit enregistrer l'événement dans un lot

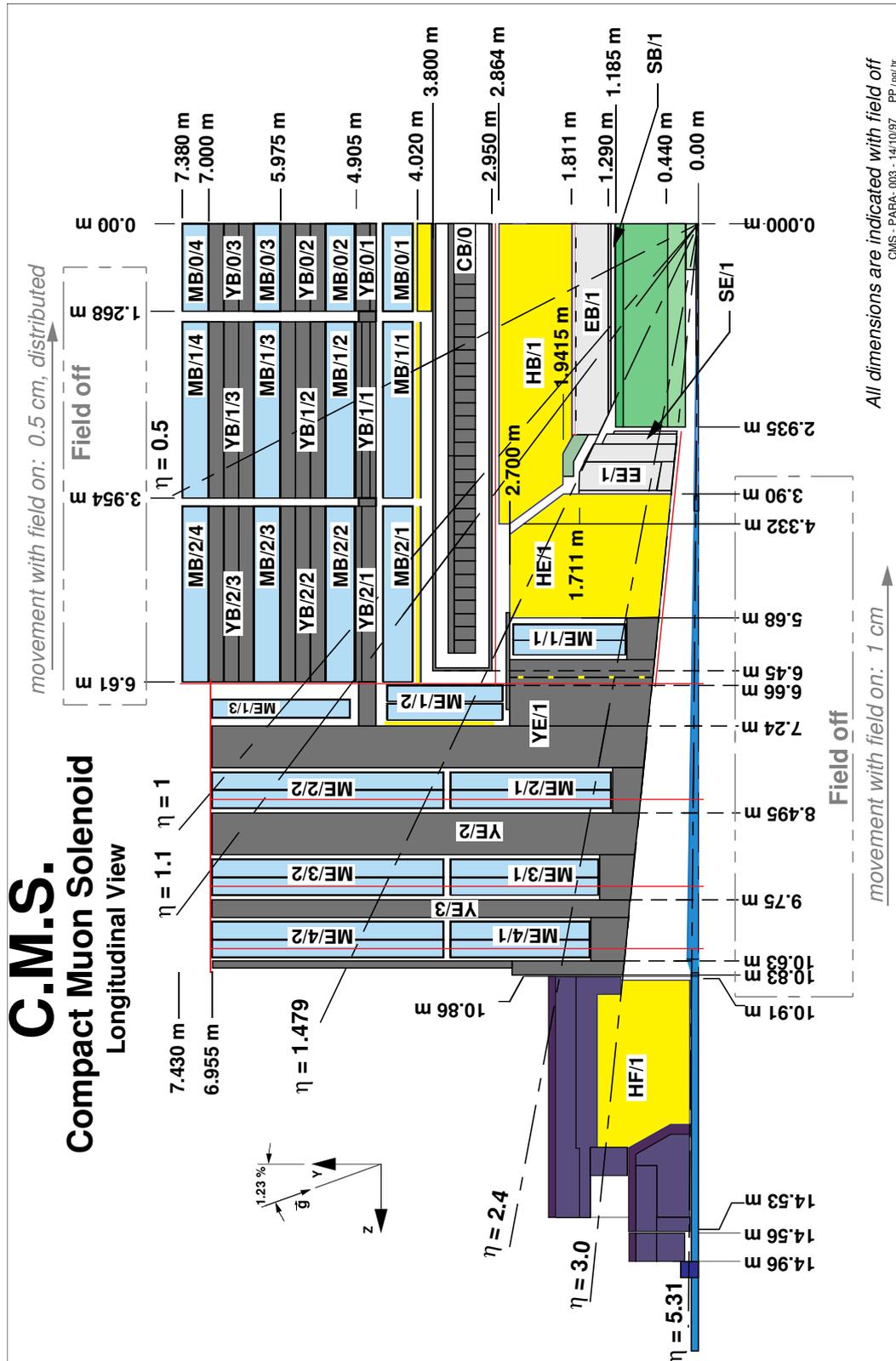


Fig. 2.15 – Coupe longitudinale d'un quart de l'expérience CMS.

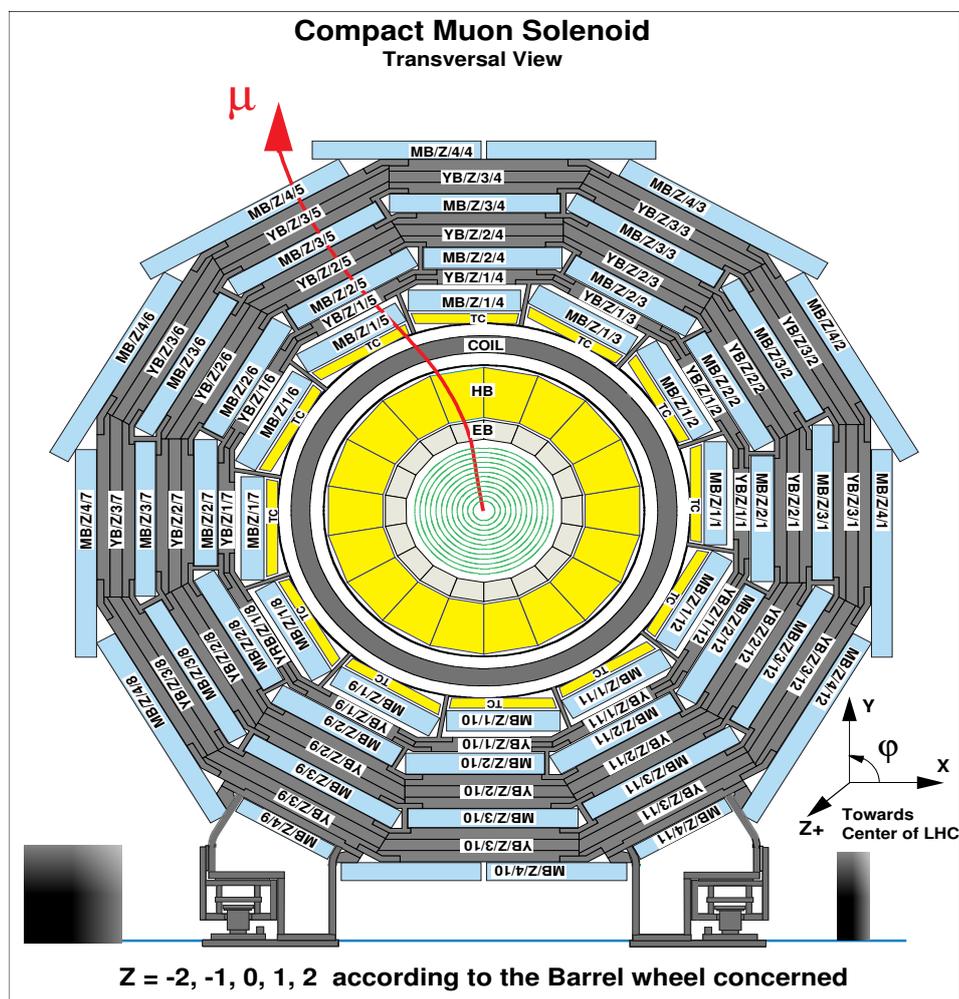


Fig. 2.16 – Coupe transversale de l'expérience CMS.

de données. On a retenu la technologie des chambres à plaques résistives qui sont constituées de deux plaques résistives dont l'intervalle est rempli de gaz. Grâce à leur temps de réponse très inférieur à 25 ns, ils permettent d'identifier avec certitude un événement avec une trace provenant d'un muon même dans des conditions de haute luminosité avec beaucoup d'empilement. Ils sont installés aussi bien dans le tonneau que dans les bouchons pour pouvoir déclencher dans toute les directions.

### Performances

Les dernières études sur le bon fonctionnement des DT portent sur l'ensemble du lot de données 2012. La figure 2.19 présente l'efficacité du déclenchement local des RPC en fonction du  $p_t$  et  $\eta$  du muon. C'est à dire la présence d'un déclenchement primaire pour une station donnée lorsqu'un segment de trace est reconstruit dans cette même station avec au moins 4 signaux.

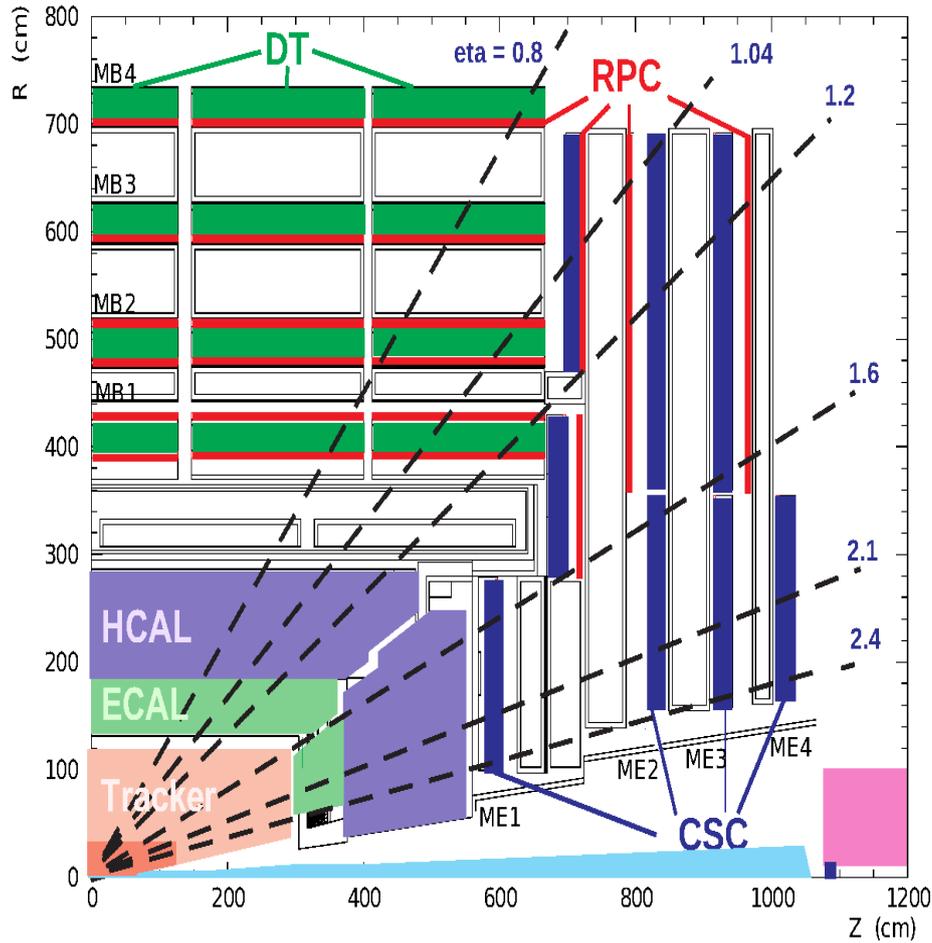


Fig. 2.17 – Vue en coupe des chambres à muons présentant l’agencement des tubes à dérives (DT), des chambres à pistes cathodiques (CSC) et des chambres à plaques résistives (RPC).

### 2.3 Le système de déclenchement de CMS

En fonctionnement nominal, il y a un événement toutes les 25 ns (à chaque croisement de paquets) ce qui représente une fréquence de 40 MHz. On estime que la taille d’un événement sera environ de 1 Mo, à 14 TeV en comptant un empilement moyen de 20 événements. Pour enregistrer l’ensemble des collisions il faudrait donc prévoir une capacité de stockage de 40 To par seconde. Non seulement il n’est pas possible techniquement de stocker une telle quantité de données mais en plus ce n’est pas souhaitable. En effet, comme nous l’avons déjà vu, la plupart de ces événements ne sont pas intéressants du point de vue physique (événements minbias). On a donc besoin d’opérer une sélection drastique des événements, et de ramener le taux d’événements enregistrés à 300 Hz. Cette sélection doit, de plus, être très rapide

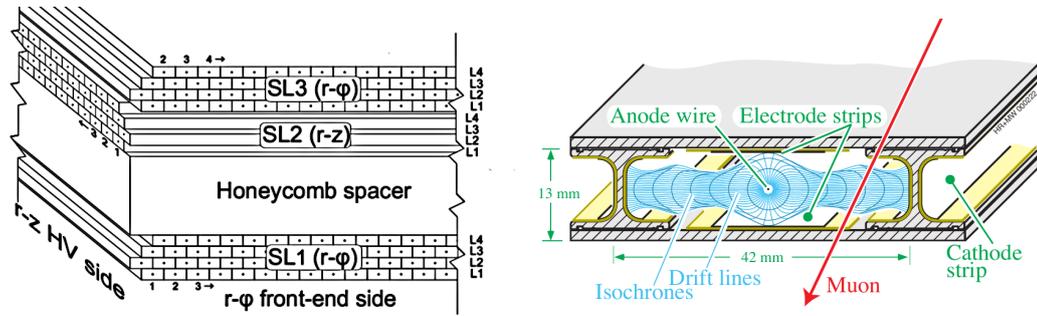


Fig. 2.18 – À gauche : vue schématique en coupe de l’agencement des tubes à dérives. À droite : section d’un tube à dérive avec les lignes de dérives et les isochrones.

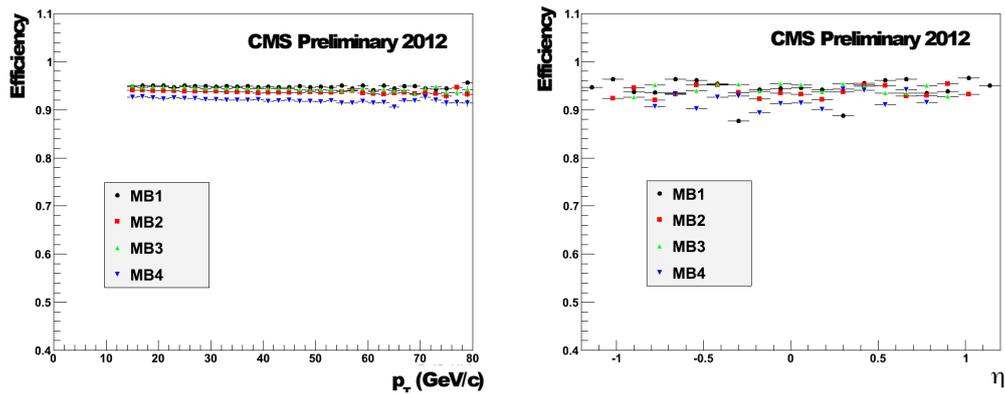


Fig. 2.19 – Efficacité du déclenchement local des RPC en fonction du  $p_t$  (à gauche) et de la pseudorapidité  $\eta$  (à droite) quand un segment de trace est identifié dans les DT.

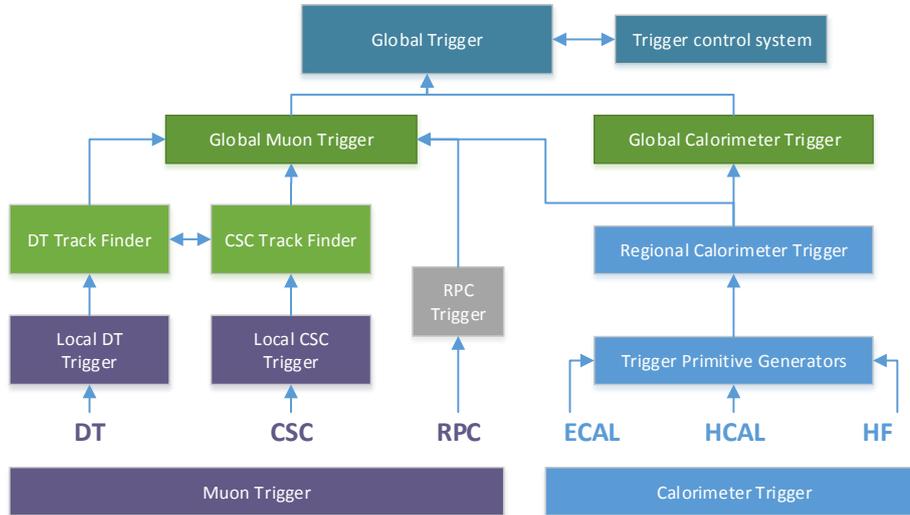


Fig. 2.20 – Architecture du premier niveau de déclenchement pour CMS.

car on doit garder en mémoire toutes les caractéristiques d'un événement au moins jusqu'à ce qu'à la décision de l'écarter ou de le garder soit prise.

Pour dépasser ces contraintes, on a créé un système de déclenchement à deux niveaux. Le niveau 1 (L1 pour Level one) est un système fait d'électronique programmable qui permet de réduire le taux à 100 kHz. Ce premier déclenchement est effectué en  $3,2 \mu\text{s}$  ce qui ne permet pas d'utiliser toutes les quantités reconstruites. Il se base principalement sur les informations du calorimètre et des RPC. Pendant ces  $3,2 \mu\text{s}$  toute l'information se rattachant à l'événement est conservée en réserve (« pipeline memory »).

Si l'événement passe le L1, il est alors passé en revue par le système de déclenchement de haut niveau (HLT pour High Level Trigger). Cet algorithme est appliqué par une ferme de calcul située à la surface à la verticale du détecteur pour minimiser le temps de traitement. La durée d'application du HLT varie selon le type d'événement mais elle vaut en moyenne 50 ms. Ce temps est suffisant pour faire appel à des quantités reconstruites combinant l'information de plusieurs sous-détecteurs. Il nécessite donc une très grande bande passante pour accéder rapidement à toutes les caractéristiques de l'événement (environ 1 To/s). Si finalement, l'événement passe le déclenchement de haut niveau, il est stocké définitivement et envoyé vers le **tier-0**.

Détaillons maintenant plus avant les deux niveaux de déclenchement.

### 2.3.1 Le déclenchement de premier niveau L1

Il est constitué de deux chaînes parallèles de déclenchement (cf Fig. 2.20). Une traitant l'information des calorimètres et l'autre celle des chambres à muons. Les

deux chaînes se divisent en 3 étapes : une reconstruction locale, régionale puis globale. Le résultat des deux parties est ensuite regroupé en une décision globale du L1.

1. La composante locale du L1 est donnée par TPG (Trigger Primitive Generator). Il s'agit des signaux électroniques générés par les différents sous-détecteurs concernés (par exemple un dépôt énergétique dans une tour calorimétrique ou un segment de trace dans les RPC) ;
2. La composante régionale du L1 regroupe les informations locales pour créer des objets candidats (électrons, muons...). Il les classe ensuite en fonction de leur énergie, impulsion, isolation, pseudo-rapacité, etc ;
3. Les composantes globales du L1 compilent toute l'information des calorimètres et des chambres à muons et transfèrent les meilleurs candidats au déclenchement global (Global Trigger GT).

Ce GT doit alors prendre la décision de laisser passer ou non l'événement. Il utilise pour cela des critères préétablis regroupés dans le "menu de déclenchement". Il regroupe 128 algorithmes qui doivent fournir les bits de décision en fonction des caractéristiques des candidats L1. Par exemple L1\_Single\_Iso\_Mu6 requiert un muon isolé d'au moins 6 GeV. Ces bits de décision sont ensuite reliés par des « ou » logiques afin d'atteindre la limite de 100 kHz d'événements enregistrés en sortie.

### 2.3.2 Le déclenchement de haut niveau HLT

Il forme lui aussi des candidats, mais en utilisant cette fois plusieurs observables reconstruites par des algorithmes similaires à ceux utilisés lors de la reconstruction offline pour former les échantillons RECO (détaillés dans la partie 2.4).

Le système de déclenchement de haut niveau utilise aussi un menu de déclenchement prédéfini. Il est constitué de 150 algorithmes différents. En les combinant, on peut trier les événements par classe et on les étiquette pour qu'ils soient enregistrés dans l'échantillon (dataset) correspondant. Certains événements peuvent correspondre à plusieurs classes et sont donc stockés plusieurs fois. Cet étiquetage permet cependant d'économiser des ressources de calcul au final. En effet les analystes auront déjà à leur disposition un sous échantillon correspondant à leur état final et ne seront pas obligés de tourner leur code sur l'ensemble des événements enregistrés par CMS depuis sa mise en marche.

On voit bien que le système de déclenchement (et notamment celui de haut niveau) doit être élaboré en étroite collaboration avec les analystes. Il ne faut, par exemple, pas trop relever le seuil en  $p_t$  de certain trigger sous peine de rendre impossible l'étude de certains canaux physiques. Il faut trouver le juste milieu entre volume de données et pertinence physique des événements.

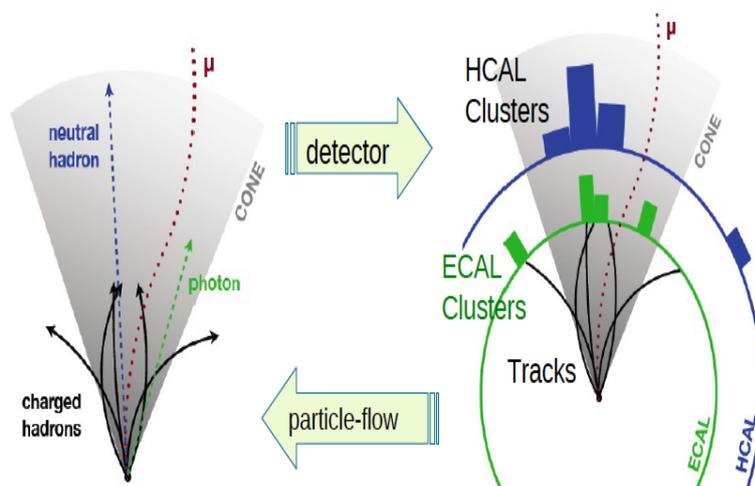


Fig. 2.21 – Principe de fonctionnement de l’algorithme de particle flow. Dans le détecteur, les particules sont vues comme des traces et des dépôts d’énergies.

## 2.4 Reconstruction d’événements dans CMS

Dans cette partie, je m’attacherai à décrire les algorithmes et méthodes utilisés pour reconstruire les candidats particules à partir des données brutes du détecteur. Les objets les plus importants pour mon analyse seront : les leptons, les jets et l’énergie transverse manquante. Comme il a été présenté dans les parties précédentes, le signal numérique provenant de chaque sous-détecteur est enregistré pour les événements qui satisfont les conditions de déclenchement. Examinons maintenant comment ces données brutes sont transformées en candidats particules et en observables physiques qui sont le matériel de base de l’analyse de données. Dans un premier temps, nous présenterons les étapes génériques du particle flow (PF) puis nous verrons les spécificités liées à l’identification de chaque objet.

### 2.4.1 L’algorithme de Particle Flow

Dans ces travaux et dans la majorité des analyses récentes de CMS, l’algorithme de particle flow [49, 50] est utilisé pour reconstruire les candidats particules. Le principe du PF est de combiner les informations de tous les sous-détecteurs pour reconstruire au mieux les particules stables d’un événement (électrons, muons, photons, hadrons chargés et neutres). Le fait d’utiliser tous les sous-détecteurs va permettre une mesure plus précise des caractéristiques (trajectoire, impulsion, énergie) de chaque candidat. La liste des particules établie permettra entre autres de former les jets avec des algorithmes dédiés, d’identifier les leptons  $\tau$  ou de calculer la  $\cancel{E}_T$ ... Le particle flow est ainsi une alternative puissante aux algorithmes de reconstruction classiques.

1. Trajectographie itérative (Iterative tracking)

On utilise comme point de départ les triplets de hits dans le trajectographe ou les doublets avec une condition supplémentaire (distance au vertex principal par exemple) [51]. On extrapole ensuite les trajectoires à la recherche de hits compatibles dans les autres sous-détecteurs en commençant par des conditions strictes. Une fois tous les multiplets passés en revue, on retire les hits qui ont été associés à des traces de façon certaine. On réitère alors la première étape sur la sous collection des hits non appareillés avec des critères moins drastiques pour l'appartenance à une trace. Après 5 itérations successives, on obtient des traces avec au minimum 3 hits, une distance de 50 cm de l'axe des faisceaux et une impulsion de 150 MeV. L'intérêt de réaliser cette collection de traces en relâchant progressivement les critères de qualité est de ne pas dégrader les caractéristiques de meilleures traces à cause des traces les moins précises ;

2. Le regroupement calorimétrique (calorimeter clustering)

Le but de cet algorithme est de mesurer l'énergie et la direction des particules stables non chargées et de les séparer des dépôts énergétiques des hadrons chargés dans les calorimètres. Il permet par ailleurs d'identifier les électrons et les photons rayonnés par bremsstrahlung puis de déterminer l'énergie déposée par les hadrons chargés de haut  $p_t$ . Le regroupement s'effectue pour chaque sous parties des calorimètres : le ECAL tonneau et bouchons, le HCAL tonneau et bouchon, le détecteur de pied de gerbe. Dans le HF il n'y a pas de regroupement, pour l'instant. Chaque cellule est considérée individuellement. Ces algorithmes de regroupement partent des cellules dont l'énergie dépasse le seuil fixé de bruit électronique. Ils associent ensuite les cellules de proche en proche pour former autant de groupes que de « graines » ;

3. Algorithme de lien (link algorithm)

Pour une particule donnée, il y a donc plusieurs informations de particle flow associées (provenant de chaque étape). L'algorithme de lien est chargé de trier toutes ces informations afin de reconstruire chaque particule et d'éviter le double comptage entre les détecteurs. Pour ce faire, il va envisager toutes les paires d'informations possibles et calculer une distance symbolique entre elles pour quantifier leurs compatibilités. Il constitue alors des blocs d'informations les plus compatibles entre elles. Un bloc contient en général 2 ou 3 types d'informations. Quand un lien est établi entre deux éléments, on extrapole la trajectoire du candidat aux autres sous-détecteurs et on regarde si d'autres informations peuvent être ajoutées. On reconstruit alors un candidat particule avec le maximum d'informations compatibles provenant de tous les sous-détecteurs.

Une fois que ces blocs d'informations ont été constitués, on procède à l'identification et à la reconstruction des particules candidats dans chaque bloc. Il en résulte une liste de particules reconstruites qui constitue une description précise de chaque événement (voir figure 2.21). Prêtons maintenant attention aux spécificités

de l'algorithme PF pour chaque objet.

## 2.4.2 Identification et reconstruction des leptons

### 2.4.2.1 Electrons

Quand un électron arrive à la surface du ECAL, une gerbe électromagnétique se développe dans les premiers centimètres du cristal. La plus grande partie de l'énergie de l'électron est alors collectée par une matrice de cristaux autour du cristal que l'électron est venu frapper. Cependant, l'électron a souvent déjà perdu une partie de son énergie dans le trajectographe en rayonnant des photons par bremsstrahlung. Ce faisant, il perd de la vitesse et sa trajectoire se retrouve plus courbée par le champ magnétique. Il en résulterait une erreur importante sur le calcul de l'impulsion de cet électron. Pour obtenir une mesure précise de l'impulsion, il est donc primordial d'identifier les photons rayonnés et d'ajouter leur énergie à celle de l'électron incident. C'est le but de la première étape de reconstruction des électrons, le superclustering. Il y a deux algorithmes différents (selon la zone d'impact), appelés hybride et d'îlots (hybrid and island algorithms) qui s'occupent de regrouper les cristaux et les photons rayonnés associés à un même électron [52]. Le point de départ (graines) des deux algorithmes sont les cristaux avec un dépôt d'au moins 1 GeV d'énergie transverse.

1. L'algorithme hybride, utilisé dans le tonneau, cherche dans le plan  $\eta - \phi$  à partir de la graine un domino  $1 \times 3$  ou  $1 \times 5$  de cristaux avec au moins 100 MeV d'énergie transverse. Les dominos sont alors alignés avec la graine suivant  $\eta$  et ils sont étendus de  $\pm 17$  cristaux. On associe ensuite ces dominos selon  $\phi$  quand un vide de moins de 100 MeV les sépare. On obtient finalement des clusters de clusters appelés superclusters.
2. L'algorithme d'îlots, employé dans les bouchons, constitue des clusters en connectant des colonnes en  $\phi$  de cristaux dont l'énergie décroît continuellement en s'éloignant du cristal graine. Des superclusters sont alors formés en agrégeant des colonnes (îlots) toujours selon  $\phi$  dans les deux directions.

Une fois qu'on a identifié un supercluster à l'aide d'un des deux algorithmes, on continue avec l'étape de reconstruction de trace. On envisage à chaque fois les deux hypothèses de charge  $\pm 1$ . L'énergie et la position du supercluster sont propagés réversiblement vers le vertex d'interaction en tenant compte des conditions de champ magnétique. On cherche alors des hits compatibles avec ce supercluster dans le trajectographe. Quand une paire compatible est identifiée, on crée une pré-trace pour l'électron. On essaye alors d'associer cette pré-trace à d'autres hits dans les autres niveaux du trajectographe en tenant compte de la probabilité d'émission de photon par bremsstrahlung. C'est le travail d'un algorithme dédié, le GSF (pour gaussian Sum Filter) [53] qui se base sur le modèle de Bethe-Heitler [54]. On réitère cet algorithme jusqu'à ce que deux niveaux consécutifs ne présentent pas de coups compatibles. Un minimum de 5 hits est nécessaire pour créer une trace pour l'électron.

Pour finir, on regroupe les informations du supercluster et de la trace associée. On combine la mesure de l'énergie dans le calorimètre  $E_{sc}$  avec la mesure du  $p_t$  dans le trajectographe  $p_{tk}$ . On améliore ainsi la mesure finale car ces deux quantités ne sont pas affectées de la même façon par le bremsstrahlung (selon l'énergie de l'électron).

#### 2.4.2.2 Muons

Pour les muons, le phénomène de bremsstrahlung est supprimé avec un facteur  $(\frac{m_e}{m_\mu})^2 \simeq 2,3 \times 10^{-5}$  par rapport aux électrons. Comme ils sont par ailleurs insensibles à la force forte, les muons traversent toutes les couches du détecteur avant d'interagir par interaction faible dans le spectromètre à muons. Les muons sont donc uniquement détectés par le trajectographe et le spectromètre à muons. L'algorithme de reconstruction des muons [55] se sert de l'information de déclenchement pour identifier des régions d'intérêt où il effectue une reconstruction locale. Il utilise les coups reçus dans les modules CSC et DT afin de reconstruire des segments de traces par interpolation linéaire. Trois types de reconstruction mènent à trois types d'objets muons complémentaires : « stand alone », « global » et « tracker ».

**Les muons « stand alone »** Cette reconstruction utilise uniquement les informations provenant du spectromètre à muons : les traces fournis par les DT et les CSC ainsi que les RPC pour l'interface tonneau bouchon. La trajectoire est reconstruite itérativement de l'intérieur vers l'extérieur du détecteur. On complète l'interpolation par l'application du filtre de Kalman qui prend en compte les pertes d'énergies, les diffusions multiples et la non homogénéité du champ magnétique.

**Les muons « global »** Ils utilisent à la fois l'information du trajectographe et des chambres à muons. L'algorithme extrapole les paramètres d'un muon stand alone avec une contrainte au vertex d'interaction dans le trajectographe. Cela permet de définir une région d'intérêt dans le trajectographe à l'intérieur de laquelle on recherchera ensuite la trace reconstruite qui s'associe le mieux au muon stand alone. Si l'on trouve une telle trace, les paramètres du muon « stand alone » sont alors ajustés.

**Les muons « trackers »** A cause du puissant champ magnétique, les muons de bas  $p_t$  peuvent ne pas laisser suffisamment d'informations dans le spectromètre à muon et être ainsi perdus par l'algorithme « stand alone ». On peut donc examiner toutes les traces du trajectographe et rechercher des indices de compatibilité avec les informations du calorimètre et des chambres à muons. Par exemple, si on peut associer une trace du trajectographe avec au moins un segment de trace dans les chambres à muons (insuffisant en lui même pour faire un muon « stand alone »), on peut le considérer comme un muon « tracker ».

Une étude récente portant sur tout le lot de données 2012 a permis d'évaluer la qualité de l'identification des muons PF dans CMS. Les résultats de la figure 2.22

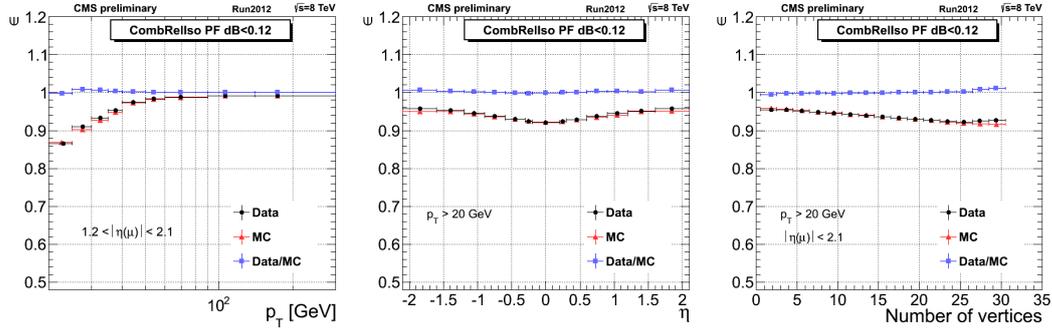


Fig. 2.22 – Efficacité de reconstruction et facteurs correctifs (data/MC) des muons en fonction du  $p_t$  (à droite), de la pseudorapidité (au milieu) et du nombre de vertex (à droite) sur l'ensemble des données 2012.

ont été obtenus avec une coupure sur l'isolation relative (delta bcorr) des muons de 0.12 (la même qui sera utilisée dans notre analyse).

### 2.4.3 Les jets

#### 2.4.3.1 Les algorithmes de formation des jets

En physique des hautes énergies, il existe deux grandes classes d'algorithmes de jets, les algorithmes de cône et ceux de regroupement séquentiel (sequential-clustering). Les deux types sont implantés dans CMSSW (le software de CMS). On retrouve l'Iterative cône et le SIScone pour la première catégorie et le  $k_t$  le Cambridge/Aachen ou l'anti- $k_t$  pour la seconde.

Les algorithmes de cône ont pour point de départ la tour (ECAL+HCAL) qui contient le plus d'énergie transverse. Ensuite, on regroupe toutes les tours dans un cône de rayon R dans le plan  $(\eta, \phi)$  autour de cette tour pour créer un premier proto jet. On définit ensuite une nouvelle direction en faisant le barycentre des tours pondérées par leur énergies. On regroupe à nouveau les tours autour de cette direction et ainsi de suite jusqu'à ce que l'énergie et la direction du jet restent stables. Tous les constituants de ce jet sont alors retirés de la liste des calotowers et on relance l'algorithme à partir de la tour la plus énergétique restante. On arrête le processus quand il ne reste plus aucune tour avec au moins 1 GeV d'énergie transverse déposée.

Les algorithmes de regroupement séquentiel sont eux basés sur la combinaison successive de paires de particules (et pseudo particules parallèles au faisceau) par rapport à leur distance réciproque. On utilise la distance  $\Delta R$  définie dans le plan  $(\eta, \phi)$ . On commence par trouver les particules qui présentent la distance minimum entre elles. Si c'est une distance entre particules on les regroupe pour former un pseudo-jet. Si c'est une distance entre une particule et l'axe des faisceaux, cette particule est déclarée comme jet et on l'enlève de la liste. On répète cet algorithme jusqu'à ce que toutes les particules aient été regroupées en jets.

Ces deux classes d'algorithme ne sont pas sensibles aux divergences infra-rouges

et colinéaires (voir chapitre MC). Ils permettent donc une interprétation robuste des événements et peuvent être appliqués aussi bien à la simulation que sur les données pour obtenir des comparaisons viables. La seule limite est qu'ils sont très gourmands en ressource CPU. Il existe différents types de jets. Pour les calojets, seuls les dépôts calorimétriques sont utilisés en entrée de l'algorithme. Pour les Jets Plus Traces (JPT) on utilise en plus l'information du trajectographe pour profiter de sa meilleure résolution en  $p_t$ . On associe les dépôts calorimétriques avec les traces en extrapolant les trajectoires. Quand on trouve une paire cohérente on retranche l'énergie de la trace aux tours calorimétriques. Enfin, les PF jets présentent la meilleure résolution en prenant en compte les informations de tous les sous-détecteurs. Ils sont maintenant largement utilisés dans les analyses de physique au vu de leurs performances.

### 2.4.3.2 Les corrections en échelle d'énergie

En général, l'énergie des jets mesurée grâce aux calorimètres sera différente de l'énergie exacte des particules constituant le jet. En effet les calorimètres ne donnent pas l'énergie absolue des particules. L'amplitude des signaux électriques émis par les PM sont seulement proportionnels à l'énergie incidente des jets. Il est donc nécessaire d'étalonner ces dispositifs afin que la simulation reproduise bien les données. Cette calibration s'appelle la correction en échelle d'énergie des jets (JES pour Jet energy scale). Dans CMS cette calibration a été séparée en quatre étapes distinctes [56] :

1. Corrections de niveau L1 dites de compensation : pour enlever l'excès d'énergie dû au bruit électronique et à l'empilement.
2. Corrections de niveau L2 dites relatives : pour atténuer les variations de la réponse des calorimètres selon  $\eta$ .
3. Corrections de niveau L3 dites absolues : pour corriger la dépendance en  $p_t$  de la réponse des calorimètres.
4. Corrections de niveau L2/L3 dites résiduelles : uniquement appliquées aux données afin de prendre en compte les différences avec la simulation.

Ces coefficients correcteurs sont déterminés grâce à de nombreuses techniques comme, par exemple, la méthode de « tag and probe » dans des topologies particulières. L'idée est de sélectionner des événements avec exactement un jet et une autre particule dont on connaît bien les caractéristiques et d'utiliser la conservation de l'énergie dans le plan transverse. En étudiant les variations de la réponse en fonction de  $\eta$ ,  $\phi$  ou  $p_t$  on peut corriger l'énergie calorimétrique. Malgré tous ces efforts, l'échelle d'énergie des jets reste l'incertitude systématique la plus contraignante (de l'ordre de 10%). Il est à noter qu'on corrige aussi la résolution des jets qui est différente dans les données et la simulation. On utilise le même genre de technique pour déterminer les coefficients correctifs pour la résolution (JER pour jet energy resolution).

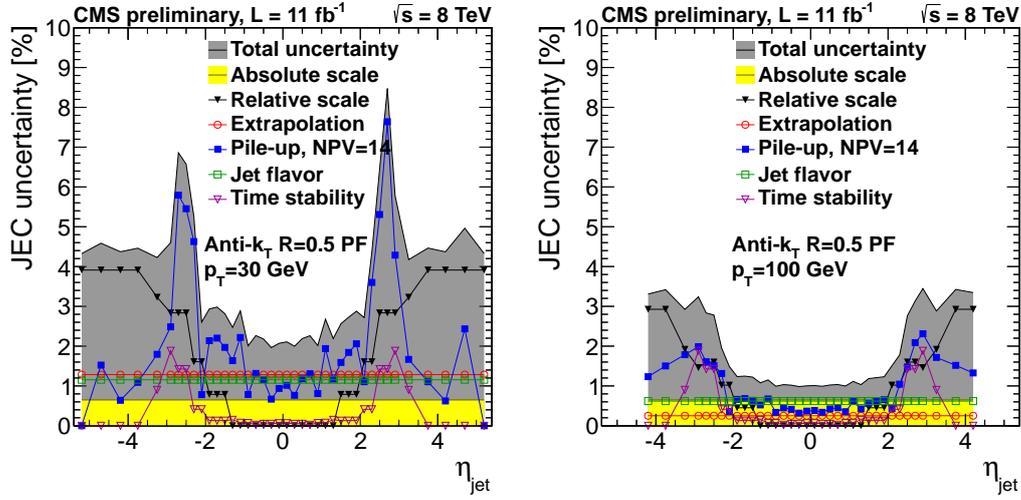


Fig. 2.23 – Incertitudes sur les corrections d'échelle d'énergie d'un jet de 30 GeV (à gauche) et de 100 GeV (à droite) en fonction de la pseudorapidité  $\eta$ .

Les incertitudes sur ces facteurs de correction en énergie évaluées sur environ la moitié des données 2012 sont présentées sur la figure 2.23

#### 2.4.4 Énergie transverse manquante

L'énergie transverse manquante est l'observable utilisée pour obtenir des informations sur les particules qui n'interagissent pas avec les détecteurs comme les neutrinos ou des particules stables prédites par des modèles de nouvelle physique. En effet l'énergie transverse étant conservée, comme elle est nulle avant collision (l'impulsion transverse des partons dans le proton est négligeable), elle doit l'être après. Ainsi dans l'hypothèse d'un détecteur parfaitement hermétique, si une particule n'interagit pas avec les détecteurs on peut estimer que toute son énergie transverse est manquante. De même, les particules mal reconstruites ou qui tombe dans des angle morts du détecteur déposeront seulement une partie de leur énergie dans le détecteur. Il faut donc distinguer l'énergie transverse manquante réelle emportée par des particules qui n'interagissent avec aucun détecteur et celle qui provient des imperfections du détecteur et des mesures. La  $\cancel{E}_T$  est ainsi très sensible à la calibration des détecteurs (notamment celle du calorimètre), à la perte de certaines cellules des détecteurs et plus généralement au bruit électronique. Comme pour les jets, il existe plusieurs manières d'estimer la  $\cancel{E}_T$  selon les informations qu'on utilise pour la calculer. On peut citer la calo  $\cancel{E}_T$  qui n'utilise que l'information des calorimètres, la JPT  $\cancel{E}_T$  (jets plus track) qui ajoute l'information tracker pour reconstruire les particules et enfin la PF  $\cancel{E}_T$  qui est la plus précise car elle utilise les données de tous les sous-détecteurs. Cette dernière est la somme vectorielle des impulsions de chaque élément de la collection de PF particules. L'équation (2.12)

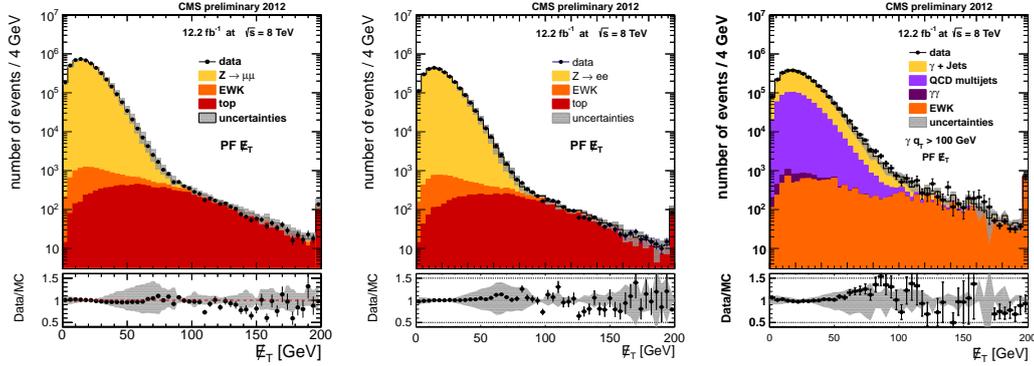


Fig. 2.24 – Distribution de  $\cancel{E}_T$  pour des événements  $Z \rightarrow \mu\mu$  (à gauche),  $Z \rightarrow ee$  (au milieu) et avec photons (à droite).

explicite la formule utilisée pour calculer cette  $\cancel{E}_T^{PF}$  :

$$\cancel{E}_T^{PF} = - \left\| \sum_{PF\text{particles}} \vec{p}_t \right\| \quad (2.12)$$

Les performances de la  $\cancel{E}_T$  calculée grâce à l’algorithme de PF sur  $12 \text{ fb}^{-1}$  des données 2012 sont présentées sur la figure 2.24. Il s’agit ici uniquement de la  $\cancel{E}_T$  dues aux erreurs de mesure sur les objets (muons, électrons, photons et jet) car on sélectionne des topologies qui ne comportent théoriquement pas d’énergie transverse manquante. On observe ainsi que la résolution en  $\cancel{E}_T$  pour l’expérience CMS est de l’ordre de 20 GeV. Mais cette résolution est très dépendante du PU et de l’activité générale dans le détecteur.

## 2.5 Traitement de données et grille de calcul

### 2.5.1 CMSSW et le format de données EDM

L’ensemble de tous les programmes utilisés pour la reconstruction ou l’analyse des données de CMS forment ce qu’on appelle CMSSW (CMS Software). C’est un framework codé en langage C++ à l’intérieur duquel de nombreux outils ont été développés pour simplifier le travail des expérimentateurs. Il est bâti autour du format de donnée de l’événement (EDM pour Event Data Model). CMSSW n’a qu’un seul exécutable pour analyser et produire les données, cmsRun. Il est configuré au moment de son exécution par un fichier en langage PYTHON qui lui spécifie :

1. sur quel échantillon tourner.
2. quels modules exécuter ainsi que leurs paramètres.
3. l’ordre d’exécution des modules.
4. à quel endroit et sous quelle forme il doit enregistrer la sortie.

L'EDM est centré autour du concept d'événement. Un événement est un conteneur C++ pour toutes les données brutes (RAW) et reconstruites associées à une collision. Les données sont transmises d'un module à l'autre par l'intermédiaire de cette structure. Tous ces événements sont ensuite enregistrés dans un fichier au format ROOT. On peut ensuite ouvrir ce fichier et naviguer facilement entre les différentes collections stockées dans l'événement et ce, à plusieurs stades de l'analyse.

Quand l'événement vient d'être enregistré il est au format brut (RAW). Seuls les signaux électriques provenant des différents sous-détecteurs sont stockés. Puis commence la phase de reconstruction, où de nombreuses quantités physiques sont calculées à partir de ces signaux électriques. On passe alors au format RECO qui contient des informations supplémentaires. Il contient notamment un certain nombre de méta-données qui décrivent la configuration de CMSSW ou les conditions de calibration utilisées pour la reconstruction. Dans la pratique, on travaille plutôt sur un sous ensemble de ces échantillons RECO appelés AOD (Pour Analysis Object Data). Ils contiennent toutes les informations nécessaires au travail d'analyse mais sont bien moins volumineux<sup>6</sup>.

### 2.5.2 La grille de calcul

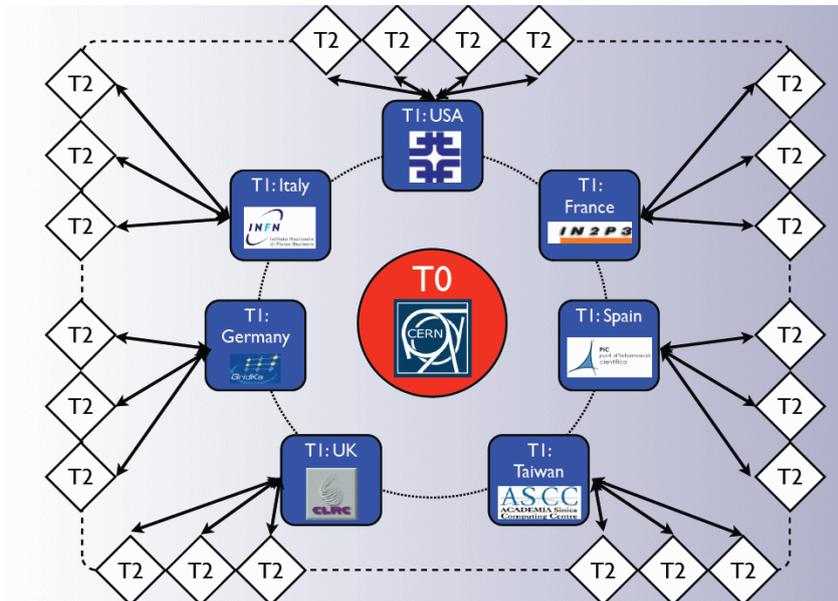


Fig. 2.25 – Organisation des différents niveaux de tiers de la grille de calcul CMS.

Les calculs nécessaires pour la reconstruction des événements et pour les analyses physiques se font sur la grille de calcul WLCG (Worldwide LHC Computing Grid) [57]. En effet, la collaboration LHC toutes expériences confondues, doit traiter environ 15 Po de données par an. Tous les nœuds de la grille sont reliés par un réseau de fibre

6. On ne conserve plus l'information sur les hits dans le trajectographe mais seulement les traces reconstruites par exemple

optique dédié très performant. La WLCG regroupe 100000 processeurs (a minima) répartis sur 32 pays. Elle est organisée selon plusieurs niveaux concentriques de tailles décroissantes :

1. Le **Tier-0** est unique, il est basé au CERN. Il est chargé de traiter les données brutes (RAW) de l'expérience CMS et de les stocker rapidement en garantissant leur intégrité. Il permet ensuite la reconstruction rapide (prompt) des données brutes pour produire les différents échantillons primaires (primary datasets) au format RECO. Son rôle est enfin de distribuer ces données brutes et reconstruites sur les différents **Tier-1**. En effet, toutes les données de CMS doivent être copiées à au moins deux endroits différents pour plus de sécurité ;
2. Les **Tier-1** sont les 7 grands centres de calculs nationaux des pays membres de la collaboration CMS (dont le CC-IN2P3 à Lyon). Ils permettent de conserver une copie des données RAW et RECO de CMS et d'organiser leur transfert aux différents **Tier-2**. Ils permettent aussi de stocker et de transférer les données simulées centralement (mais pas pour les productions MC locales). Ils sont aussi chargés des opérations de re-traitement (reprocessing) des données (RERECO) pour profiter de versions améliorées du software de CMS et de nouvelles calibrations. Enfin, ils ont aussi la charge d'extraire les AOD du format RECO (voir partie 2.5.1) ;
3. Les **Tier-2**, sont des centres de calcul universitaires de taille plus réduite mais pourvoyant une bonne capacité de calcul dédiée aux analyses de physique des laboratoires et aux productions locales de données simulées ;
4. Les **Tier-3**, sont souvent plus modestes (à l'échelle d'un laboratoire ou d'un institut), et permettent aux groupes locaux de mener à bien leurs analyses.

Cette organisation en plusieurs niveaux et le fait de séparer les échantillons selon leur contenu physique permet de répartir les calculs sur l'ensemble de la grille et d'éviter un engorgement. Aucun centre de calcul au monde ne peut aujourd'hui rivaliser avec la puissance combinée des différents tiers de la grille de calcul du LHC.



# Simulation Monte-Carlo avec MADGRAPH pour la collaboration CMS

## Sommaire

<b>3.1</b>	<b>La génération Monte-Carlo en physique des particules . . .</b>	<b>60</b>
3.1.1	Intégration numérique par tirage Monte-Carlo . . . . .	60
3.1.2	La méthode Monte-Carlo en physique des particules . . . . .	62
<b>3.2</b>	<b>La chaîne de génération d'événements Monte-carlo . . . . .</b>	<b>63</b>
3.2.1	Les fonctions de densité partonique (PDF) . . . . .	64
3.2.2	L'événement dur . . . . .	66
3.2.3	La gerbe partonique (Parton Shower ou PS) . . . . .	67
3.2.4	Appariement des jets . . . . .	68
3.2.5	Événements sous jacent et interaction partonique multiple . .	69
3.2.6	Hadronisation et désintégration de hadrons . . . . .	70
<b>3.3</b>	<b>Simulation du détecteur . . . . .</b>	<b>71</b>
3.3.1	Simulation rapide . . . . .	71
3.3.2	Simulation complète . . . . .	71
<b>3.4</b>	<b>Présentation du format LHE et des générateurs utilisés . .</b>	<b>72</b>
3.4.1	Le format LHE . . . . .	72
3.4.2	Le générateur MADGRAPH . . . . .	72
3.4.3	Le générateur PYTHIA . . . . .	73
3.4.4	Interface MADGRAPH PYTHIA . . . . .	74
<b>3.5</b>	<b>Génération et validation d'événements Monte-Carlo à grande échelle pour la collaboration CMS . . . . .</b>	<b>76</b>
3.5.1	Production d'événements MC et post processing . . . . .	76
3.5.2	Un exemple de validation d'événements physiques . . . . .	78
3.5.3	Production locale de gridpacks pour les échantillons $t\bar{t}$ systématiques . . . . .	79
3.5.4	Maîtriser l'appariement : détermination de la valeur optimale de $Q_{MS}$ . . . . .	80
3.5.5	Conclusion . . . . .	81

## 3.1 La génération Monte-Carlo en physique des particules

### 3.1.1 Intégration numérique par tirage Monte-Carlo

Je présenterai dans un premier temps le principe de l'intégration numérique par tirage aléatoire. Puis je détaillerai les difficultés qui surviennent quand on passe à des intégrales multidimensionnelles (ce qui est le cas en physique des particules). Enfin je présenterai quelques techniques d'optimisation de cette méthode. Une présentation plus détaillée de l'ensemble des techniques Monte-Carlo peut être trouvée en référence [58].

#### Intégration numérique à une dimension

Le but est de résoudre numériquement une intégrale  $I$  du type (3.1) dont on ne connaît pas de primitive analytique.

$$I = \int_a^b dx f(x) \quad (3.1)$$

On utilise pour ce faire, la relation (3.2) qui relie cette intégrale à la moyenne de la fonction sur ce même intervalle.

$$\langle f(x) \rangle_{x \in [a,b]} = \frac{\int_a^b dx f(x)}{b-a} = \frac{I}{b-a} \quad (3.2)$$

Il suffit alors de générer des nombres aléatoires à l'intérieur de cet intervalle et d'évaluer la fonction  $f$  en chacun de ces points. Si on génère suffisamment de points  $x_n \in [a, b]$  on obtiendra une valeur approchée de la moyenne de la fonction  $f$  grâce à la formule (3.3).

$$\langle f(x) \rangle_{x \in [a,b]} \simeq \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (3.3)$$

Ceci constitue le principe de base de l'intégration numérique par tirage Monte-Carlo. Il reste cependant à déterminer quelle erreur est faite sur l'intégrale pour un nombre de tirages donné. La réponse est fournie par le théorème central limite. Ce théorème fondateur de la statistique stipule que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire Gaussienne (où distribuée normalement). Dans notre cas, il garantit que si la distribution des nombre aléatoires est bien uniforme sur l'intervalle  $[a,b]$ , alors l'intégrale numérique convergera vers la bonne valeur et l'erreur commise sera liée à la variance selon l'équation (3.4).

$$I = \int_a^b dx f(x) = (b-a) \langle f \rangle \pm \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}} \quad (3.4)$$

L'erreur sur l'intégrale avec cette méthode Monte-Carlo se comporte comme  $N^{-1/2}$  ce qui est bien moins performant que d'autres méthodes d'échantillonnage

dédiées au calcul d'intégrale à une dimension. On citera par exemple la méthode des trapèzes qui converge en  $N^{-2}$ . Elle consiste à discrétiser la plage d'intégration en intervalles réguliers et à remplacer l'aire sous la courbe par l'aire du trapèze formé par les  $f(x_n)$ . La méthode Simpson qui utilise un polynôme quadratique pour approximer la fonction sur chaque intervalle permet, elle, d'obtenir une convergence en  $N^{-4}$  [58]. Pour le problème à une dimension, cette méthode Monte-Carlo d'intégration numérique est donc bien trop longue à converger et n'est, en pratique, jamais utilisée. Examinons maintenant le cas à  $n$  dimensions.

### Intégration numérique à $n$ dimensions

Dès qu'on passe à des dimensions supérieures à quatre le calcul numérique se complique. En effet pour les méthodes consistant à approximer la fonction étudiée par une fonction plus simple (méthode des trapèzes, Simpson...), il faut absolument évaluer la fonction pour chaque borne des intervalles d'intégration sur chaque dimension. Ainsi pour une intégrale à  $D$  dimensions, les méthodes des trapèzes et de Simpson convergeront respectivement comme  $N^{-2/D}$  et  $N^{-4/D}$ . En revanche la méthode Monte-Carlo se contente d'évaluer une seule fois la fonction pour chaque tirage aléatoire. La seule différence est que l'on tirera maintenant  $D$  chiffres aléatoires à chaque tirage. La méthode Monte-Carlo conserve donc sa vitesse de convergence en  $N^{-1/2}$  ce qui en fait la méthode numérique la plus efficace pour les intégrales à nombreuses dimensions. La généralisation à  $D$  dimensions de l'équation (3.4) est donnée par l'équation (3.5) où  $V$  est l'hyper volume d'intégration.

$$I = \int_V d\Omega f = V \times \langle f \rangle \pm V \times \sqrt{\frac{\langle f^2 \rangle - \langle f \rangle^2}{N}} \quad (3.5)$$

En plus de converger plus vite que les autres méthodes, elle facilite grandement l'implémentation des domaines d'intégration les plus compliqués. En effet certaines topologies peuvent poser problème quand on essaye d'approximer la fonction sur des intervalles présentant des discontinuités. Avec la méthode Monte-Carlo il suffit de rejeter les tirages aléatoires des  $n$ -uplets qui ne sont pas dans l'intervalle de définition. Attardons nous maintenant sur l'échantillonnage préférentiel (sampling), largement utilisées dans les générateurs MC en physique des particules car il permet d'améliorer les performances de l'algorithme en réduisant l'erreur.

### L'importance de l'échantillonnage

Comme on l'a vu précédemment, l'erreur faite avec la méthode Monte-Carlo peut être écrite comme  $E_{MC} = \frac{\sigma_{MC}}{\sqrt{N}}$ , où  $\sigma_{MC}$  représente la déviation standard sur l'intégrale (équation (3.6)).

$$\sigma_{MC}^2 = V^2 \int_V d\Omega f^2 - V^2 \left( \int_V d\Omega f \right)^2 \quad (3.6)$$

On peut donc améliorer la convergence de la méthode en réduisant la variance de la fonction à intégrer. En pratique, on utilise un changement de variable astucieux

qui va nous permettre de sélectionner les zones de l'espace des phases qui contribuent le plus à l'intégrale originale. Pour expliquer le principe de cet échantillonnage préférentiel, replaçons nous à une dimension comme dans l'équation (3.7).

$$\int_a^b dx f(x) = \int_a^b dx g(x) \left( \frac{f(x)}{g(x)} \right) = \int_a^b dx g(x) h(x) \quad (3.7)$$

Le but est alors de trouver une fonction  $h(x)=f(x)/g(x)$  qui varie moins vite que  $f$  (donc avec une plus petite variance). On peut montrer [58] que le meilleur choix pour  $g$  est de prendre une fonction proportionnelle à  $|f|$ . Enfin, il faut procéder au changement de variable en lui même et réécrire l'intégrale comme dans l'équation (3.8).

$$\int_a^b dx g(x) h(x) = \int_{G(a)}^{G(b)} dy h(G^{-1}(y)) \quad (3.8)$$

Toute la difficulté est donc de bien choisir la fonction  $g$ . Elle doit être inversible, intégrable et positive sur l'intervalle d'intégration, alors qu'à la base on ne sait pas intégrer  $f$  analytiquement. Il est à noter que si on génère des événements avec cette méthode, on obtiendra des événements pondérés car on a parcouru l'espace des phases de façon préférentiel. Il faut donc garder une trace de ces poids notamment pour calculer des observables cohérentes et exploitables du point de vue physique. Dans la pratique on applique la transformation inverse sur les événements et on travaille finalement sur des événements non pondérés (unweighted events).

### 3.1.2 La méthode Monte-Carlo en physique des particules

Les générateurs Monte-Carlo sont utilisés en physique des particules pour faire des prédictions théoriques du résultat des collisions à haute énergie. On peut ainsi les comparer directement aux données qui sortent des détecteurs. Des programmes informatiques ont été créés spécialement pour cette tâche, les générateurs d'événements Monte-Carlo. On retrouve ainsi deux chaînes de production d'événements parallèles (voir figure 3.1). Une pour les données qui sortent des détecteurs (voie de gauche). Et l'autre pour les événements simulés (à droite). Le point de départ de celle ci sont les générateurs d'événements MC comme MADGRAPH[59], ALPGEN[60] ou PYTHIA[61]...

Ils fournissent en sortie les quadri-impulsions de tous les partons dans l'état final d'une collision simulée. On reproduit ensuite très précisément l'effet qu'auraient eu ces événements dans le détecteur (voir partie 3.3). Les deux chaînes se réunissent ensuite et on tourne exactement les mêmes algorithmes de reconstruction sur les données et la simulation. C'est la pierre angulaire du système d'analyse de données qui va nous permettre de faire des comparaisons data/MC cohérentes et finalement d'en tirer des enseignements sur la validité de tel ou tel modèle.

Plus généralement les générateurs d'événements MC sont aussi utilisés en amont des analyses physiques. Ils permettent notamment de :

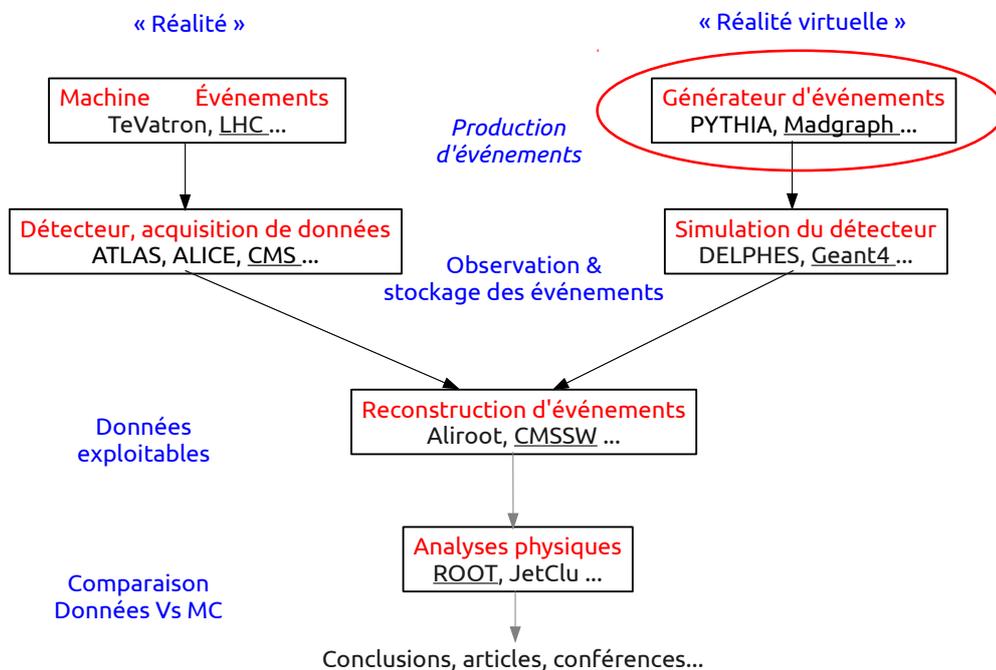


Fig. 3.1 – Comparaison de la chaîne de génération des événements pour les données et la simulation.

1. Prédire les nombres d'événements et leurs topologies (pour les modèles BSM par exemple) ;
2. Simuler les fonds pour affiner les stratégies d'analyse ;
3. Étudier les besoins des futurs détecteurs (différentes solutions techniques, mise au point du système de déclenchement...);
4. Étudier les imperfections des détecteurs pour être encore plus cohérent avec les données (évaluer les corrections en acceptance par exemple).

Examinons maintenant les différentes étapes de la production d'un événement Monte-Carlo complet.

## 3.2 La chaîne de génération d'événements Monte-carlo

Toutes les interactions au LHC font intervenir des hadrons que ce soit en tant que particules incidentes ou dans les produits de désintégration. Ces réactions impliquant des hadrons sont cependant difficiles à modéliser. En effet, selon leur énergie, on ne peut pas utiliser le même formalisme ni les mêmes approximations pour les décrire. Par ailleurs, seuls les constituants (quarks et gluons) des protons incidents interagissent en réalité (partie 3.2.1). Ces partons peuvent être considérés comme

libres seulement s'ils sont impliqués dans une collision dure avec grand transfert d'impulsion transverse (régime de liberté asymptotique de la QCD). On peut, dans ce cas là, utiliser la théorie des perturbations pour calculer exactement les sections efficaces des processus pour un certain ordre (voir partie 3.2.1). Si en revanche, on s'intéresse aux radiations de quarks et gluons il faut changer de formalisme et employer une approche phénoménologique (partie 3.2.3). La transition entre ces deux régimes n'est pas triviale et a fait l'objet d'une partie de mon travail de service (partie 3.2.4 et 3.5.4). Une fois ces partons formés, ils vont évoluer en suivant la règle du confinement de la QCD. C'est à dire que seules les particules blanches du point de vue de la couleur (avec les 3 charges de couleurs) sont autorisées. Couplée avec la désintégration des produits instables, l'évolution de ces partons forment ce qu'on appelle l'hadronisation des particules (partie 3.2.6). Enfin, on verra qu'il faut aussi se préoccuper des partons qui n'ont pas pris part à l'interaction dure. Leur présence et leur possible interaction secondaire ont des effets non négligeables, et doivent être prises en compte pour reproduire les données. Il s'agit respectivement de l'événement sous-jacent (underlying event) et de l'interaction partonique multiple (MPI), ils seront présentés dans la partie 3.2.5. Une représentation de toutes les particules produites au cours de ces différentes étapes se trouve en figure 3.2. Une étude récente couvrant en détail les différents aspects de la génération Monte-Carlo au LHC est développée dans la référence [62].

### 3.2.1 Les fonctions de densité partonique (PDF)

La génération d'événements Monte-Carlo pour un collisionneur hadronique est basée sur le modèle des partons de la QCD. Il a été postulé dès les années 60 par Bjorken et Feynman [64, 65]. Il permet une description précise des interactions ayant lieu dans les collisionneurs à haute énergie [66]. Les seules particules sensibles à l'interaction forte que l'on peut détecter expérimentalement sont des combinaisons « blanches » de quarks appelés hadrons. Ils sont constitués de partons qui peuvent être des quarks de valence (fixés pour un type de hadron) ou des gluons et des paires de quarks antiquarks provenant de la mer. A basse énergie, la constante de couplage forte est si grande que l'on ne peut pas utiliser la théorie des perturbations pour décrire un hadron. Mais quand on accélère suffisamment ce hadron le couplage fort diminue et on peut traiter le hadron comme un assemblage de partons libres (régime de liberté asymptotique de la QCD). Il reste à connaître la composition de cet assemblage pour savoir avec quelle probabilité un type de parton interagira et avec quelle fraction de l'énergie du hadron. Les fonctions qui décrivent la distribution de ces partons à l'intérieur du hadron en fonction de leur impulsion sont appelés fonctions de densité partonique (PDF)[63]. Au premier ordre,  $f_i(x_i, Q^2)$  peut être considéré comme la probabilité de trouver un parton  $i$  (gluon ou quark) avec une impulsion  $x_i$  dans un hadron sondé à l'énergie  $Q^2$  (appelée échelle de factorisation). L'évolution des PDF en fonction de cette énergie de factorisation est décrite perturbativement. Donc si on la connaît pour une certaine énergie on peut la calculer pour une autre par extrapolation [67]. Par contre, les PDF en elles mêmes ne peuvent pas être calculées

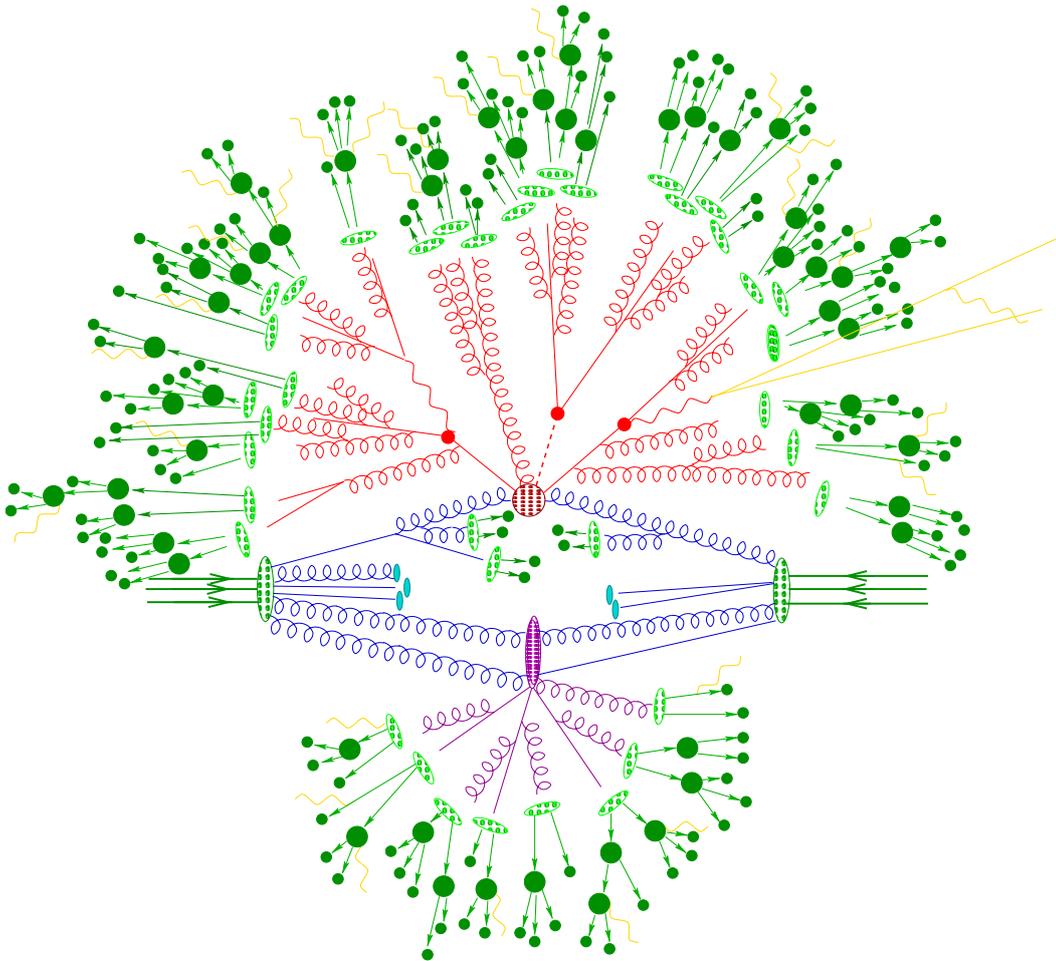


Fig. 3.2 – Représentation d'un événement Monte-Carlo produit par un générateur.

L'interaction dure (bulle rouge au centre) est réalisée par l'interaction de deux partons provenant de protons incidents (en vert), puis suivie par la désintégration de 3 partons produits (en rouge). Des radiations ISR (en bleu) et FSR (en rouge) sont ajoutées et une interaction secondaire avec le reste des partons incidents a lieu (bulle violette). On distingue ensuite l'hadronisation des partons produits (bulles vertes), puis la désintégration des hadrons instables (vert foncé). Des radiations de photons sont ajoutées à différentes étapes (en jaune).

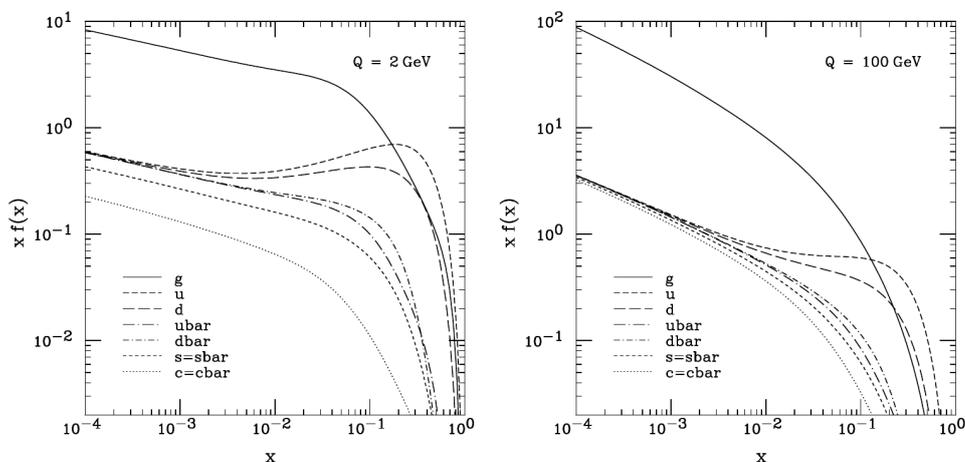


Fig. 3.3 – Fonction de densité partonique CTEQ6M [63] pour  $Q=2$  GeV(gauche) et  $Q=100$  GeV(droite) [63].

théoriquement (de manière perturbative) car elles sont associées au comportement du hadron confiné. Elles sont déterminées à partir d'ajustements globaux sur les données provenant de nombreuses expériences (voir figure 3.3). On utilise notamment les résultats des expériences de diffusion inélastique des leptons sur les hadrons qui permettent de sonder leur structure interne. Plus l'énergie de la collision est importante plus on explore la région à grand  $Q^2$ . On continue ainsi à mettre à jour les informations sur les PDF avec chaque nouvelle expérience présentant des conditions expérimentales différentes ( $E_{CM}$ , particules incidentes...). Les deux principaux groupes fournissant les paramètres des PDF sont CTEQ [63] et MSTW [68]. Il est à noter que les conditions inédites de fonctionnement du LHC forcent à extrapoler les PDF dans la zone à grand  $Q^2$  non encore explorée.

### 3.2.2 L'événement dur

La génération de l'événement dur est la première étape de la chaîne de génération (figure 3.1), elle consiste à calculer la section efficace du processus principal de l'événement. Dans un collisionneur hadronique, on se restreint à l'interaction entre deux partons  $i$  et  $j$  qui donne un état final  $X$  ( $ij \rightarrow X$ ). Le générateur Monte-Carlo doit envisager tous les processus permettant d'obtenir l'état final  $X$  et calculer leurs sections efficaces  $\hat{\sigma}_{ij \rightarrow X}$ . D'après le théorème de factorisation [69], on peut séparer l'interaction de deux hadrons entre les PDF non perturbatives (effets à longue distance) et la diffusion perturbative des partons (effets à courte distance). Ce théorème constitue la base du modèle des partons. On obtient alors la section efficace inclusive de production de l'état final  $X$  en sommant sur tous les partons incidents (spin, couleur, saveur) et en intégrant sur leurs impulsions  $x_{i/j}$  tout en

pondérant par les PDF comme dans l'équation (3.9)

$$\sigma_X = \sum_{i,j} \int \int f_i(x_1, Q^2) f_j(x_2, Q^2) \times \hat{\sigma}_{ij \rightarrow X} \quad (3.9)$$

où  $\hat{\sigma}_{ij \rightarrow X}$  est la section efficace partonique pour des partons initiaux spécifiques et  $f_{i/j}(x_{i/j}, Q^2)$  représentent la probabilité de trouver un parton  $i/j$  avec la fraction d'impulsion  $x_{i/j}$  à l'échelle  $Q^2$  (voir partie 3.2.1)

Pour calculer  $\hat{\sigma}_{ij \rightarrow X}$ , on utilise l'équation suivante issue de la théorie quantique des champs :

$$\hat{\sigma}_{ij \rightarrow X} = \int_{cuts} d\hat{\sigma}_{ij \rightarrow X} = \int_{cuts} \left[ \prod_{i=1}^N \frac{d^3 q_i}{(2\pi)^3 2E_i} \right] \delta^4 \left( p_1 + p_2 - \sum_i q_i \right) \left| \mathcal{M}_{p_1 p_2 \rightarrow \vec{q}}^{ij} \right|^2 \quad (3.10)$$

Ce calcul se factorise en deux étapes distinctes : le calcul de l'élément de matrice en lui même et l'intégration multi dimensionnelle sur tout l'espace des phases cinématique. On réalise le premier grâce à un calcul perturbatif exact (régime de liberté asymptotique de la QCD) que l'on peut pousser à différents ordres selon la précision souhaitée. Le principe est d'inventorier tous les diagrammes de Feynman contribuant à ce processus précis puis de sommer quadratiquement leurs contributions pour obtenir la section efficace. Dans les cas  $2 \rightarrow 2$  et  $2 \rightarrow 3$  les calculs restent relativement simples. Ensuite, le nombre de diagrammes augmente factoriellement en fonction du nombre de partons dans l'état final. A partir de quatre particules dans l'état final le temps de calcul devient prohibitif et d'autres méthodes doivent être employées. En pratique, les générateurs Monte-Carlo de dernière génération fournissent des résultats NNLO (next to next to Leading Order) ce qui correspond à des diagrammes de Feynman à deux boucles. C'est à cause de cette étape du calcul qu'on nomme les programmes chargés de cette tâche les générateurs à éléments de matrice (Matrix Element ou ME) ou générateurs au niveau partonique (Parton Level).

L'intégration multi-dimensionnelle sur les paramètres cinématiques est quant à elle réalisée grâce à des méthodes MC présentées plus haut (voir partie 3.1). Au LHC la multiplicité dans l'état final est de l'ordre de  $\mathcal{O}(100)$ . Il y a ici un équilibre à trouver entre la rapidité de convergence du calcul et la bonne couverture de l'espace des phases. Il est à noter que les coupures sur des observables peuvent compliquer la tâche en créant des géométries complexes pour l'espace des phases qui empêchent les algorithmes de converger. On utilise les techniques d'intégration préférentielle par canaux [70] pour résoudre ces divergences et accélérer les calculs.

### 3.2.3 La gerbe partonique (Parton Shower ou PS)

La génération de l'événement dur, aussi précise soit elle, n'est cependant pas suffisante pour décrire l'ensemble de la collision. La suite du travail est confiée aux logiciels qui décrivent les gerbes partoniques. Ils sont basés sur une approche probabiliste gouvernée par les équations DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi)

[71]. Elles décrivent l'évolution de la fonction de densité partonique et peuvent être interprétées de façon probabiliste sous la forme d'une chaîne de Markov. Cette méthode phénoménologique fournie par la QCD est uniquement valide dans l'approximation d'émission colinéaire des radiations par rapport au parton. On peut schématiser leur principe comme suit : sachant qu'un parton est caractérisé par un paramètre  $t'$ , il faut déterminer pour quelle paramètre  $t < t'$  il pourrait le plus probablement se séparer en deux partons fils en précisant leur saveurs et leurs impulsions. On peut utiliser plusieurs observables pour ordonner cette gerbe partonique mais ils seront toujours proportionnels à la masse invariante du système ( $t \simeq m^2$ ). Certains générateurs utilisent l'angle entre les partons et d'autres le moment transverse relatif. Ces branchements sont ensuite appliqués récursivement aux partons jusqu'à une valeur du paramètre pré déterminée (cutoff scale = (1GeV)).

Dans la collaborations CMS, les deux outils les plus fréquemment employés pour simuler cette gerbe partonique sont les générateurs PYTHIA [61] et HERWIG [72]

### 3.2.4 Appariement des jets

Comme nous venons de le voir, il y a donc deux approches complémentaires pour modéliser les multiples particules produites lors d'une collision. D'un coté les calculs perturbatifs d'éléments de matrice, qui décrivent la production de particules pour un processus donné par le biais des diagrammes de Feynman. De l'autre, les logiciels à gerbe partonique qui décrivent l'émission de partons secondaires grâce à une approche phénoménologique et probabiliste. Le premier calcul présente l'avantage de bien décrire les émissions à grand angle (ou dures) alors que le second traite de façon réaliste les émissions colinéaires (ou douces). Seule une combinaison maîtrisée des deux permet une bonne couverture de l'espace des phases et fournit donc des résultats en adéquation avec les données. Une des premières approches pour combiner ces deux méthodes dans les processus  $e^+e^- \rightarrow hadrons$  fut développée notamment par S. Catani et F. Krauss au début des années 2000 [73]. Ces méthodes furent ensuite étendues aux interactions hadroniques [74]. Dans les deux cas, l'espace des phases de la production de particules est scindée en deux régimes distincts. Une partie des partons est décrite uniquement par les éléments de matrice correspondants alors qu'une autre est régie par les gerbes partoniques. La grande difficulté de ce processus de combinaison et de réaliser une transition fluide entre les deux régimes et d'éviter les doubles comptages des états finals. En effet, un même état final peut être réalisé de nombreuses façons différentes selon que les jets proviennent des ME ou du PS. Il faut donc procéder à un tri des événements générés pour écarter ceux qui sont en trop. Il faut les voir comme des artifices de calcul sans réalité physique et ils détérioreraient les comparaisons avec les données. Une description plus complète et appliquée de ces techniques de matching se trouve dans la partie 3.5 puisqu'elles ont constitué une partie de mon travail de service dans la collaboration.

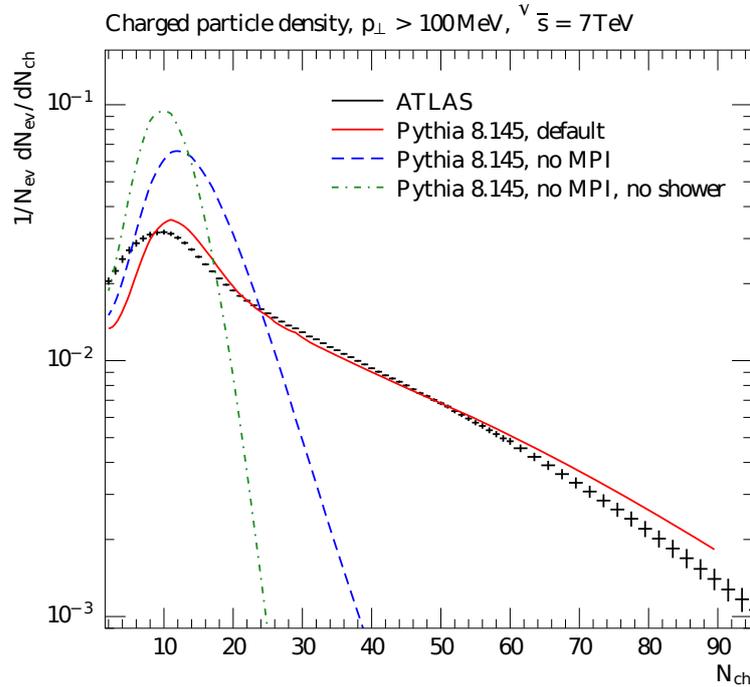


Fig. 3.4 – Comparaison des multiplicités en particules chargées entre les données et différents scénarios pour les simulations dans l'expérience ATLAS [75].

### 3.2.5 Événements sous jacent et interaction partonique multiple

#### 3.2.5.1 Evenement sous jacent

L'événement sous jacent est composé de l'activité produite dans le détecteur par les partons provenant de protons incidents qui n'ont pas participé à la collision dure. Il est important de modéliser cette activité supplémentaire pour bien reproduire les données. En effet, pour les partons portant des charges de couleur, le rayonnement de quarks et de gluons va rajouter après hadronisation de nombreuses particules dans l'état final.

#### 3.2.5.2 Interactions Partoniques Multiples (MPI)

Dans une collision proton proton, les partons n'intervenant pas dans la collision dure peuvent aussi interagir entre eux. Ces collisions parasites se superposent à la collision principale. Si on ne les simule pas correctement, à cause des radiations provenant de ces interactions, on obtient de grandes différences en multiplicité entre les données et la simulation. Pour illustrer ce fait, la figure 3.4 représente la comparaison entre les données d'ATLAS et différents scénarios de simulation (avec ou sans MPI). On voit clairement que la multiplicité prédite par les générateurs sans MPI est bien trop faible ce qui légitime leur utilisation de façon ad hoc.

Différents types de modèles [76, 77] ont été développés et implémentés dans les générateurs puis validés par la comparaison avec les données. Il s'agit à chaque fois

de modèles phénoménologiques qui doivent être paramétrés en fonction des données.

### 3.2.6 Hadronisation et désintégration de hadrons

Après la fin de l'application de la gerbe partonique, on se retrouve avec un ensemble de partons colorés dont l'énergie est de l'ordre de quelques GeV. Ces partons doivent aussi être convertis en hadrons. Il n'existe pas de théorie qui décrit ce phénomène à proprement parler mais plutôt des modèles phénoménologiques. On peut les diviser en deux catégories, les modèles basés sur des cordes de Lund (Lund string model) [78] et les modèles de regroupement hadronique [79] (cluster-hadronisation). Le premier utilise les flux de couleurs reliant les partons qui s'écartent après la collision. A cause de la forme du potentiel de QCD ( $\alpha_S$ ), plus ces flux de couleurs s'allongent plus l'énergie augmente. Quand l'énergie potentielle est suffisante pour la création d'une paire quark antiquark la ligne de flux se rompt et une paire est créée. On recommence ensuite avec de nouvelles lignes de flux reliant les nouvelles particules aux anciennes. La deuxième classe de modèle est basée sur l'idée qu'il y a une continuité dans les nombres quantiques portés par les partons produits et les hadrons dans l'état final (parton-hadron duality). On réalise alors des clusters de partons « blancs » (neutre du point de vue de la QCD). Ils forment des précurseurs pour les hadrons ; on parle de pré-confinement. Si la masse d'un cluster est proche de celle d'un hadron connu et si les propriétés correspondent (spin, saveur...), on réalise la transition directement. Si la masse est trop importante, on cherche une résonance entre deux hadrons de masse plus petite et on applique la désintégration.

Dans les deux cas, on finit avec un ensemble de hadrons stables et instables. Il est à noter que cette notion de stabilité est très relative et dépendra des expériences ou de la physique étudiée. En pratique on se fixe une limite pour la durée de vie des particules dans notre simulation. Au delà, on considère la particule comme stable et on simule ensuite son interaction avec les sous détecteurs. En deçà, on va réaliser la chaîne de désintégration de la particule jusqu'à tomber sur des produits considérés comme stables. A cause du nombre de hadrons instables dans l'état final, il est impossible de calculer les probabilités exactes de désintégration (comme on l'a fait pour l'événement dur). On se sert alors des rapports de branchement de désintégration des hadrons pour reconstituer la chaîne de désintégration. Il faut par ailleurs générer aléatoirement une cinématique pour chaque désintégration tout en conservant les corrélations de spin et de couleur. Devant la multitude de hadrons, on serait tenter de traiter chaque désintégration de façon individuelle. Cependant l'expérience montre que certaines corrélations à longue distance existent (principalement de spin couleur et saveur). Certains générateurs les prennent donc en compte grâce à des algorithmes dont la complexité dépasse le champ de cette thèse [80].

## 3.3 Simulation du détecteur

En sortie de la chaîne de production des événements simulés, on obtient les quadri-impulsions de toutes les particules stables dans l'état final. On va alors simuler l'interaction de ces particules avec notre détecteur à l'aide de logiciels dédiés (comme GEANT 4 [81]). Ils sont chargés de simuler :

- la traversée des détecteurs par les particules ;
- les déviations occasionnées par le champ magnétique ;
- les gerbes produites dans les calorimètres ;
- les effets de l'interaction particule/Matière (IPM) notamment dans les absorbeurs.

Il existe deux grands types de simulation du détecteur CMS :

- La simulation rapide, qui va à l'essentiel mais qui ne reproduit pas l'ensemble des signaux dans les détecteurs. On passe directement à l'étape d'événements reconstruits. Elle a l'avantage d'être très économe en terme de temps et elle reste très utilisée ;
- La simulation complète, réalisée à l'aide du logiciel GEANT 4. Une réplique très précise de CMS est codée dans le logiciel. On obtient donc une réponse très proche de celle du vraie détecteur. La simulation est ici bien plus longue.

### 3.3.1 Simulation rapide

Elle correspond au module de CMSSW : CMS Fast Simulation [82]. C'est un logiciel très efficace qui permet de traiter environ un événement à la seconde. On estime que l'erreur liée à son utilisation est de l'ordre du pour-cent sur les différentes observables par rapport à la simulation complète. La simulation de l'interaction des particules est simplifiée ainsi que la géométrie du détecteur. Les dépôts d'énergie des particules sont paramétrés au lieu d'être calculés. Cependant, les différentes approximations sont validées et optimisées à l'aide de la simulation complète. Tous les résultats sont vérifiés sur les données des bancs de tests (type particle gun) et des tests sur faisceau. Au final cette reconstruction rapide donne de bons résultats et reste très utilisée dans CMS notamment pour les échantillons concernant l'étude des erreurs systématiques.

### 3.3.2 Simulation complète

Elle permet, à l'aide du logiciel GEANT 4, de décrire la propagation des particules dans le détecteur en prenant en compte les déviations dues au champ magnétique et aux IPM ainsi que les désintégrations en vol. Elle simule, par ailleurs, les pertes d'énergie des particules que ce soit dans les absorbeurs ou les zones non instrumentées du détecteur. Le logiciel reproduit les interactions électromagnétiques et hadroniques des particules avec les sous-détecteurs ce qui permet de simuler précisément les signaux (hits) que ces événements auraient générés. La simulation du détecteur est, de plus, évolutive et mise à jour fréquemment en utilisant les

calibrations laser les plus récentes. Toutes ces informations sont sauvegardées à l'intérieur d'une étiquette globale qui caractérise les conditions de fonctionnement du détecteur (`globaltag`). Cette étiquette doit ensuite être utilisée de manière cohérente entre les données et la simulation (mêmes conditions de fonctionnement du détecteur). `GEANT 4` est aussi chargé de simuler la réponse électrique du système d'acquisition de données afin de se rapprocher au plus près de la réalité. Le seul désavantage de cette méthode reste son temps d'exécution de l'ordre de plusieurs minutes par événement.

## 3.4 Présentation du format LHE et des générateurs utilisés

### 3.4.1 Le format LHE

Le format Les Houches Event (LHE) [83], est un format standardisé de données pour décrire un événement au niveau partonique. Il a été créé afin d'uniformiser l'articulation avec les différents générateur PS. Le développement de cette norme a permis aux analystes de choisir librement leur générateur à élément de matrice et leur générateur PS. Avant l'avènement de ce format, seul certaines combinaisons ME/PS étaient possibles. Il résulte d'une intense collaboration entre théoriciens (qui implémentent leur modèle au niveau ME) et analystes de données qui doivent eux reproduire toute la chaîne de désintégration pour pouvoir les comparer aux données des accélérateurs. Ce format de données a été calqué sur les blocks Fortran `HEPRUP` et `HEPEUP` [84] utilisés dans `PYTHIA 6` et `HERWIG 6`. Les fichiers LHE se composent d'une part d'un bloc d'en-tête (header) qui comporte les informations dédiées à chaque générateur ME. Et d'autre part, d'un bloc qui décrit les événements de façon standardisé. Pour chaque événement on a la liste des particules impliquées dans le processus avec toutes leurs caractéristiques (PDGID, quadri-impulsion, particules mère, facteur de couleur...). Aujourd'hui, la plupart des générateurs ME produisent des événements au format LHE. Et réciproquement, tous les générateurs généralistes ont intégré un module pour lire ces fichiers en entrée et les incorporer à leur chaîne de production d'événements.

### 3.4.2 Le générateur `MADGRAPH`

`MADGRAPH` [85, 59] est un générateur à éléments de matrices LO pour les processus de physique à hautes énergies. On spécifie, le processus en terme de particules dans l'état final et initial. Puis `MADGRAPH` génère tous les diagrammes de Feynman correspondant au premier ordre et fournit la formule pour évaluer l'élément de matrice correspondant pour un point de l'espace des phases. L'évaluation de l'élément de matrice respecte les amplitudes et les fonctions d'ondes de l'Hélicité. `MADGRAPH` produit en plus le schéma des diagrammes de Feynman afin de pouvoir contrôler visuellement tous les sous processus pris en compte. Les premières versions de `MADGRAPH` étaient écrites en Fortran<sup>77</sup> mais sa dernière mouture

MADGRAPH 5 a été portée en Python. Cependant, lors de ma tâche de service, la version officielle était MADGRAPH4 et tout ce qui sera dit par la suite se rapportera à cette avant dernière version. Dans MADGRAPH 4, la génération de diagrammes suit les étapes suivantes :

1. Générer toutes les topologies avec le bon nombre d'extrémités.
2. Assigner des particules aux extrémités.
3. Identifier les vertex qui ont au plus une trace non assignée.
4. Vérifier s'il existe une interaction dans le modèle choisi qui permet de produire ce vertex avec ces particules.
5. Si oui : étape 6  
Si non : le diagramme n'est pas retenu.
6. Si le vertex comporte une trace non assignée, on la complète avec la particule appropriée (selon le modèle) et on passe au vertex suivant.  
Sinon toutes les traces sont associées à des particules, le diagramme est complet et retenu.

A la sortie de ce générateur, on récupère la liste des particules au niveau partonique avec leur quadri-impulsion sous forme d'un fichier LHE. Il peut alors être passé à un programme chargé du PS et de l'hadronisation comme PYTHIA ou HERWIG. Ou bien être passé après conversion dans une simulation simplifiée du détecteur comme DELPHES [86]. Cette deuxième méthode doit être utilisée avec précaution car il manquera alors beaucoup d'éléments physiques dans les événements. MADGRAPH est un environnement très flexible qui permet l'implémentation de nouveaux modèles par l'intermédiaire du format Universal FeynRules Output (UFO) [87]. Il est ainsi très apprécié des théoriciens et largement utilisé par les analystes dans CMS (couplé avec un programme chargé du PS).

### 3.4.3 Le générateur PYTHIA

PYTHIA est un générateur d'événements MC généraliste très utilisé dans la communauté de physique des particules. Il a été développé à l'origine par Torbjörn Sjöstrand [61]. Il a été écrit à la base en Fortran 77, mais a depuis été porté en C++. Il est capable de produire toute la chaîne de génération d'un événement. C'est un générateur LO capable de modéliser plus de 300 processus dont plusieurs modèles BSM (SUSY, Leptoquarks...). Il est surtout optimisé pour les processus  $2 \rightarrow 1$  et  $2 \rightarrow 2$  ce qui limite parfois son utilisation uniquement à la partie PS et hadronisation. Une fois l' $E_{CM}$ , la PDF et le processus spécifiés, la génération d'événements commence par la simulation du processus dur suivie de la désintégration des particules instables. Il est à noter que PYTHIA possède sa propre bibliothèque de fonctions de densité partonique mais qu'il accepte aussi les bibliothèques externes plus précises. Puis PYTHIA ajoute les corrections ISR et FSR par l'intermédiaire du PS. La simulation de l'événement sous jacent est alors prise en compte pour les partons qui n'ont pas participé à l'interaction dure. Il simule enfin l'hadronisation

des produits de désintégration selon le modèle de Lund. PYTHIA offre par ailleurs la possibilité de lire le processus dur produit par un autre générateur (MADGRAPH, SHERPA, CALCHEP ..) au format LHE et de prendre en charge la suite de la génération (PS puis hadronisation). A la fin de la chaîne on obtient la liste des particules stables générées ainsi que leurs quadri-impulsion. On classe donc PYTHIA dans la catégorie des générateurs au niveau hadronique, et la liste des particules peut directement être lue par un programme de simulation du détecteur.

### 3.4.4 Interface MADGRAPH PYTHIA

Comme nous l'avons vu, les générateurs d'éléments de matrice simulent mieux l'événement dur alors que les générateurs à gerbe partonique modélisent mieux la désintégration et l'hadronisation des produits de la collision. Le but est alors d'utiliser les avantages de chaque type de générateur pour modéliser nos événements. Il va ainsi falloir articuler les deux de façon maîtrisée pour éviter le double comptage de jets ou les trous dans l'espace des phases. La première approche naïve de simplement tourner le PS sur la sortie du ME ne donne pas de bons résultats. Le problème principal est que ces deux générateurs répondent à des logiques différentes. Les générateurs ME ont en effet une approche exclusive et donnent la probabilité qu'exactly N jets soient produits. Alors que les générateurs PS procèdent par une approche inclusive ils génèrent tous les processus qui ont au moins N jets (et calculent la probabilité du total). Il faut par ailleurs que l'espace des phases soit correctement parcouru. C'est à dire sans laisser de trou ni en comptant deux fois le même état final. Un événement donné avec N+1 jets dans l'état final peut, en effet, être produit de plusieurs manières différentes. Soit par l'évolution PS de N+1 jets à partir des N+1 partons générés par le ME. Soit par l'évolution de N jets provenant de N partons plus un jet supplémentaire émis séparément généré par le PS. Il reste donc une étape cruciale appelée appariement (matching) qui va permettre de réaliser la transition entre les deux types de générateurs et qui a occupé une partie de ma tâche de service. Le principe de base de ces méthodes d'appariement, est de séparer l'espace des phases en deux. Une partie de la production de jets sera prise en charge par le ME et le reste sera effectuée par le PS. Il existe plusieurs méthodes d'appariement, CKKW, MLM, pseudo-shower... [88, 89]. Mais nous allons uniquement décrire celle qui est utilisée à l'interface de MADGRAPH et PYTHIA, l'appariement ktMLM [90, 91].

L'appariement ktMLM est basée sur l'algorithme de jet  $k_t$  ([92]). Il repose sur le choix de la définition d'un jet (et donc d'un seuil  $k_t^{min}$ ) à deux niveaux de production. D'abord au niveau partonique en sortie du ME (appelé  $Q_{cut}$ ) puis après application du PS mais sans hadronisation (appelé  $Q_{MS}$  pour Merging Scale). Le seuil au niveau ME dépend du processus mais a été déterminé empiriquement grâce aux campagnes de production précédentes. Par contre, le  $Q_{MS}$  peut être choisi librement tant qu'il satisfait  $Q_{cut} < Q_{MS}$ . Il est cependant crucial de comprendre que c'est du choix de ce deuxième seuil que va dépendre la qualité de l'appariement. Envisageons un processus  $pp \rightarrow X + Njets$ . L'algorithme d'appariement ktMLM se déroule alors

comme suit :

1. Calculer la section efficace et générer les événements associés au processus de façon inclusive. Les événements sont générés avec un couplage  $\alpha_S^{ME}$  fixe et une multiplicité partonique maximale fixe ( $N$ ) (On appelle  $n \leq N$  le nombre de partons identifiés après ME).
2. Faire tourner l'algorithme  $k_t$  une première fois avec  $\Delta R_{ij}^{min} = Q_{cut}$  pour regrouper les partons en sortie du générateur ME.
3. Appliquer l'algorithme de gerbe partonique (PYTHIA) aux produits du ME.
4. Tourner une deuxième fois l'algorithme  $k_t$  avec cette fois  $\Delta R_{ij}^{min} = Q_{MS}$  sur les produits du PS (avant hadronisation).
5. Parcourir la liste des partons regroupés par énergie décroissante et les associer avec les jets regroupés ( $\delta R(parton, jet) < Q_{MS}$ ).
6. Les événements où tous les partons sont finalement appariés avec un jet sont conservés et les autres événements sont rejetés.
7. Si  $n < N$  tous les événements qui présentent des jets supplémentaires sont éliminés (On ne garde que les échantillons exclusifs avec autant de partons que de jets).  
Si  $n = N$  On accepte les événements avec  $n_{jets} > N$  seulement si les jets supplémentaires non appariés à des partons ont tous des impulsions transverses inférieures à celles des jets appariés.
8. Après sélection des événements, on reconstitue l'échantillon inclusif en combinant les différentes multiplicités partoniques.

Quelques remarques sur cet algorithme :

- La raison pour laquelle on veut  $Q_{cut} < Q_{MS}$  est que certains événements juste en dessous du seuil après ME peuvent passer au dessus après le PS.
- Dans l'étape 6, on rejette par exemple les événements où deux partons sont trop proches pour générer deux jets distincts après PS. Mais aussi les événements où un parton n'est pas assez énergétique pour générer un jet après PS (d'après la définition de  $Q_{MS}$ ).
- l'étape 7 permet d'éviter le double comptage. On choisit toujours que l'état final  $n+1$  jets provienne de  $n+1$  partons après ME plutôt que de  $n$  partons plus un jet supplémentaire provenant du PS (meilleure modélisation). La seule exception est le cas  $n=N$ , dans ce cas on permet des jets supplémentaires provenant du PS car ils ne pourraient pas être générés dans des événements à plus grand nombre de partons.
- Le choix du  $Q_{MS}$  va déterminer si l'appariement fonctionne bien (voir partie 3.5.4).

### 3.5 Génération et validation d'événements Monte-Carlo à grande échelle pour la collaboration CMS

Au cours de ma première année de thèse j'ai été incorporé dans l'équipe de support du générateur ME MADGRAPH. Mon travail consistait en plusieurs missions :

- Produire des événements au niveau partonique pour les processus du MS dont la collaboration avait besoin ;
- Valider ces événements une fois produits pour vérifier que les différents paramètres de génération décrivaient bien la physique (au niveau partonique et après hadronisation) ;
- Répondre aux sollicitations des groupes de physique quand ils découvraient des anomalies dans les échantillons simulés ;
- Déterminer les paramètres d'appariement qui permettaient une bonne articulation avec le générateur PS PYTHIA.

#### 3.5.1 Production d'événements MC et post processing

Le cœur de mon travail au cours de cette année 2011 a donc été la production massive d'événements MC avec MADGRAPH pour l'ensemble de la collaboration. En effet, plus la luminosité intégrée augmente plus la statistique MC nécessaire est importante. On rappelle la formule de base pour un processus de section efficace  $\sigma$  :  $N_{attendu} = \sigma \times \mathcal{L}$ . Avant toute coupure, on estime que pour les études basées sur la comparaison données/MC on a besoin d'au moins 2 fois plus de statistique générée pour un processus que le nombre d'événements attendu. D'une part, il faut prendre en compte les incertitudes statistiques dans la génération. D'autre part, les analyses vont souvent chercher des événements dans les queues de distribution après coupure, et il faut que le nombre d'événements générés soit élevé pour qu'il en reste suffisamment à ce niveau de sélection. Les histogrammes de la figure 3.5 montre le nombre d'événements générés pour différents processus du MS pendant la campagne de production 2011 à 7 TeV. On rappelle que la luminosité intégrée pour cette période de prise de données était de  $4,98 \text{ fb}^{-1}$ . Donc l'objectif était de produire environ  $10 \text{ fb}^{-1}$  de luminosité équivalente pour les processus du MS. On comprend bien que cet objectif n'est pas réalisable pour les processus dont la section efficace est très grande (par exemple W+jets ou QCD). En effet nous n'avons pas les ressources de calcul pour générer autant d'événements. C'est là que le travail d'analyse commence, puisque on va sélectionner des zones de l'espace des phases où ces fonds principaux du MS sont fortement réduits par rapport au signal. Avec l'augmentation de la luminosité en 2012, et encore plus pour la reprise du LHC, les analyses basées uniquement sur le MC simulé seront de plus en plus dures à mener à bien à cause de l'énorme statistique nécessaire. On pourra alors se diriger vers des analyses où les nombres d'événements provenant des processus du MS sont évalués dans des zones de contrôle plutôt qu'intégralement simulés.

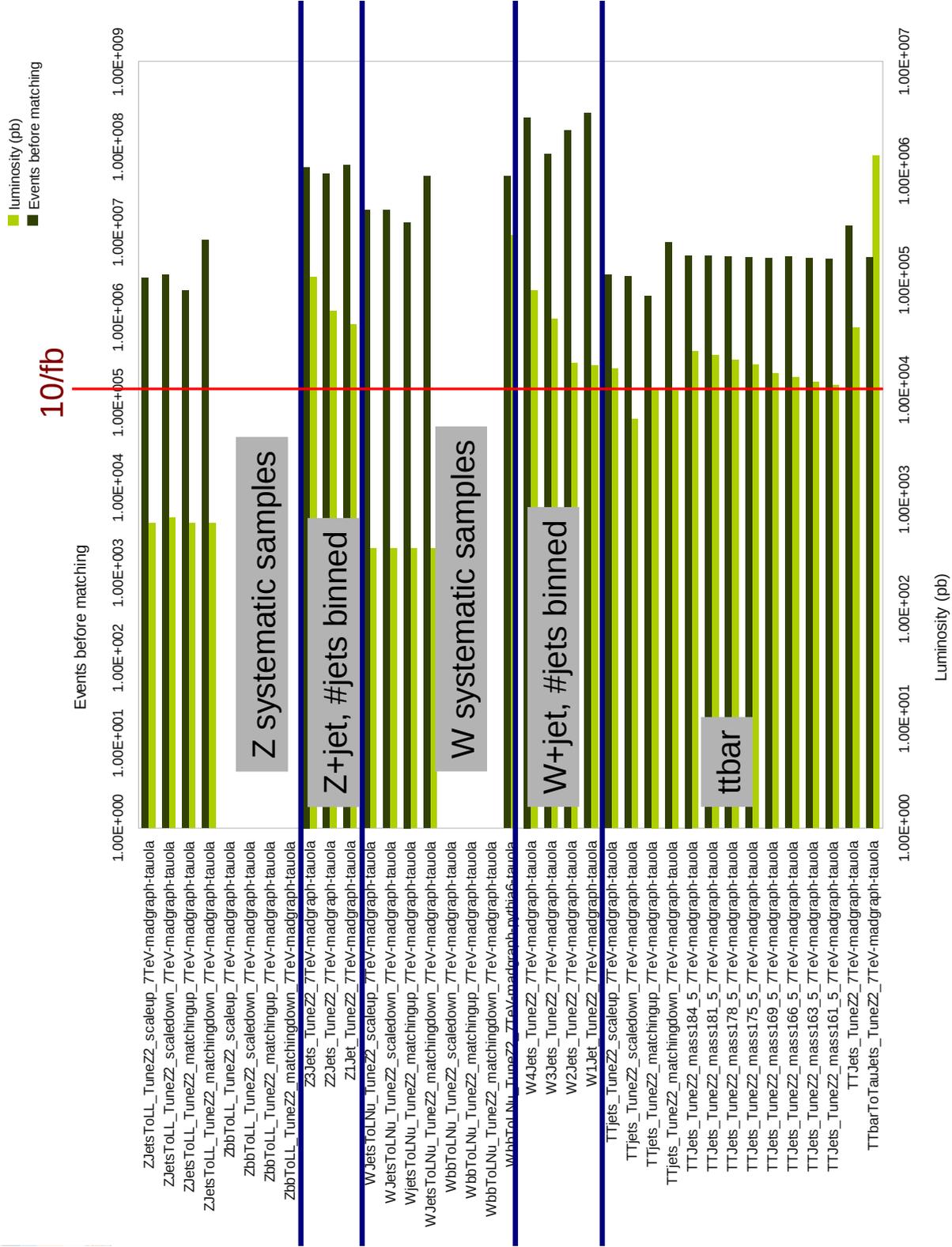


Fig. 3.5 – Événements MC générés avec MADGRAPH avant et après matching pendant la campagne 2011 à 7 TeV pour des processus Z+jets, W+jets et tt.

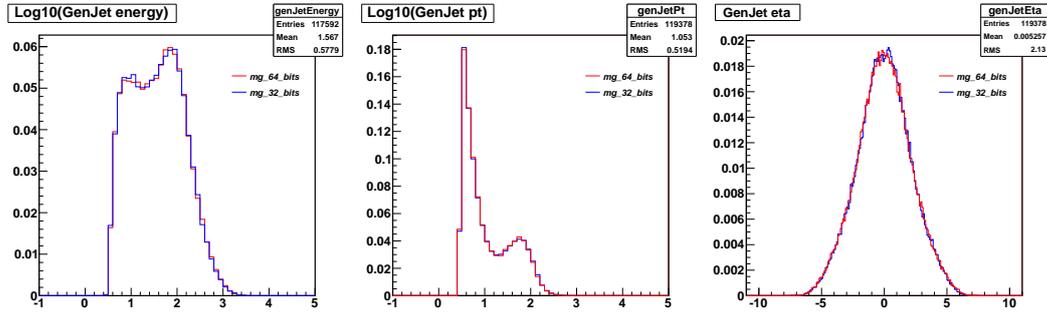


Fig. 3.6 – Distributions de validation partonique.

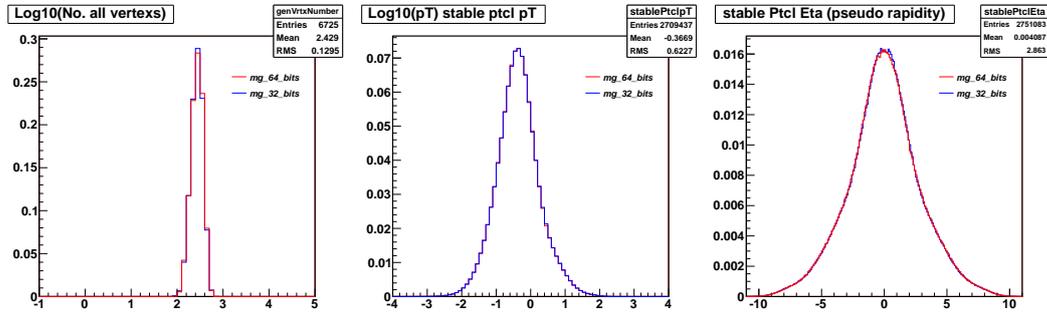


Fig. 3.7 – Distributions de validation hadronique.

### 3.5.2 Un exemple de validation d'événements physiques

Le code de MADGRAPH est en évolution continue. Une partie de mon travail au sein de l'équipe de support MADGRAPH a été de valider ces changements dans le code pour s'assurer que les distributions physiques n'étaient pas affectées. Je voudrais surtout mettre en avant le travail collaboratif entre les développeurs du code et les analystes de données. Les développeurs essaient d'optimiser leur code et de prendre en compte des processus de plus en plus compliqués. Mais ils ont besoin d'un retour sur expérience détaillé de la part des utilisateurs. Le rôle de cette équipe support était donc d'être à l'interface entre les deux communautés. Dès la parution d'une nouvelle version du code notre tâche était de générer des échantillons pour les processus MS principaux et de vérifier que les distributions étaient toujours valides. Les histogrammes présentés sur la figure 3.6 donne un exemple de ces distributions de validation au niveau partonique pour le passage du code de MADGRAPH de 32 à 64 bits. Sur la figure 3.7, on peut voir d'autres exemples de ces distributions mais cette fois au niveau hadronique après hadronisation par PYTHIA dans le cadre de CMSSW.

Malgré ces vérifications, certains bugs très spécifiques ne peuvent être découverts que par les analystes qui utilisent des variables plus complexes ou vont étudier des points très spécifiques dans l'espace des phase. Là encore notre rôle était d'examiner les problèmes que faisaient remonter les différents groupes de physique et d'identifier

à quel niveau ils se situaient. Si le problème venait effectivement d'une nouvelle caractéristique du code nous travaillions alors avec les développeurs pour fournir et valider le plus rapidement possible une nouvelle version du code.

### **3.5.3 Production locale de gridpacks pour les échantillons $t\bar{t}$ systématiques**

En période intense de production d'événements Monte-Carlo, l'ensemble des ressources de la grille de calcul doivent être mis à contribution. Lors de la campagne de production 2011, les ressources dédiées du CERN étaient saturées par la production des gridpacks et des fichiers LHE correspondant aux processus du MS. Ces échantillons étaient attendus par toute la collaboration et ne pouvaient souffrir aucun retard. Cependant, d'autres échantillons utilisés pour l'étude des incertitudes systématiques dans le groupe TOP devaient aussi être produits au plus vite. Une partie du travail qui m'a été confié cette année là a donc été d'adapter les processus de production centraux de ces fichiers sur le tiers 3 de mon laboratoire à Lyon. Il s'agissait ainsi de délester les machines de production du CERN en utilisant les ressources de calcul locales. Pour que la génération d'événements soit indépendante des configurations des différentes machines de la grille, la solution technique des gridpack a été adoptée pour la production officielle. Ce sont des archives de fichiers pré-compilés du logiciel MADGRAPH qui correspondent à chaque processus du MS. L'avantage est que la production de fichier LHE à partir de ces gridpacks donne le même résultat sur toutes les machines indépendamment des bibliothèques installées. Le défi pour moi était cependant de créer ces gridpacks dans des conditions identiques à ceux du CERN pour obtenir des lots d'événements cohérents. En collaboration avec l'ingénieur informatique de mon laboratoire, nous avons donc du adapter le serveur MADGRAPH local pour qu'il utilise exactement les mêmes bibliothèques et la même configuration que les plateformes LSF (Platform Load Sharing Facility) du CERN utilisées pour la production centrale de gridpacks. Après avoir résolu les contraintes techniques, j'ai mené une phase de validation de la production locale de nos événements en les comparant avec les fichiers LHE centraux. Je me suis ensuite chargé de la production des gridpacks pour les différents échantillons systématiques du groupe TOP<sup>1</sup>. Finalement j'ai envoyé ces gridpacks sur l'ensemble de la grille afin de produire la quantité de fichiers LHE nécessaire pour atteindre la statistique demandée par le groupe TOP pour chaque échantillon. Après une phase d'agrégation des fichiers et de validation, ces échantillons ont été mis à disposition de l'ensemble de la collaboration par l'intermédiaire d'un « wiki » dont j'avais la charge.

---

1. Variation de l'échelle de renormalisation, variation du  $Q_{MS}$  (voir partie suivante) et variation de la masse du top

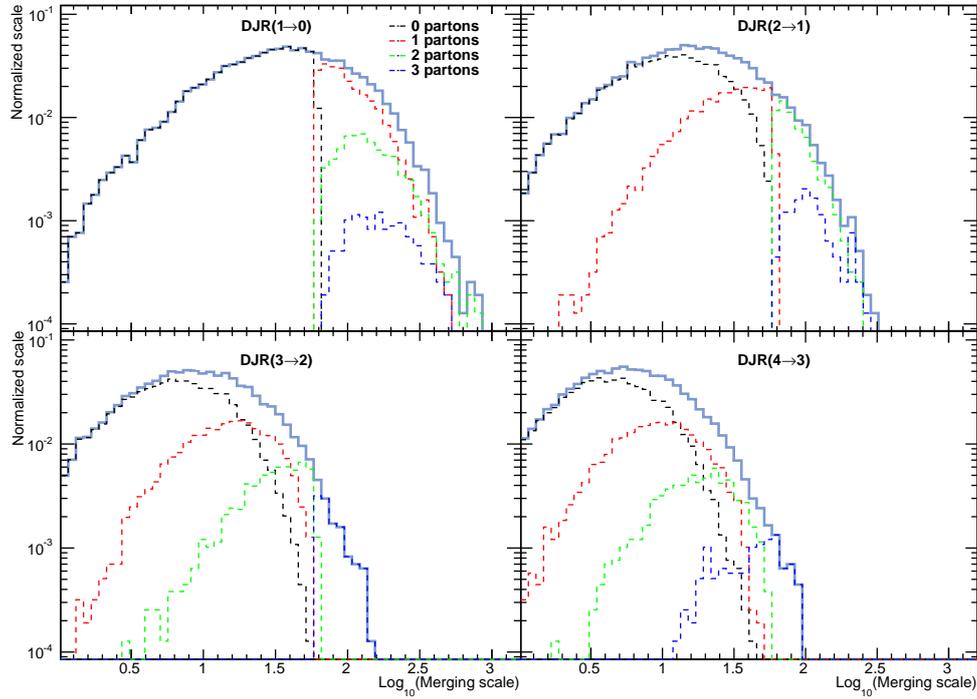


Fig. 3.8 – Diagrammes DJR pour l’échantillon  $t\bar{t}$  réalisé avec MADGRAPH et hadronisé avec PYTHIA pour un  $Q_{MS}$  approprié (60 GeV).

### 3.5.4 Maîtriser l’appariement : détermination de la valeur optimale de $Q_{MS}$

Pour s’assurer que l’appariement a bien fonctionné, nous allons nous intéresser au taux de jet différentiel. Dans un algorithme de jets type  $k_t$ , on spécifie une distance minimum appelée  $Q_{MS}$  entre les jets. Dans le cas d’une articulation ME/PS, tous les jets dont la distance est supérieure à cette valeur seront générés par le ME et tous ceux dont la distance est inférieure avec le PS. Quand on change la valeur de  $Q_{MS}$ , le nombre de jets dans l’événement peut changer. Le taux de jets différentiel  $n \rightarrow n - 1$  est la valeur du  $Q_{MS}$  pour laquelle on passe de  $n$  jets à  $n-1$  jets dans un événement donné. En faisant cela pour chaque événement, on obtient les diagrammes DJR (Delta Jet Rate). Pour vérifier que la valeur du  $Q_{MS}$  a été bien choisie, on vérifie que ce diagramme DJR est continu autour de cette valeur seuil pour les différentes transitions  $n \rightarrow n - 1$ . Si c’est le cas (voir figure 3.8), on peut considérer que l’appariement a bien fonctionné, sinon les événements générés ne pourront pas décrire précisément les collisions vraies (voir figure 3.9).

Une partie de mon travail de service a donc été de déterminer la valeur du  $Q_{MS}$  optimum pour différents processus une fois que les événements ME avaient été générés. En pratique on part du fichier LHE et on l’hadronise à l’aide de PHYTIA dans le cadre de CMSSW avec différents  $Q_{MS}$ . Puis on produit les diagrammes DJR et on vérifie visuellement leur continuité. Pour trouver la valeur optimum du

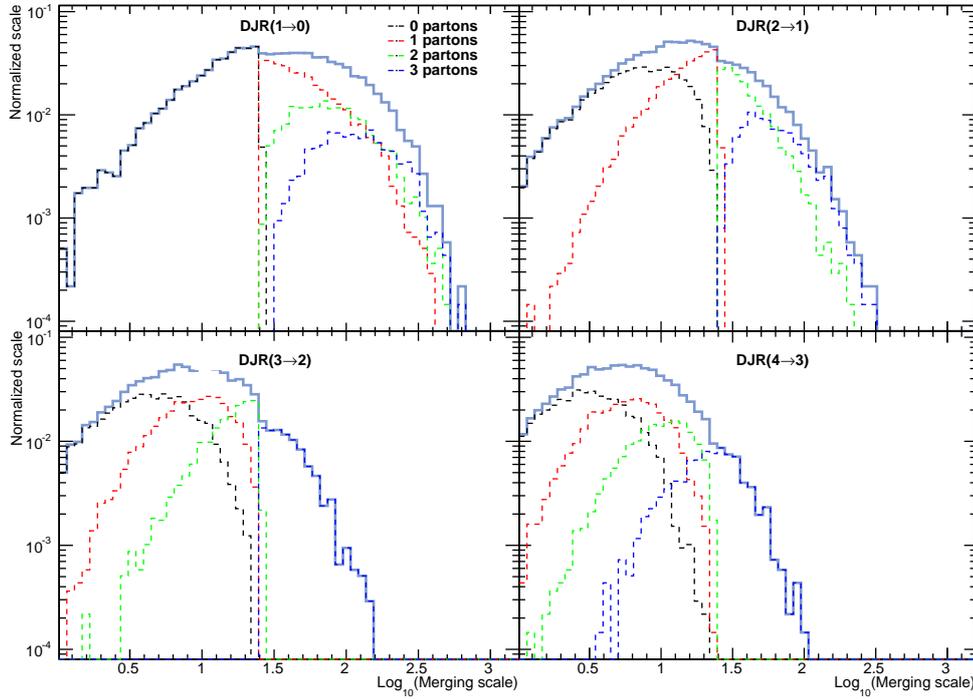


Fig. 3.9 – Diagrammes DJR pour l'échantillon  $t\bar{t}$  réalisé avec MADGRAPH et hadronisé avec PYTHIA pour un mauvais  $Q_{MS}$  (10 GeV).

$Q_{MS}$  parmi celles qui laissent ces diagrammes continus, on rajoute une condition sur la section efficace. En effet, la variation du  $Q_{MS}$  engendre une variation de cette dernière (différent nombre de jets). Parmi celles qui laissent les diagrammes DJR continus, on décide alors de retenir comme valeur pour  $Q_{MS}$  celle qui correspond au centre du plateau des sections efficaces. La valeur ainsi déterminée sera ensuite utilisée pour la production centralisée d'événements MC ce qui assure à toute la collaboration de travailler avec des échantillons simulés produits en cumulant les avantages des générateurs ME et PS de façon cohérente.

### 3.5.5 Conclusion

En conclusion de cette partie portant sur mon travail de service dans le groupe générateur, je voulais souligner la position centrale qu'il m'a permis d'occuper. En effet, j'ai eu le privilège de travailler en étroite collaboration avec les créateurs de MADGRAPH mais aussi en relation directe avec tous les groupes de physique de CMS. Cette première année de thèse m'a, de plus, permis d'acquérir la maîtrise d'outils indispensables à mon travail d'analyse de données comme le calcul sur grille, l'utilisation poussée de ROOT ou encore les fondamentaux de la simulation d'événements MC en physique des particules. Enfin, cette position charnière m'a fait prendre conscience des enjeux politiques et des relations parfois tendues entre les différents groupes de physique. La compréhension de ces mécanismes inhérents à

toutes les grandes collaborations me sera sans aucun doute très utile dans le reste de ma carrière.

# Recherche de leptoquarks de 3 ème génération dans l'expérience CMS

## Sommaire

<b>4.1</b>	<b>Phénoménologie des leptoquarks au LHC . . . . .</b>	<b>84</b>
4.1.1	Modes de production des Leptoquarks au LHC . . . . .	84
4.1.2	Sections efficaces de production des leptoquarks scalaires . . . . .	85
4.1.3	Simulation du signal . . . . .	88
<b>4.2</b>	<b>Simulation MC des événements de fonds du MS . . . . .</b>	<b>88</b>
<b>4.3</b>	<b>Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 . . . . .</b>	<b>90</b>
4.3.1	Lots de données utilisés et facteurs correctifs . . . . .	90
4.3.2	Contrôle des fonds du modèle standard . . . . .	94
<b>4.4</b>	<b>Recherche de Leptoquarks de 3 ème génération se désintégrant en top <math>\tau</math> . . . . .</b>	<b>100</b>
4.4.1	État final et canal d'analyse . . . . .	100
4.4.2	Pré-sélection des événements . . . . .	102
4.4.3	Les incertitudes statistiques . . . . .	104
4.4.4	Les incertitudes systématiques . . . . .	105
4.4.5	Optimisation finale . . . . .	107
4.4.6	Masse limite et interprétation des résultats . . . . .	110
<b>4.5</b>	<b>Recherche de Leptoquarks de 3 ème génération se désintégrant en top <math>\nu</math> . . . . .</b>	<b>110</b>
4.5.1	État final et canal d'analyse . . . . .	110
4.5.2	Pré-sélection des événements . . . . .	114
4.5.3	Optimisation finale et incertitudes . . . . .	115
4.5.4	Masse limite et interprétation des résultats . . . . .	115
<b>4.6</b>	<b>Recherche de Leptoquarks de 3 ème génération se désintégrant en top <math>\mu</math> . . . . .</b>	<b>118</b>
4.6.1	État final et canal d'analyse . . . . .	118
4.6.2	Pré-sélection des événements . . . . .	118
4.6.3	Optimisation finale et incertitudes . . . . .	121
4.6.4	Masse limite et interprétation des résultats . . . . .	125
<b>4.7</b>	<b>Conclusions et perspectives . . . . .</b>	<b>127</b>

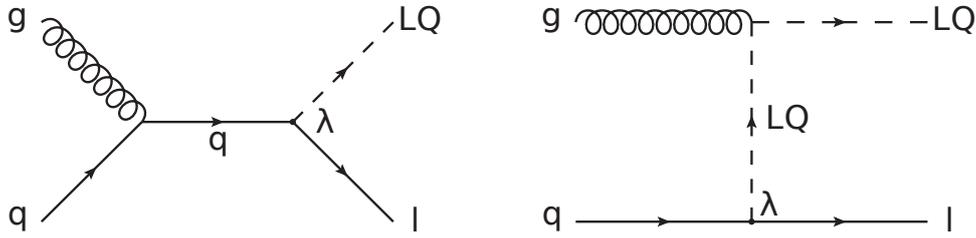


Fig. 4.1 – Diagrammes de feynman de production de LQ célibataire au premier ordre en voie s (à gauche) et voie t (à droite).

## 4.1 Phénoménologie des leptoquarks au LHC

### 4.1.1 Modes de production des Leptoquarks au LHC

Dans les collisions protons protons du LHC, il y a deux modes de production principaux des leptoquarks : la production de LQ célibataire et la production par paire. La production de LQ célibataire (voir diagrammes 4.1) fait apparaître un vertex  $LQ - \ell - q$  et sera donc proportionnelle au couplage de Yukawa  $\lambda_{LQ-\ell-q}$ . Cette section efficace est donc dépendante du modèle. De plus, la valeur de ce couplage est attendue comme très faible par rapport au couplage de l'interaction forte. La production de LQ célibataire est donc possible au LHC mais elle est beaucoup plus rare que la production par paire [93]. En fait, elle ne devient significative à 14 TeV que pour des masses de l'ordre de 2 TeV.

La production par paire (voir diagrammes 4.2) par contre, ne contient pas de vertex  $LQ - \ell - q$ , elle se produit principalement via l'interaction avec un gluon. La production par paire de LQ scalaires est entièrement régie par la QCD ce qui la rend très indépendante du modèle de LQ choisi. Dans le cas de LQ vecteurs par contre, il existe en plus une dépendance de la section efficace aux couplages anormaux  $k_G$  et  $\lambda_G$ , ce qui complexifie l'espace des paramètres et rend la recherche de LQ vecteur dépendante du modèle étudié. Dans cette étude, nous choisissons donc de nous intéresser uniquement à la production par paire de LQ scalaires pour rester le plus inclusif possible. Ceci présente l'avantage d'une simplification de l'espace des phases et d'une recherche indépendante du modèle qui prédit les LQ. Cependant, en cas de découverte, on ne pourrait trancher sur le type de LQ découvert. Cette production est dominée par les processus de fusion de gluons ou d'annihilation quark anti-quark qui dépendent uniquement de la constante de couplage de l'interaction forte. Seules des paires de LQ conjuguées de charge peuvent être produites par ce biais. Pour les basses masses de LQ la fusion de gluons domine. L'annihilation quark anti-quark devient non négligeable à haute masse. Elle atteint même 30% de la section efficace totale pour des masses de leptoquarks de 1,5 TeV. Les seuls paramètres pertinents pour cette étude sont la masse du leptoquark  $m_{LQ}$  et le rapport de branchement de désintégration vers la configuration recherchée. Par la suite nous fixerons le rapport de branchement de désintégration à 100 % laissant pour seul paramètre libre  $m_{LQ}$ .

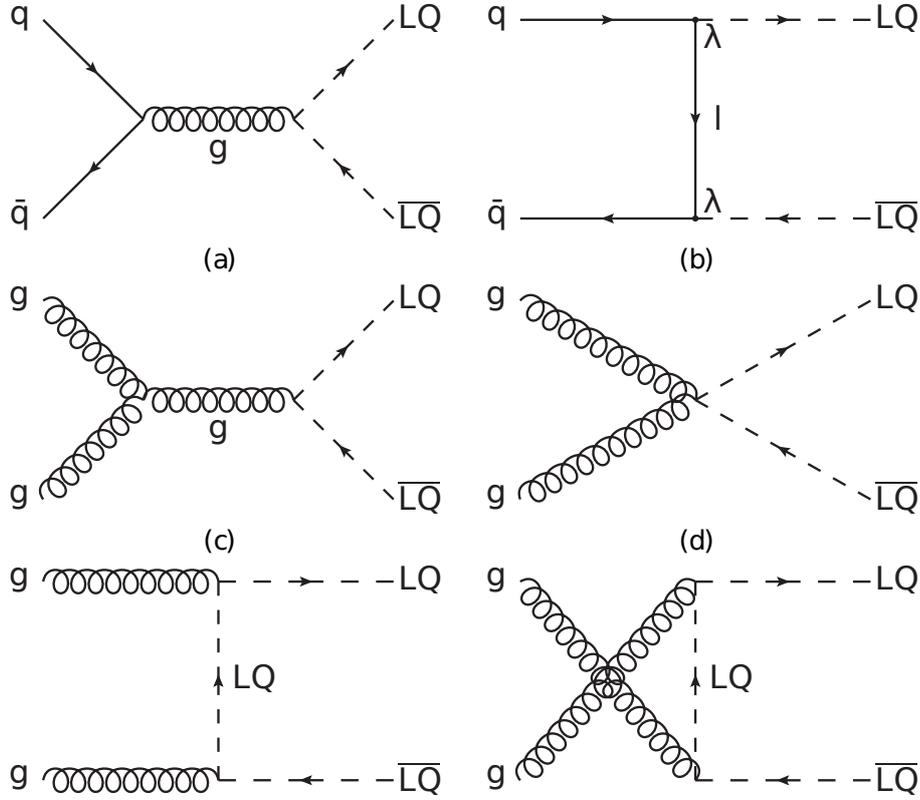


Fig. 4.2 – Diagrammes de Feynman de production de paires de LQ au premier ordre par annihilation quark anti-quark (a), par voie t (b) et par fusion de gluons (c, d, e, f).

#### 4.1.2 Sections efficaces de production des leptoquarks scalaires

Comme nous l'avons vu, la production de paires de leptoquarks par fusion de gluons et par annihilation quark anti-quark domine au LHC. Les sections efficaces au premier ordre pour ces deux modes de productions de leptoquarks scalaires sont [94, 95] :

$$\hat{\sigma}_{gg}^{LO} = \frac{\alpha_s^2 \pi}{96 \hat{s}} \left[ \beta(41 - 31\beta^2) + (18\beta^2 - \beta^4 - 17) \log \frac{1+\beta}{1-\beta} \right] \quad (4.1)$$

$$\hat{\sigma}_{q\bar{q}}^{LO} = \frac{2\alpha_s^2 \pi}{27 \hat{s}} \beta^3 \quad (4.2)$$

avec  $\hat{s}$  l'énergie dans le centre de masse et  $\beta = (1 - 4m_{LQ}^2/\hat{s})^{1/2}$  le facteur de Lorentz.

$$\sigma_{pp \rightarrow LQ\bar{L}Q}^{LO} = \sum_{i,j} \int \int f_i(x_i, Q^2) f_j(x_j, Q^2) \times \hat{\sigma}_{ij}^{LO} dx_i dx_j \quad (4.3)$$

avec  $f_{i/j}(x_{i/j}, Q^2)$  la probabilité de trouver un parton i/j avec la fraction d'impulsion  $x_{i/j}$  à l'échelle  $Q^2$  et  $\hat{\sigma}_{ij}^{LO}$  une des deux sections efficaces vues plus haut

$M_{LQ}(\text{GeV})$	$\sigma(\mu = M_{LQ})$	$\delta(PDF)$	$\sigma(\mu = M_{LQ}/2)$	$\sigma(\mu = M_{LQ} \times 2)$
200	17,4	1,2	15,0	19,7
300	1,89	0,21	1,63	2,13
400	0,342	0,052	0,295	0,385
500	0,082	0,016	0,070	0,092
600	0,0235	0,0053	0,0201	0,0265
700	0,0076	0,0020	0,0064	0,0086
800	0,00269	0,00081	0,00228	0,00304
900	0,00101	0,00035	0,00086	0,00115
1000	0,00040	0,00016	0,00034	0,00045

TABLE 4.1 – Les colonnes 2 et 3 présentent les sections efficaces NLO (en pb) accompagnées des incertitudes dues aux pdf pour les différents échantillons générés. Les colonnes 3 et 4 indiquent les sections efficaces NLO (en pb) calculées en faisant varier l'échelle de renormalisation/factorisation.

(équations 4.1 et 4.2).

Les sections efficaces NLO de la production de paires de leptoquarks scalaires ont été déterminées comme dans [95]. Ces sections efficaces sont la somme  $\sigma_{qq}^{LO} + \sigma_{gg}^{LO}$  calculées par PROSPINO [96] puis corrigées pour obtenir le résultat à l'ordre supérieur des perturbations (NLO) [95]. Les contributions NLO sont de plusieurs types :

- ajout des diagrammes dans la voie t avec des interactions quarks gluons ;
- prise en compte des radiations de gluons dans l'état final (FSR) et dans l'état initial (ISR) ;
- calcul des corrections virtuelles correspondant à l'ajout de boucles de particules dans les diagrammes.

Pour appliquer ces corrections, il faut impérativement passer par l'étape de renormalisation (notamment à cause de l'émission de gluons colinéaires qui font apparaître des divergences). Le paramètre  $\mu$  qui correspond à l'échelle d'énergie de renormalisation est choisi égale à la masse du LQ ( $\mu = m_{LQ}$ ). Le résultat du calcul de section efficace dépend donc du choix de cette échelle d'énergie. Il faut ainsi prendre en compte l'incertitude systématique associée à la variation de  $\mu$ . Les contributions à l'ordre supérieur de la section efficace totale permettent de définir le facteur  $k = \frac{\sigma_{NLO}}{\sigma_{LO}}$ . Ces corrections sont intégrées numériquement dans PROSPINO en utilisant les fonctions de structure CTEQ6M [97]. Le tableau 4.1 récapitule les valeurs de ces sections efficaces et leurs incertitudes pour chaque masse de leptoquark entre 200 et 1000 GeV. L'évolution de la section efficace et des incertitudes en fonction de  $m_{LQ}$  est présentée sur le graphique 4.3.

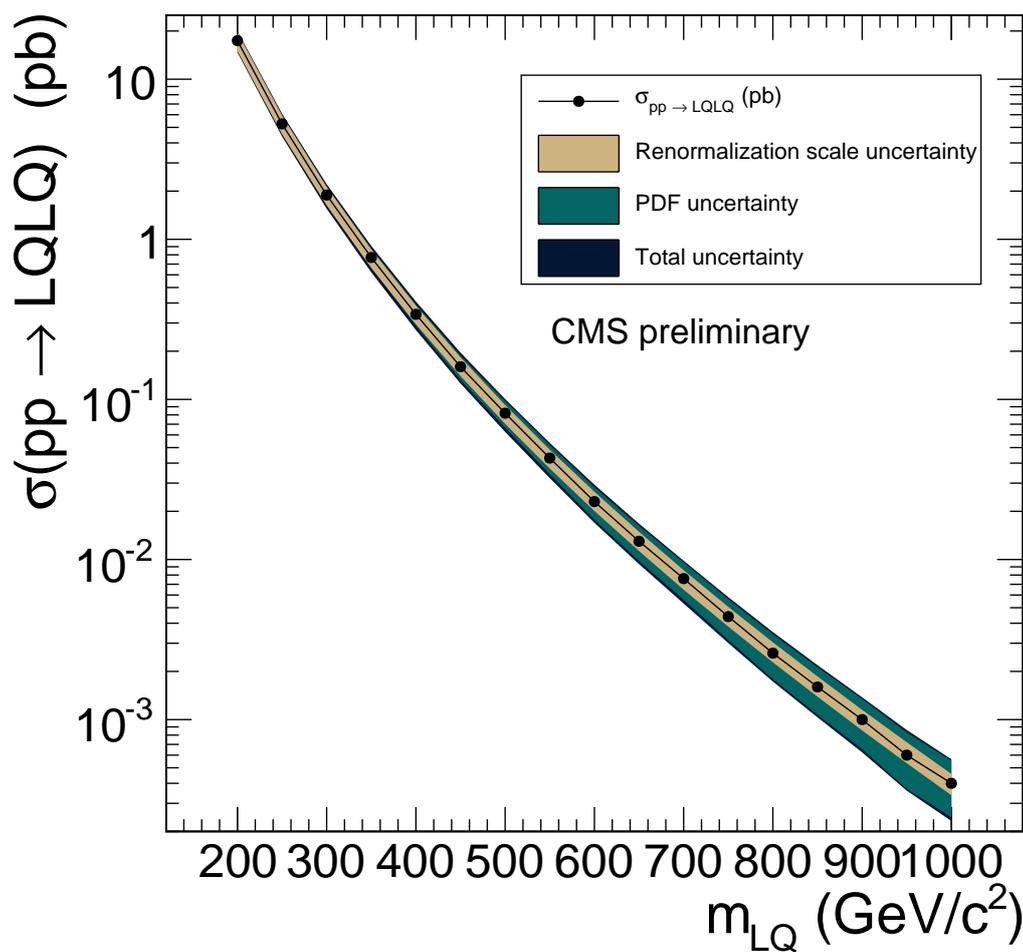


Fig. 4.3 – Section efficace théorique de production de paire de LQ à 8 TeV au LHC. Les incertitudes hautes et basses liées à l'échelle de renormalisation et aux PDF sont respectivement représentées par les bandes de couleurs beige et turquoise.

### 4.1.3 Simulation du signal

Les échantillons de signaux ont été réalisés avec le générateur PYTHIA 6 [61] et avec les fonctions de distributions partoniques CTEQ6M [97]. Les échantillons couvrent l'intervalle de masses [200 GeV, 1000 GeV] avec des pas de 100 GeV. Pour chaque valeur de masse, 25000 événements sont produits au niveau générateur.

Pour vérifier la validité de la génération des événements de signal on compare les distributions pour différents points de masse de LQ. La figure 4.4 montre l'impulsion transverse au niveau généré (avant simulation du détecteur) des LQ, des muons, des jets ainsi que la  $\cancel{E}_T$  pour 3 masses de LQ différentes dans le cas de la désintégration des LQ en top  $\tau$ . Plus la masse des LQ augmente, plus les muons et les jets produits seront énergétiques. De même, le neutrino produit dans la désintégration du  $\tau$  emporte plus d'énergie ce qui occasionne une  $\cancel{E}_T$  plus importante. Les différents scénarios envisagés pour la désintégration des LQ seront détaillés dans les analyses correspondantes (parties 4.4, 4.5 et 4.6).

## 4.2 Simulation MC des événements de fonds du MS

Dans l'étude de la désintégration de paire LQ de troisième génération en quark top et lepton, la principale difficulté vient de la paire  $t\bar{t}$  produite avec les leptons. En effet, la désintégration de cette paire va conduire à des états final très proche de certains processus du modèle standard qui constitueront le bruit de fond de notre analyse. On retrouve 6 sources principales de bruits de fond correspondant à notre état final. Ce sont, par ordre d'importance dans notre sélection optimale :

- la production de paires  $t\bar{t}$  qui ressemble le plus à notre signal et qui sera donc le plus difficile à réduire ;
- la production de bosons  $W^\pm$  avec des jets associés. On peut espérer réduire ce fond en demandant explicitement des quarks b ;
- la production de top célibataire peut produire un lepton isolé, mais présente en général une faible multiplicité en jets ;
- la production de bosons  $Z^0$  avec des jets associés produit deux leptons isolés. Comme pour le fond W on espère le réduire grâce à la multiplicité en jets ;
- les productions de dibosons (WW, WZ, ZZ) peut amener jusqu'à 4 leptons isolés mais leurs sections efficaces sont faibles.

Le tableau 4.3 présente les différents échantillons MC utilisés, les nombres d'événements générés et les sections efficaces de production. Pour savoir comment ont été obtenues ces section efficaces et connaître leurs incertitudes, voir aussi le tableau 4.6 dans la partie 4.4.4.

Il est à noter que les processus QCD peuvent occasionnellement produire des topologies proches de notre signal. Mais ils sont très largement supprimés par la demande d'un jet de b et par la coupure sur la masse transverse du W (voir partie 4.3.2.2). Il a donc été décidé de ne pas utiliser ces fonds QCD qui représentent énormément de ressources de calcul pour très peu d'événements au niveau de sélection final. Ce choix est d'ailleurs validé a posteriori par le très bon accord données/MC

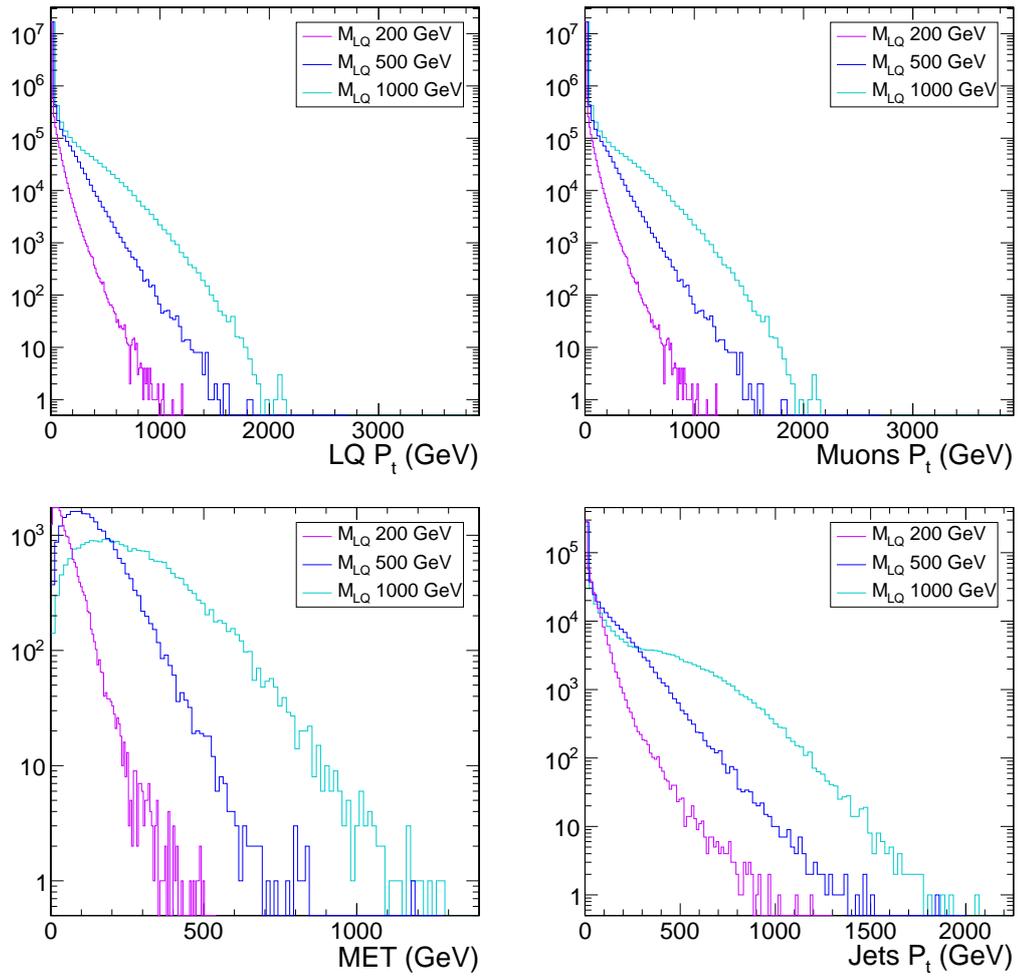


Fig. 4.4 – Impulsion transverse au niveau généré des LQ (en haut à gauche), des muons (en haut à droite), des jets (en bas à droite) et  $\cancel{E}_T$  (en bas à gauche) dans le cas  $LQ \rightarrow t\mu$

obtenu avant optimisation finale.

## 4.3 Prérequis pour l'analyse des données de CMS à 8TeV Run 2012

### 4.3.1 Lots de données utilisés et facteurs correctifs

#### 4.3.1.1 Conditions de déclenchement

Afin de réduire l'ensemble des données fournies par CMS on les subdivise en lots de données associés à différents états finals. Ainsi, selon l'analyse physique que l'on veut mener, on s'intéresse uniquement aux événements qui ont passé un certain chemin de déclenchement. Pour les analyses  $LQ \rightarrow t\tau$  et  $LQ \rightarrow t\mu$ , le canal le plus indiqué est 1 muon accompagné de plusieurs jets ce qui correspond au lot de donnée MuHad et au filtre MuPlusjet du lot de donnée SingleMu. Pour l'analyse  $LQ \rightarrow t\nu$ , on utilisera aussi cet échantillon de données en supposant qu'au moins un des deux quarks top produira un muon dans sa chaîne de désintégration. Afin de contrôler la qualité de nos muons, on utilisera aussi le canal DoubleMu (voir partie 4.3.2.1). Les différents lots de données utilisés et les chemins de déclenchement associés sont présentés dans le tableau 4.3.

#### 4.3.1.2 Pré-sélection des objets

Pour la sélection des objets, nous suivrons les recommandations officielles du groupe top de CMS dont l'état final semi leptonique ressemble beaucoup à celui de nos analyses. Les muons sélectionnés doivent ainsi :

- être reconstruits par les algorithmes de particle flow ;
- avoir été identifiés comme Muon global (voir partie 2.4) ;
- avoir une impulsion transverse supérieure à 26 GeV ;
- être centraux  $|\eta| < 2.1$  ;
- être associés à une trace de bonne qualité :  $\chi^2 < 10$ , nombre de hits dans le trajectographe  $> 5$ , nombre de stations correspondantes  $> 1$  ;
- être associés à une trace qui passe à moins de 2 centimètres du vertex primaire dans l'absolu et à moins de 5 centimètres selon l'axe Z ;
- présenter une isolation basée sur le PF (corrigée selon  $\Delta\beta$ ) inférieure à 0.12 pour un cône de 0.4.

Les jets sélectionnés, quant à eux, doivent :

- avoir une impulsion transverse supérieure à 40 GeV (Voir plus loin pour plus de précision) ;
- être centraux  $|\eta| < 2.5$  ;
- avoir été corrigés énergétiquement aux niveaux L1,L2 et L3 (+L2L3 résiduel dans les données) (voir partie 2.4.3.2) ;
- avoir été corrigés du point de vue de la résolution (les jets dans le MC ont toujours une meilleure résolution que dans les données) ;

### 4.3. Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 91

Processus	Datasets	$\sigma(pb)$	Événements générés	$\mathcal{L}$ correspondante ( $fb^{-1}$ )
$t\bar{t}$	TTJets_FullLeptMGDecays_8TeV-madgraph	26,6	12011428	452,3
	TTJets_SemiLeptMGDecays_8TeV-madgraph	115,6	24953451	215,9
	TTJets_HadronicMGDecays_8TeV-madgraph	115,6	31223821	270,1
$W \rightarrow l\nu + jets$	W1JetToLNu_8TeV-madgraph	6662,8	23141598	3,5
	W2JetsToLNu_8TeV-madgraph	2159,2	34044921	15,8
	W3JetsToLNu_8TeV-madgraph	640,4	15539503	24,3
	W4JetsToLNu_8TeV-madgraph	264,	13382803	50,7
Top célibataire	Tbar_s-channel_8TeV-powheg	1,8	139974	77,7
	Tbar_t-channel_8TeV-powheg	30,7	1935072	63,0
	Tbar_tW-channel_8TeV-powheg	11,1	493460	44,5
	T_s-channel_8TeV-powheg	3,8	259961	68,4
	T_t-channel_8TeV-powheg	56,4	3758227	66,6
	T_tW-channel_8TeV-powheg	11,1	497658	44,8
	DY1JetToLL_M-50_8TeV-madgraph	666,3	24045248	36,1
$Z \rightarrow ll + jets$	DY2JetsToLL_M-50_8TeV-madgraph	215,0	21852156	101,6
	DY3JetsToLL_M-50_8TeV-madgraph	60,7	11015445	181,4
	DY4JetsToLL_M-50_8TeV-madgraph	27,3	6402827	234,5
	TTWJets_8TeV-madgraph	0,232	196046	845,0
$t\bar{t}V$	TTZJets_8TeV-madgraph	0,206	210160	1021,7
	WW_8TeV_pythia6	54,8	10000431	182,5
	WZ_8TeV_pythia6	33,2	10000283	301,2
Dibosons	ZZ_8TeV_pythia6	8,1	979990	121,0

TABLE 4.2 – Les échantillons simulés du modèle standard utilisés avec leurs sections efficaces, le nombre d'événements générés et la luminosité équivalente.

lot de données (22Jan2013 rereco)	Trigger utilisé pour remplir le lot de données	$\mathcal{L}$ intégrée (fb)
/DoubleMu/Run2012A	HLT_Mu17_Mu8_v* ou HLT_Mu17_TkMu8_v*	0,876
/DoubleMuParked/Run2012B	HLT_Mu17_Mu8_v* ou HLT_Mu17_TkMu8_v*	4,407
/DoubleMuParked/Run2012C	HLT_Mu17_Mu8_v* ou HLT_Mu17_TkMu8_v*	7,016
/DoubleMuParked/Run2012D	HLT_Mu17_Mu8_v* ou HLT_Mu17_TkMu8_v*	7,369
<hr/>		
/MuHad/Run2012A	HLT_IsoMu17_eta2p1	
	+ _TriCentralPFJet30_v*	0,876
/SingleMu/Run2012B-TOPMuPlusJets	+ _TriCentralPFNoPUJet30_v*	ou
	+ _TriCentralPFNoPUJet30_30_20_v*	4,412
/SingleMu/Run2012C-TOPMuPlusJets	+ _TriCentralPFNoPUJet30_30_20_v*	ou
	+ _TriCentralPFNoPUJet45_35_25_v*	7,044
/SingleMu/Run2012D-TOPMuPlusJets	+ _TriCentralPFNoPUJet30_30_20_v*	ou
	+ _TriCentralPFNoPUJet45_35_25_v*	7,368
<hr/>		
/ElectronHad/Run2012A-22Jan2013	HLT_Ele25_CaloIdVT_CaloIsoT_TkIdT_TkIsoT	
	+ _TriCentralPFJet30_v* ou + _TriCentralPFNoPUJet30_v*	0,876
/SingleElectron/Run2012B-TOPElePlusJets	+ _TriCentralPFNoPUJet30_v*	ou
	+ _TriCentralPFNoPUJet30_30_20_v*	4,399
/SingleElectron/Run2012C-TOPElePlusJets	+ _TriCentralPFNoPUJet30_30_20_v*	ou
	+ _TriCentralPFNoPUJet45_35_25_v*	7,022
/SingleElectron/Run2012C-TOPElePlusJets	+ _TriCentralPFNoPUJet30_30_20_v*	ou
	+ _TriCentralPFNoPUJet45_35_25_v*	7,369

TABLE 4.3 – Les échantillons de données utilisés ainsi que les chemins de déclenchement et les luminosités intégrées correspondantes.

### **4.3. Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 93**

---

— respecté les critères de bonne qualité des jets.

Ces critères de base assurent le rejet d'un certain nombre d'événements avec des observables de mauvaise qualité qui pourraient détériorer nos accords entre les données et les événements simulés. Il faut aussi souligner que ces restrictions permettent déjà de réduire une partie des fonds du modèle standard. Par exemple les processus QCD dont la section efficace est très importante produisent très rarement des muons isolés (mais plutôt à l'intérieur de jet). Ils seront donc fortement supprimés par la coupure sur l'isolation des muons. De plus, le fait de demander des objets à grande impulsion transverse et dans la partie centrale du détecteur va permettre de sélectionner préférentiellement des collisions dures plus susceptibles de produire la nouvelle physique qui nous intéresse.

#### **4.3.1.3 Facteurs correctifs**

##### **Repondération du PU**

Comme les échantillons MC sont simulés avec une distribution de PU hypothétique (selon les prédictions au moment de la production de l'échantillon), on effectue une repondération en fonction de la vraie distribution de PU extraite des données. On pourrait repondérer le nombre de vertex reconstruits dans le MC à celui dans les données, mais on serait alors sensible aux différences dans les efficacités de reconstruction. On préfère donc repondérer directement le nombre d'interactions générées dans le MC (vérité MC) au nombre d'interactions réelles pour chaque croisement de faisceau. Ce nombre est extrait des données grâce à la luminosité instantanée par croisement de faisceau selon la formule :  $\mu = \sigma_{pp} \times \mathcal{L} / f_c$  (voir partie 2.1.4). On voit que la section efficace inélastique rentre en jeu dans ce calcul. Comme elle n'est pas connue parfaitement, il faudra en tenir compte pour le calcul des incertitudes. En la faisant varier de  $\pm\sigma$  on obtiendra les systématiques inhérentes à la repondération du PU. La figure 4.5 montre la distribution du nombre de vertex reconstruits après cette repondération pour une sélection d'événements comportant exactement deux muons.

On valide ainsi notre méthode de repondération du PU, en observant que la distribution du nombre de bons vertex reconstruits est bien en accord pour les données et la simulation.

##### **Repondération selon le $p_t$ du top**

Au cours de l'année 2013, on s'est rendu compte que le  $p_t$  du quark top n'était pas parfaitement simulé dans l'échantillon MADGRAPH officiel. En effet lors de la mesure de la section efficace différentielle  $t\bar{t}$  dans plusieurs canaux, des différences dans la forme des distributions entre les données et le MC sont apparues [98]. Le spectre en  $p_t$  des quarks tops générés était en effet plus énergétique que celui des données. C'est-à-dire que les quarks tops avaient été simulés avec trop de boost. Pour corriger ces différences, et comme il était impossible de reproduire l'ensemble

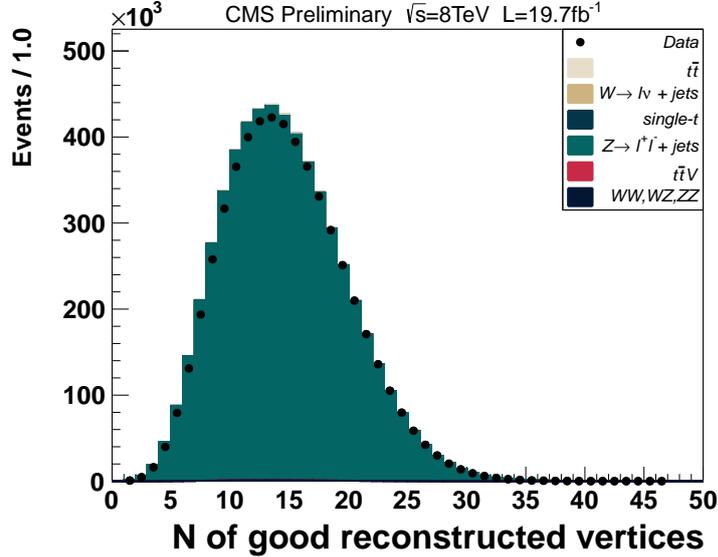


Fig. 4.5 – Distribution du nombre de vertex dans les données et la simulation après repondération pour une sélection dimuon.

des événements simulés  $t\bar{t}$ , il a été décidé de dériver des facteurs correctifs pour les événements MC à partir des données. Pour ce faire, on calcule le rapport de la section efficace différentielle dans les données et la simulation en fonction du  $p_t$  du top (voir [99]). Puis on ajuste ce rapport avec une fonction exponentielle. Il suffit alors d'inverser la formule pour obtenir l'accord données/MC pour le  $p_t$  du top. Au final, on obtient les équations 4.4 et 4.5 qui fournissent un poids global à appliquer aux événements MC  $t\bar{t}$  en fonction des  $p_t$  du quark top et de l'antiquark top au niveau généré.

$$\text{Facteur correctif} = \sqrt{f(p_T(t)) \times f(p_T(\bar{t}))} \quad (4.4)$$

$$\text{avec } f(x) = \exp(0.159 - 0.0141x) \quad (4.5)$$

### 4.3.2 Contrôle des fonds du modèle standard

Notre analyse basée sur une optimisation des coupures nécessite un accord données MC très fin. Pour s'en assurer, on va vérifier que la simulation reproduit bien les données dans un certain nombre de régions de contrôle. La base de notre sélection est de demander un bon muon isolé. Les processus principaux du MS contenant un muon isolé dans l'état final sont  $Z^0 \rightarrow \mu\bar{\mu}$ ,  $W^\pm \rightarrow l^\pm\nu_l$  et  $t\bar{t}$ . L'état final du signal contient principalement un bon lepton et de nombreux jets il faut donc s'assurer que les distributions les caractérisant sont bien modélisées dans le MC.

### 4.3. Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 95

#### 4.3.2.1 Sélection $Z^0 \rightarrow \mu^+ \mu^-$

L'idée principale qui sous tend cette première région de contrôle est de vérifier la qualité des muons. Pour cela nous allons nous appuyer sur une des chandelle du modèle standard, le boson Z. Elle permet par ailleurs de vérifier la qualité de la simulation du fond Z en dimuons. Nous demandons donc exactement deux muons dans l'état final. Pour être cohérent avec la simulation on demande en plus que la masse dimuons reconstruite soit supérieure à 50 GeV <sup>1</sup>. Cette coupure sera logiquement conservée dans toutes les sélections ultérieures (quand deux bons muons sont identifiés dans l'événement). Enfin on s'assure que les événements simulés auraient déclenché un des deux triggers dimuon utilisés pour former le lot de données DoubleMu. Il s'agit de HLT\_Mu17\_Mu8 ou de HLT\_Mu17\_TkMu8 qui requièrent un muon de 17 GeV et un muon de 8 GeV ou respectivement, un muon de 17 GeV et un tracker muon de 8 GeV.

La figure 4.6 (à gauche) présente la masse invariante des deux muons. Cette distribution est largement dominée par la résonance du boson Z autour de 91 GeV. Cependant, la figure de gauche montre un saut important dans le rapport données/MC autour de la masse du boson Z. Ceci vient du fait que la modélisation du détecteur est imparfaite. Par exemple, les problèmes d'alignement, la connaissance sur le champ magnétique ou la modélisation imparfaite de certaines parties du détecteur vont altérer la réponse et la résolution des muons dans les données par rapport à celle dans le MC. Il faut donc modifier légèrement l'impulsion des muons générés pour corriger ce décalage et reproduire au mieux les données. Un algorithme dédié, appelé MUSCLEFIT [100], a été développé spécialement pour régler ces divergences. Il utilise les résonances bien connue du MS ( $J/\psi$ , Y, Z) pour comparer et corriger les caractéristiques des muons dans la simulation afin de supprimer les différences observées par rapport aux données. On peut voir l'effet des ces corrections sur la figure 4.6 (à droite).

Vérifions maintenant la qualité des distributions caractérisant les muons dans les figures 4.7 et 4.8. La multiplicité en jets et leurs distributions en  $p_t$  pour cette sélection dimuon est présentée sur la figure 4.9. Enfin, le tableau 4.4 récapitule le nombre d'événements dans les données et la simulation pour cette sélection dimuon.

Que ce soit en terme de profil des distributions ou en terme de nombre d'événements, cette sélection dimuon nous a permis de contrôler la qualité de nos muons, de vérifier le bon fonctionnement de la repondération du PU ainsi que de valider le lot de données simulés  $Z \rightarrow \mu\mu$ . Les jets qui satisfont les critères de qualité du groupe top semblent, par ailleurs, tout à fait robustes.

#### 4.3.2.2 Sélection W + 4 jets

Nous allons maintenant étudier la deuxième région de contrôle  $W \rightarrow \mu\nu_\mu + 4jets$ . Maintenant que nous avons vérifié la qualité de nos muons nous allons pouvoir relâcher la coupure en ne demandant maintenant qu'un seul bon muon isolé. Pour

---

1. Cette coupure existant au niveau générateur.

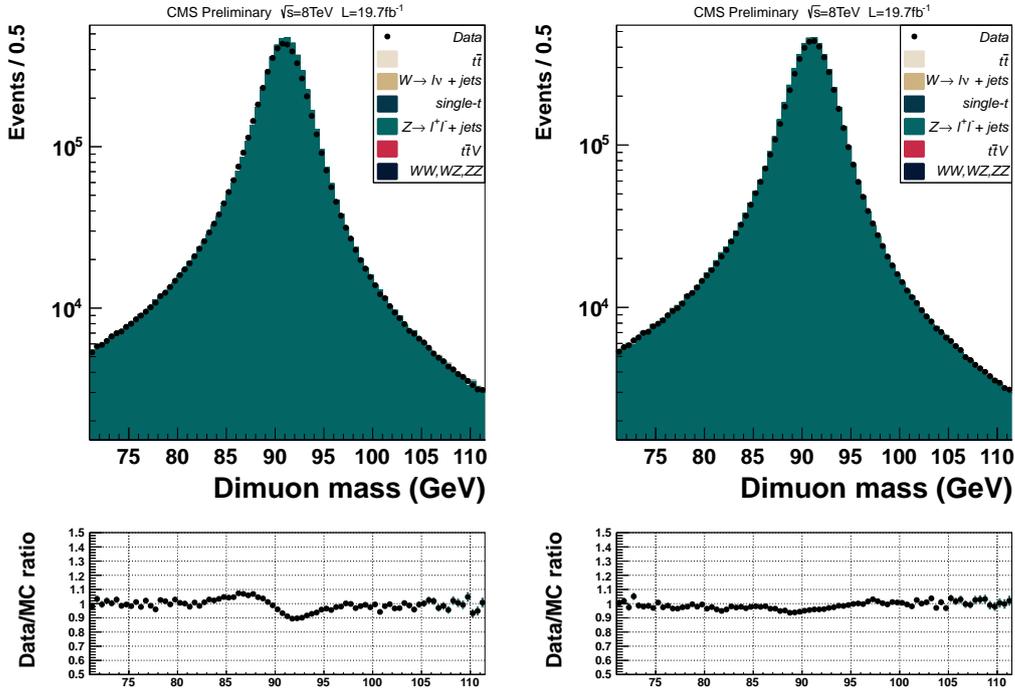


Fig. 4.6 – Masse sans correction MUSCLE fit (à gauche) et avec (à droite) du système dimuons.

Fonds	Nombre d'événements
Données	558472
Fonds du MS	$577830 \pm 3452$
Accord données/MC	$3.35 \% \pm 0.05\%$
$t\bar{t}$	$18539 \pm 30$
$t\bar{t}V$	$122 \pm 2$
$W \rightarrow l\nu + \text{jets}$	$74 \pm 31$
WW	$3424 \pm 19$
WZ	$6792 \pm 21$
ZZ	$3253 \pm 7$
Single t et $\bar{t}$	$1871 \pm 22$
$Z \rightarrow ll + \text{jets}$	$544422 \pm 3451$

TABLE 4.4 – Nombre d'événements et incertitudes statistiques associées qui passent la sélection dimuons.

se rapprocher de la topologie du signal, nous requerrons en plus 4 jets au minimum dans l'état final. Pour simuler le chemin de déclenchement suivi par les données dans le MC, les efficacités de déclenchement ont été mesurées par le groupe Top. Un module validé par ce même groupe permet de pondérer les événements MC par

### 4.3. Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 97

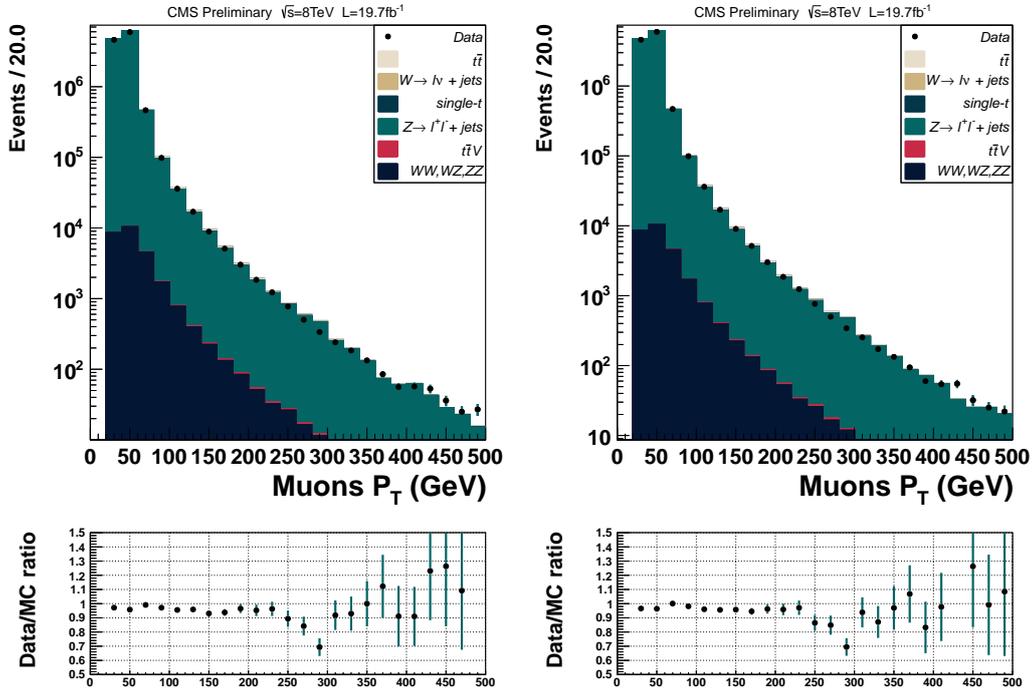


Fig. 4.7 – Impulsion transverse des muons sans correction MUSCLEFIT (à gauche) et avec (à droite).

les efficacités de trigger en fonction des caractéristiques du muon et du quatrième jet<sup>2</sup>. Pour bien utiliser ce module il faut se placer exactement au même niveau de sélection que celui qui a permis de déterminer ces poids. Nous continuons donc de suivre les recommandations du groupe top en demandant :

- au moins 4 bons jets dont l'impulsion transverse est supérieure à 30GeV ;
- que le  $p_t$  du jet le plus énergétique soit supérieur à 45 GeV ;
- que le  $p_t$  du deuxième jet le plus énergétique soit supérieur à 45 GeV ;
- que le  $p_t$  du troisième jet le plus énergétique soit supérieur à 35 GeV ;

Similairement à la sélection dimuons, une repondération du PU est effectuée en tirant cette fois la vraie distribution du nombre d'interactions du lot de données MuHad.

Cette région de contrôle a plusieurs objectifs :

- valider le fond  $W+$  jets pour des multiplicités en jets élevées ;
- s'assurer du bon fonctionnement du module fournissant les efficacités de trigger ;
- vérifier que le fond QCD (non simulé) est négligeable après la préselection ;
- confirmer que les distributions des observable globales ( $H_t$ ,  $\cancel{E}_T$ ,  $S_t$ ...) sont en accord pour les données et la simulation.

2. Le trigger utilisé pour ce lot de données demandant au moins 1 muon et 4 jets.

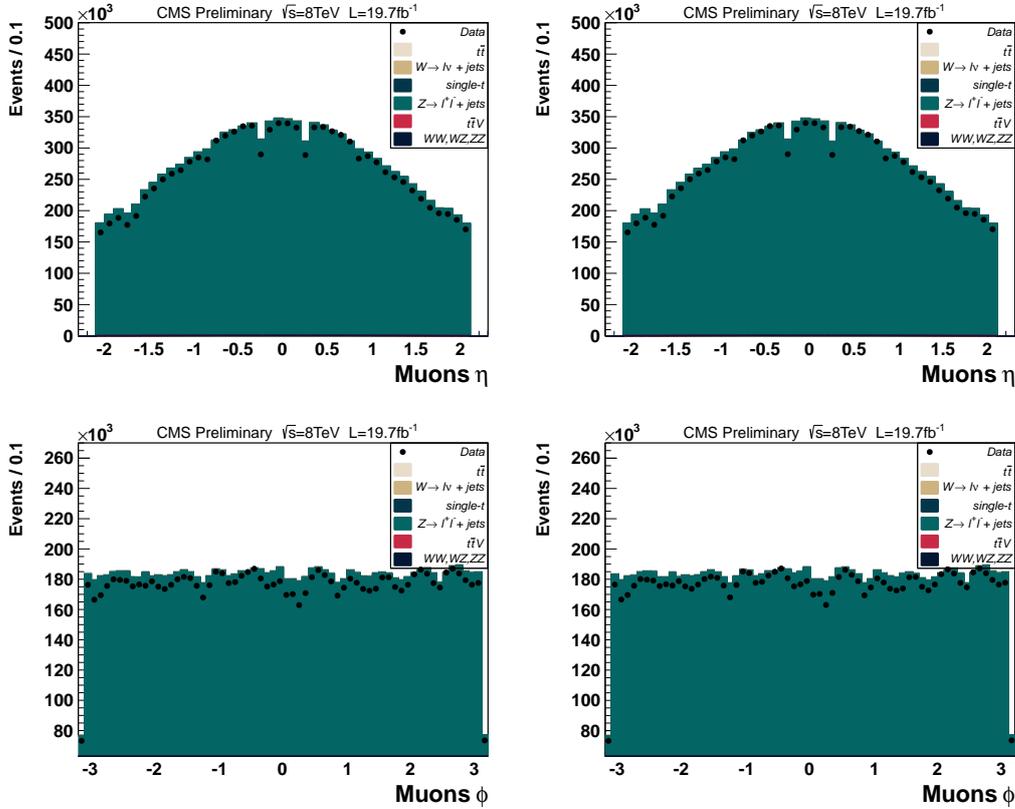


Fig. 4.8 –  $\eta$  (en haut) et  $\phi$  (à bas) des bons muons sans correction MUSCLEFIT (à gauche) et avec (à droite).

La première chose à contrôler dans cette sélection est que l'on reproduit correctement la masse transverse du boson W. Pour la calculer, on suppose que toute l'énergie transverse manquante est due au neutrino muonique provenant de la désintégration du boson W. On calcule la masse transverse à l'aide de la formule suivante :

$$M_t(W) = \sqrt{2 \times \cancel{E}_T \times P_t(\mu) \times (1 - \cos \Delta\phi(\cancel{E}_T, \mu))} \quad (4.6)$$

La figure 4.10 (à gauche), montre cette distribution de masse transverse. On voit qu'en dessous de 60 GeV il y a un écart important entre données et MC. On peut même reconnaître la décroissance caractéristique de la QCD dans les données. Tous ces événements QCD présents dans les données mais non pris en compte dans la simulation viennent biaiser les distributions comme on peut le voir dans la figure 4.11 (à gauche). Il s'agit de la somme scalaire des impulsions transverse du muon et des bons jets :

$$S_T = p_t^\mu + \sum_{jets} p_t^j \quad (4.7)$$

Nous le verrons par la suite,  $S_T$  est une variable très importante pour notre analyse. En effet, une sélection des événements avec une grande valeur de  $S_T$  permet de

### 4.3. Prérequis pour l'analyse des données de CMS à 8TeV Run 2012 99

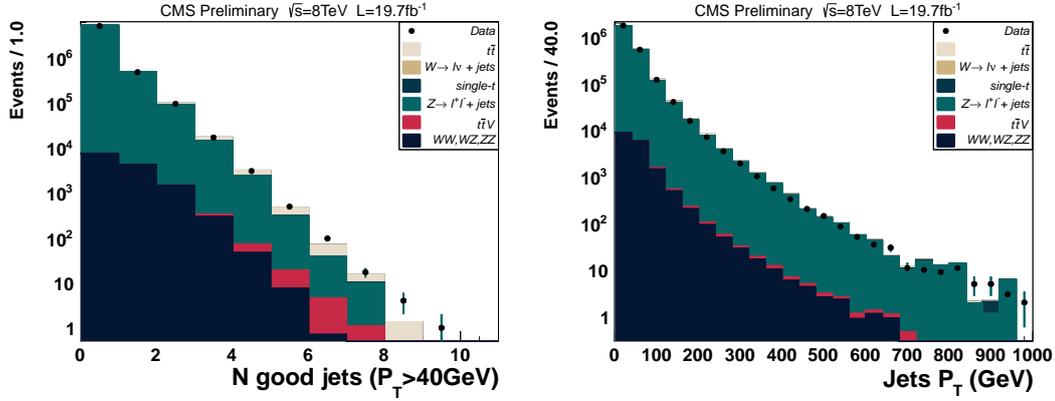


Fig. 4.9 – Nombre de jets de plus de 40 GeV (à gauche) et impulsion transverse de tous les bons jets (à droite).

supprimer une majeure partie du fond du MS tout en conservant une bonne efficacité de signal. On élimine la contribution des événements QCD en rejetant tous les événements avec une masse transverse du W inférieure à 60 GeV (voir les figures 4.10 et 4.11 (à droite)). Il faut noter qu'on effectue cette coupure uniquement s'il y a exactement un bon muon dans l'événement. Ce qui explique le reliquat d'événements à basse masse transverse quand il n'y a pas exactement un bon muon. Nous avons du mettre en place cette spécificité car une coupure générale sur la masse transverse amputerait le signal d'un trop grand nombre d'événements et faisait chuter dramatiquement la signficance au niveau de sélection optimal. Sauf contre indication, on utilisera toujours cette coupure à 60 GeV quand il n'y a qu'un seul bon muon identifié dans l'événement.

D'autre part, il existe 2 échantillons d'événements simulés avec MADGRAPH W+jets et Z+jets, l'un inclusif et l'autre par multiplicité en jets générés au niveau partonique. Comme notre sélection présente un grand nombre de jets énergétiques, on peut s'attendre à ce que les échantillons exclusifs donnent un meilleur accord données/MC car il contiennent proportionnellement plus d'événements générés avec 4 jets. C'est ce que nous vérifions sur la figure 4.12 qui présente l'impulsion transverse des muons et la  $\cancel{E}_T$  pour deux échantillons. Même si les nombres d'événements prédits par les deux échantillons sont sensiblement les mêmes (voir tableau 4.5), on voit que les échantillons inclusifs simulent moins bien les queues de distributions qui sont les zones de l'espace des phases qui correspondent à notre signal. Alors que pour les échantillons exclusifs, la majorité des événements provient de l'échantillon à 4 jets ce qui décrit donc beaucoup mieux la physique des données. Il s'agit uniquement d'un effet statistique, comme la section efficace de production d'un W et de 4 jets et assez faible, il faudrait produire une énorme quantité d'événements pour l'échantillon inclusif pour parvenir au même accord données/MC dans les queues de distribution. Dans toute la suite de cette étude, les échantillons W+jets et Z+jets exclusifs seront donc utilisés.

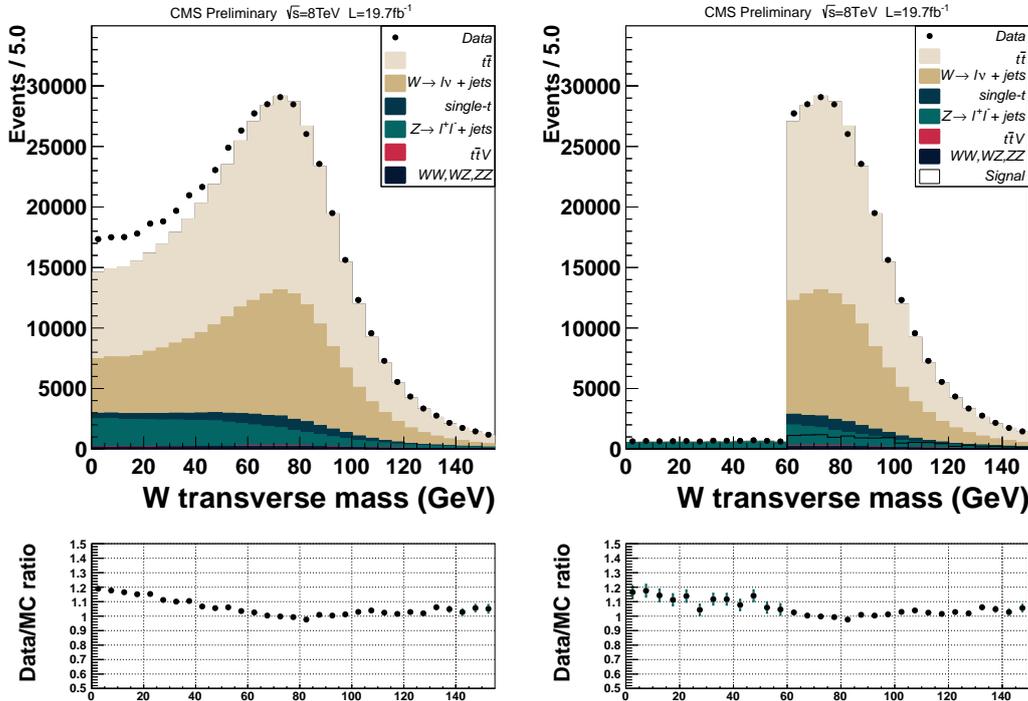


Fig. 4.10 – Distribution de la masse transverse du W avant coupure (à gauche) et après (à droite).

Il reste à vérifier que la procédure de repondération des événements pour simuler l'effet du trigger dans la simulation donne des résultats en accord avec les données. A cet effet, la figure 4.13 présente les caractéristiques des objets (muon et 4<sup>ème</sup> jet le plus énergétique) utilisés par le module qui fournit l'efficacité de déclenchement. Ces distributions montrent que l'application des coefficients d'efficacité de déclenchement permet d'obtenir un bon accord données/MC sans utiliser la simulation du trigger pour le MC. Par ailleurs il est important de s'assurer que les distributions des observables reconstruites de l'analyse sont en accord entre les données et la simulation. La figure 4.14 présente les distributions des observables sur lesquelles nous ferons les coupures d'optimisation dans la suite de l'analyse. On vérifie qu'elles sont bien simulées ce qui donne confiance quant à l'utilisation de ces variables pour l'optimisation des coupures.

## 4.4 Recherche de Leptoquarks de 3<sup>ème</sup> génération se désintégrant en top $\tau$

### 4.4.1 État final et canal d'analyse

Dans cette partie nous considérerons uniquement la désintégration du LQ scalaire de 3<sup>ème</sup> génération en quark top et lepton  $\tau$ . Les désintégrations en quark top  $\nu$  et

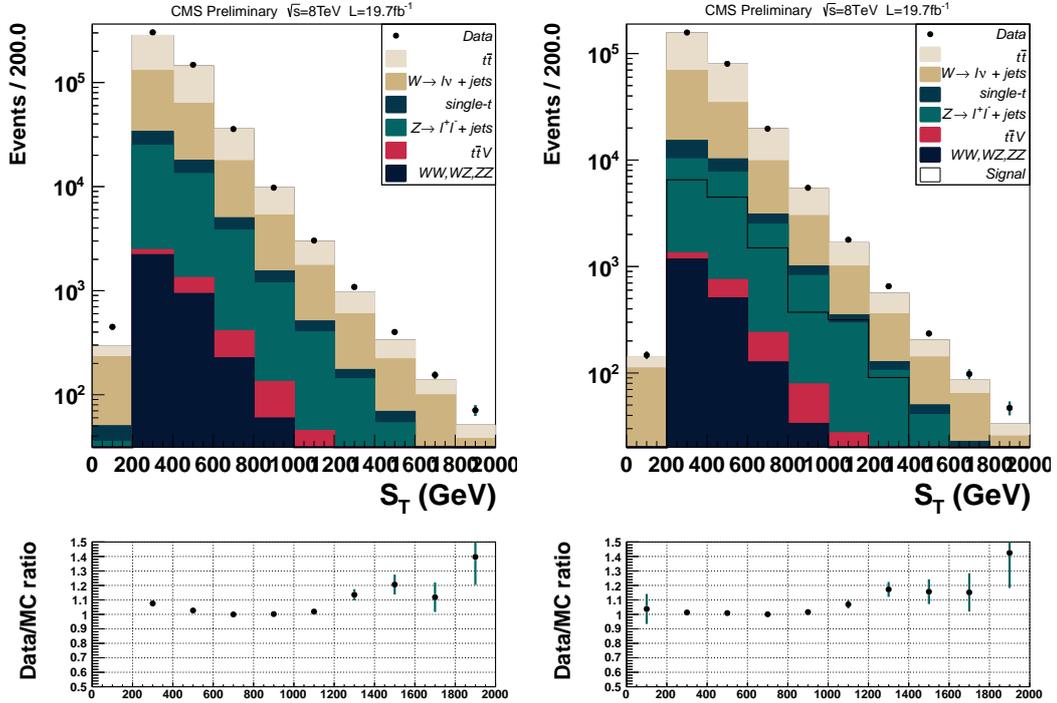


Fig. 4.11 – Distribution de la somme des impulsions transverses des muons et des jets avant coupure en masse transverse du candidat W à 60 GeV (à gauche) et après (à droite). Le signal correspond à des LQ de 300 GeV

en quark top  $\mu$  seront envisagées dans un deuxième temps.

Chaque LQ se désintègre en un lepton et un quark top ce qui mène à 3 états finals différents  $\tau\tau + jets$ ,  $\tau\nu_\tau + jets$  et  $\nu_\tau\nu_\tau + jets$ . Ces 3 états finals apparaissent avec des taux proportionnels respectivement  $\beta^2$ ,  $\beta(1 - \beta)$  et  $(1 - \beta)^2$ . Cependant dans cette partie nous nous focalisons sur la recherche avec deux  $\tau$  dans l'état final en fixant  $\beta = 1$  (premier canal ouvert uniquement). Notre état final est donc exclusivement  $LQLQ \rightarrow t\bar{t}\tau^+\tau^-$ . Les quarks tops se désintégrant presque à 100% en  $bW$ , l'état final pour la désintégration en top  $\tau$  est  $LQLQ \rightarrow b\bar{b}W^+W^-\tau^+\tau^-$ . Il faut ensuite prendre en compte la désintégration des leptons  $\tau$  et des bosons W. Les bosons W ont 32.1% de chance de se désintégrer en  $l_i\nu_i$  ( $i \in \{e, \mu, \tau\}$ ) et 67,6% en  $q\bar{q}$  [12]. Le tau est le seul lepton qui peut se désintégrer en hadrons. De plus, compte tenu de la conservation du nombre leptonique, sa désintégration contient toujours un neutrino tauique ce qui se traduira par de l'énergie transverse manquante. Le rapport de branchement du  $\tau$  en  $\nu_\tau l_i \nu_i$  ( $i \in \{e, \mu\}$ ) est de 35,2% [12]. Les 64,8% restant représentent les désintégrations hadroniques notamment par l'intermédiaire des mésons  $\pi^\pm$  et  $\pi^0$ . Dans tous les cas, à cause de la masse élevée des LQ, les leptons et les jets produits par la désintégration auront une impulsion transverse très importante.

Dans un collisionneur, la désintégration des LQ en top  $\tau$  est donc caractérisée par

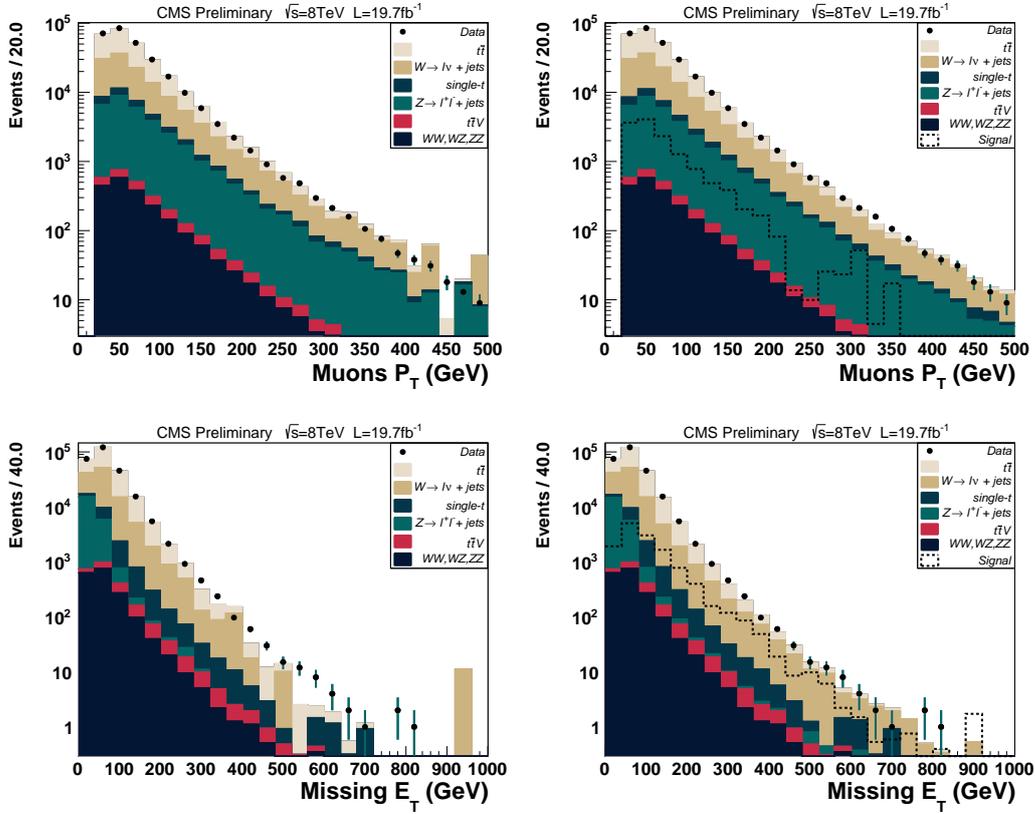


Fig. 4.12 – Impulsion transverse des muons (en haut) et  $\cancel{E}_T$  (en bas) avec les échantillons W+jets et Z+jets inclusif (à gauche) et exclusif (à droite).

des jets dont au moins 2 b jets, des leptons très énergétiques et une  $\cancel{E}_T$  conséquente.

#### 4.4.2 Pré-sélection des événements

Même si les accords données MC de la sélection précédente étaient très bons, le bruit de fond du MS domine encore largement le signal des LQ attendu. En physique des particules dans le cas d'un nombre d'événements de signal faible, on utilise souvent le degrés de signficance (équation (4.8)) pour optimiser cette proportion de fond par rapport au signal.

$$S = \frac{n_s}{\sqrt{n_s + n_b}} \quad (4.8)$$

Le but est donc de trouver une pré-sélection avant l'optimisation finale avec le meilleur S possible tout en conservant assez d'événements et un bon accord avec les données. Comme nous l'avons vu dans la partie 4.4.1, il y a au moins 2 jets de b dans l'état final. Il est donc naturel de penser à l'étiquetage de jet b pour augmenter S. Après l'essai de plusieurs coupures, nous avons déterminé le niveau de pré-sélection optimal en terme d'accord données/MC et de signficance. Cette sélection correspond

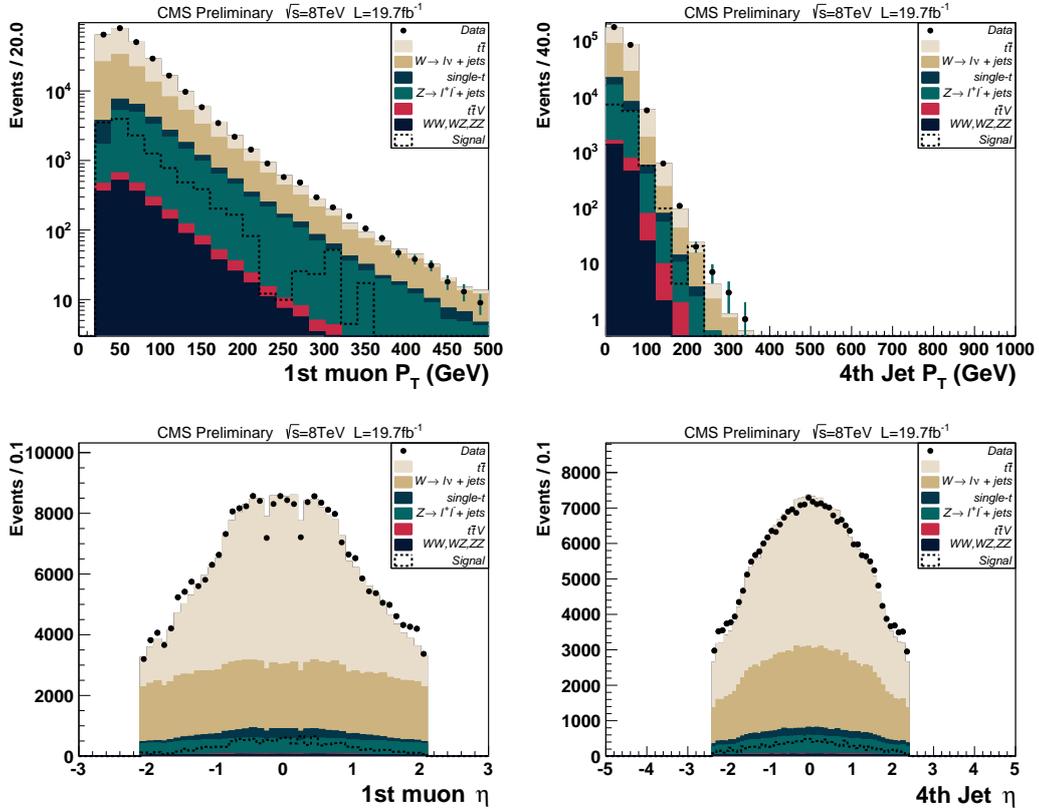


Fig. 4.13 – Caractéristiques ( $P_t$  en haut et  $\eta$  en bas) du muon (à gauche) et du quatrième jet (à droite) utilisés pour calculer l'efficacité du trigger.

à au moins un bon muon isolé, une coupure sur la masse transverse du boson W de 60 GeV, 4 jets de plus de 40 GeV correspondant à la sélection top (voir partie 4.3.2.2) et au moins un jet étiqueté comme b-jet avec la qualité medium.

A partir de là, il faut déterminer les distributions cinématiques qui vont permettre de séparer au mieux le fond du signal. Pour ce faire, on regarde uniquement la forme du fond et du signal et on cherche des disparités. L'idéal est de trouver une distribution où le signal se sépare clairement du fond. Après avoir testé de nombreuses possibilités, trois distributions apparaissent comme les plus discriminantes. Il s'agit du nombre de jets de plus de 40 GeV, de  $S_t$  et de  $\cancel{E}_T$ . Ces trois observables sont présentées au niveau de la pré-sélection optimale avec un signal correspondant à  $m_{LQ} = 400\text{GeV}$  sur la figure 4.15. Comme on peut le constater l'accord des données avec le MC est bon pour ces 3 quantités (ce qui va nous permettre d'optimiser des coupures dessus et d'obtenir des limites cohérentes). Il est logique que ces trois quantités se détachent, en effet avec des LQ massifs on s'attend à plus d'activité dans le détecteur, donc plus de jets énergétiques et une valeur de  $S_t$  plus importante. De plus les neutrinos tauiques associés à la désintégration des LQ de troisième génération emportent une grande part d'énergie non détectée ce qui occasionne plus

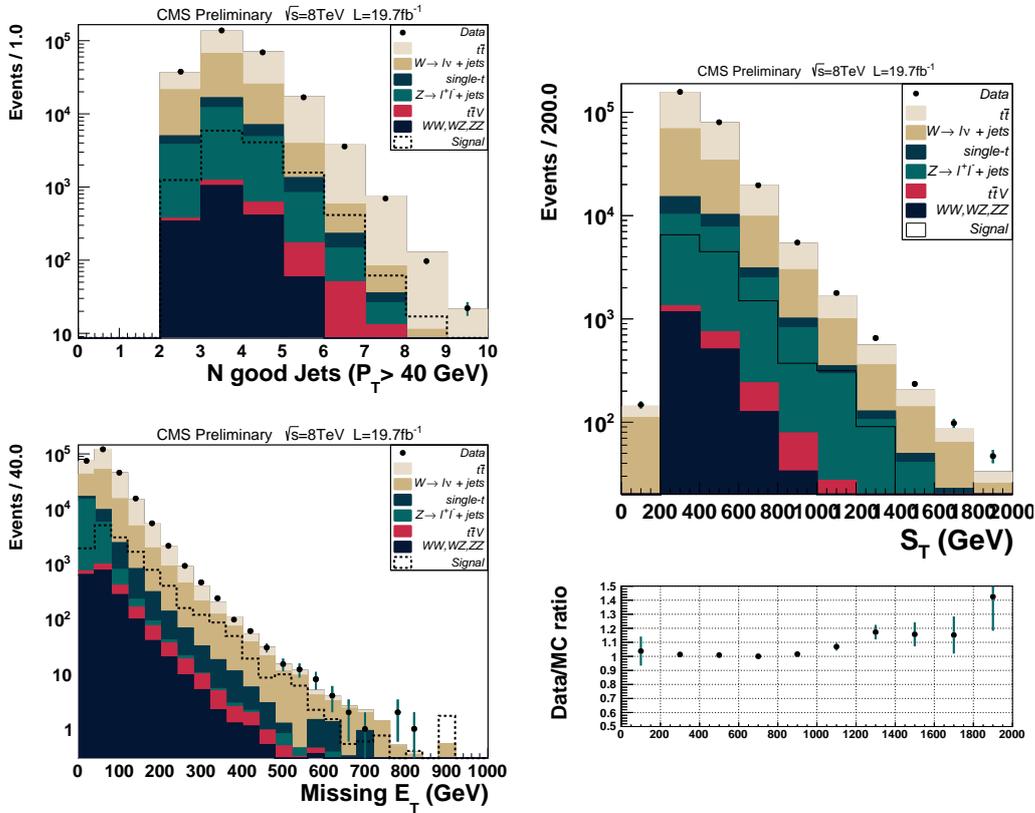


Fig. 4.14 – Multiplicité en jets de  $p_t > 40\text{GeV}$  (en haut à gauche)  $E_T^{\text{Missing}}$  (en bas à gauche) et  $S_T$  (à droite).

de  $E_T^{\text{Missing}}$  que pour le bruit de fond  $t\bar{t}$ .

### 4.4.3 Les incertitudes statistiques

Il est très important de déterminer les incertitudes avec la plus grande précision car elles rentrent en compte dans le calcul des niveaux de confiance. Plus les incertitudes seront basses, meilleures seront les valeurs de  $Cl_s$  obtenues et donc meilleures seront les limites. Les incertitudes statistiques proviennent de la taille limitée des échantillons MC générés. Comme nous l'avons dit dans la partie 3.5, à cause de la limitation en terme de ressources et de l'augmentation de la luminosité équivalente on ne peut générer qu'un nombre limité d'événements MC. Pour en tenir compte on calcul l'incertitude statistique qui dépend de l'efficacité de sélection. Il s'agit du quotient défini par le rapport entre événements sélectionnés et événements générés (équation (4.9)). L'équation (4.10) rappelle la formule utilisée pour calculer l'incertitude statistique dans le cas d'une loi binomiale. La coupure définie deux populations distinctes (les événements qui la passent et les autres) ce qui revient au cas d'une loi binomiale avec une probabilité  $p$  de succès et  $1-p$  d'échec.

Fonds	Nombre d'événement
Données	265505
Fonds du MS	262332 ± 244
Accord données/MC	1.19 % ± 0.09%
$t\bar{t}$	144974 ± 133
$t\bar{t}V$	587 ± 5
$W \rightarrow l\nu + 1\text{jet}$	688 ± 62
$W \rightarrow l\nu + 2\text{jets}$	1460 ± 42
$W \rightarrow l\nu + 3\text{jets}$	10765 ± 93
$W \rightarrow l\nu + 4\text{jets}$	74454 ± 169
$W \rightarrow l\nu + \text{jets excl}$	87368 ± 366
$W \rightarrow l\nu + \text{jets incl}$	87066.8 ± 1056
WW	1093 ± 11
WZ	629 ± 6
ZZ	126 ± 1
Single t	5512 ± 62
Single $\bar{t}$	3854 ± 56
$Z \rightarrow ll + 1\text{jet}$	141 ± 9
$Z \rightarrow ll + 2\text{jets}$	610 ± 11
$Z \rightarrow ll + 3\text{jets}$	2965 ± 18
$Z \rightarrow ll + 4\text{jets}$	15472 ± 36
$Z \rightarrow ll + \text{jets excl}$	19188 ± 74
$Z \rightarrow ll + \text{jets incl}$	20025 ± 213

TABLE 4.5 – Nombre d'événements et erreurs statistiques associées qui passent la sélection W+4jets  $M_T(W) \geq 60$ .

$$N = \varepsilon \mathcal{L} \sigma \text{ avec } \varepsilon = \frac{N_{sel}}{N_{gen}} \quad (4.9)$$

$$\Delta N_{stat} = (\mathcal{L} \sigma) \Delta \varepsilon \rightarrow \Delta N_{stat} = (\mathcal{L} \sigma) \sqrt{\frac{\varepsilon(1-\varepsilon)}{N_{gen}}} \quad (4.10)$$

Les incertitudes statistiques provenant de chaque processus MS simulé sont ensuite sommées quadratiquement pour obtenir l'incertitude statistique totale. Les contributions des incertitudes statistiques pour chaque valeur de masse de LQ au niveau de sélection optimale peuvent être trouvées dans les tableaux 4.7 et 4.8.

#### 4.4.4 Les incertitudes systématiques

Contrairement à l'erreur statistique globale qui dépend uniquement du nombre d'événements générés, les incertitudes systématiques proviennent d'une erreur faite sur chaque mesure. Elles peuvent être de deux types : théoriques ou expérimentales.

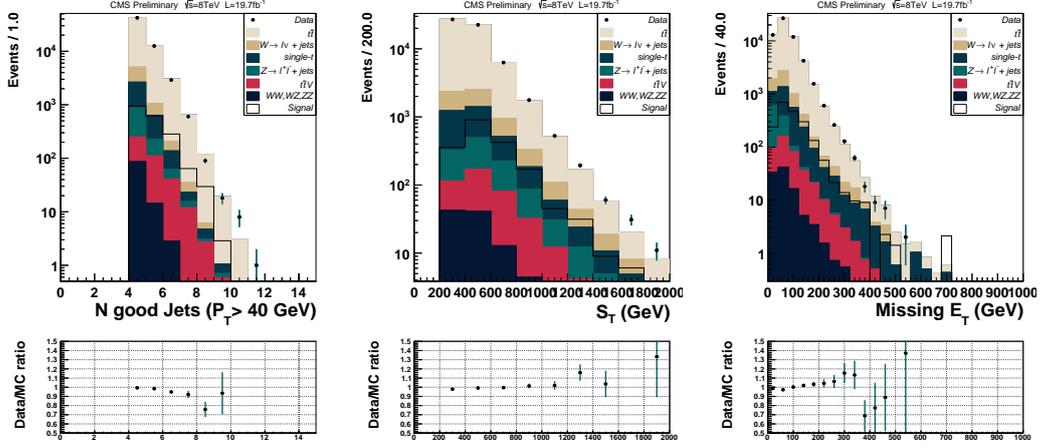


Fig. 4.15 – Multiplicité en jets de  $p_t > 40\text{GeV}$  (à gauche)  $S_t$  (au milieu) et  $\cancel{E}_T$  (à droite) pour la pré-sélection optimale. Le signal correspond ici à une masse de LQ de 400 GeV et une désintégration à 100% en top  $\tau$ .

Processus du MS	$\sigma_{NLO}(pb)$	$\Delta\sigma_{NLO}$
$t\bar{t}$	252,9	$\pm 14,5$
$t\bar{t}W$	0,23	$\pm 0,07$
$t\bar{t}Z$	0,21	$\pm 0,02$
$W \rightarrow l\nu$	37510	$\pm 1970$
WW	54,8	$\pm 1,9$
WZ	33,2	$\pm 1,3$
ZZ	17,7	$\pm 0,5$
Single t voie s	3,79	$\pm 0,15$
Single t voie t	56,4	$\pm 2,4$
Single t voie t-W	11,1	$\pm 0,8$
Single $\bar{t}$ voie s	1,76	$\pm 0,08$
Single $\bar{t}$ voie t	30,7	$\pm 1,3$
Single $\bar{t}$ voie t-W	11,1	$\pm 0,8$
$Z \rightarrow ll$	3504	$\pm 151$

TABLE 4.6 – Sections efficaces et incertitudes théoriques associées [101] (et [102] quand les incertitudes à 8 TeV n'étaient pas disponibles).

#### 4.4.4.1 Théoriques

Les incertitudes théoriques sur les sections efficaces de production des bruit de fond du modèle standard ont été prises en compte. Le tableau 4.6 récapitule les sections efficaces utilisées et les incertitudes associées.

#### 4.4.4.2 Expérimentales

**Incertitudes sur l'échelle d'énergie des jets (JES) et la résolution des jets (JER)** Conformément aux recommandations du groupe top, on fait varier la JES et la JER entre plus ou moins un écart standard ( $\sigma$ ) dans le MC. On compare ensuite pour chaque niveau de sélection variation du nombre d'événements par rapport à l'échantillon central. Le résultat de ces variations est résumé dans le tableau 4.7. On peut voir que les incertitudes liées à l'échelle d'énergie sont les plus importantes de toutes les incertitudes expérimentales (de l'ordre de 6 % pour les fonds du MS). Les incertitudes prenant en compte les différences de résolution en énergie entre les données et le MC viennent en deuxième (de l'ordre de 2 % pour les fonds du MS).

**Autres incertitudes** L'incertitude systématique globale sur le calcul de la luminosité dans CMS avec la méthode pixel pour la période de prise de donnée 2012 est de 2,6% [103]. L'incertitude systématique liée à l'utilisation du module donnant l'efficacité de trigger a aussi été prise en compte. Le module fournit les variations haute et basse sur le coefficient. Il suffit donc d'observer les répercussions de ces variations sur les nombre d'événements attendus aux différents niveaux de sélection pour estimer l'incertitude associée (voir tableau 4.7).

#### 4.4.5 Optimisation finale

Aucun excès significatif d'événements dans les données n'apparaît par rapport à la simulation (fond MS). On interprète alors les résultats dans une optique d'exclusion plutôt que de découverte. Pour chaque valeur de masse des LQ, on réalise une grille de coupures pour les trois distributions les plus discriminantes. On calcule la limite attendue pour chacun des points de la grille. Puis on sélectionne les coupures qui maximisent cette limite attendue pour chaque valeur de masse.

##### 4.4.5.1 Niveaux de confiance Cls

Pour quantifier le fait qu'un excès d'événements ne soit pas dû à des fluctuations, on mesure ce qu'on appelle un niveau de confiance ou  $CL_s$ . On utilisera dans ce travail une version modifiée [104, 105] de l'approche fréquentiste classique des intervalles de confiance. On définit un test statistique  $X$  qui varie de façon monotone avec la présence de signal. Soit  $X_{obs}$  la valeur de ce test observée expérimentalement. Dans l'approche fréquentiste classique, on définit respectivement les niveaux de confiance  $CL_{s+b}$  et  $CL_b$  comme les probabilités que  $X$  soit inférieur à  $X_{obs}$  selon les hypothèses signal plus fond ou fond uniquement (voir les équations (4.11) et (4.12)).

$$CL_{s+b} = P_{s+b}(X \leq X_{obs}) = \int_{-\infty}^{X_{obs}} \frac{dP_{s+b}}{dX} dX \quad (4.11)$$

$$CL_b = P_b(X \leq X_{obs}) = \int_{-\infty}^{X_{obs}} \frac{dP_b}{dX} dX \quad (4.12)$$

type	mLQ	stat	JES		JER		trig eff		PU		X sec
MC	200	$\pm 1,1$	-6,1	+6,5	-2,0	+2,1	-1,5	+1,6	-0,1	+0,1	$\pm 5,5$
SIG	200	$\pm 25,5$	-17,4	+0,0	-7,1	+0,0	-1,4	+1,4	-8,8	+8,4	$\pm 0$
MC	300	$\pm 1,6$	-6,0	+6,1	-2,7	+2,3	-1,5	+1,6	-0,0	+0,1	$\pm 5,5$
SIG	300	$\pm 11,9$	-5,0	+7,9	-3,9	+1,4	-1,4	+1,4	-1,8	+2,1	$\pm 0$
MC	400	$\pm 2,1$	-6,0	+6,5	-1,5	+1,9	-1,5	+1,7	-0,0	+0,1	$\pm 5,5$
SIG	400	$\pm 8,0$	-5,0	+3,9	-0,7	+4,0	-1,4	+1,4	-4,0	+3,6	$\pm 0$
MC	500	$\pm 3,0$	-6,4	+4,7	-1,8	+0,9	-1,5	+1,4	-0,1	+0,3	$\pm 5,5$
SIG	500	$\pm 6,8$	-3,7	+1,4	-0,8	+0,3	-1,4	+1,4	-0,9	+1,1	$\pm 0$
MC	600	$\pm 3,8$	-5,2	+5,1	-0,6	+0,6	-1,5	+1,4	-0,2	+0,5	$\pm 5,5$
SIG	600	$\pm 5,0$	-4,5	+3,9	-1,2	+1,4	-1,4	+1,4	-1,0	+1,3	$\pm 0$
MC	700	$\pm 4,9$	-6,6	+4,9	-1,4	+1,1	-1,5	+2,1	-0,0	+0,3	$\pm 5,5$
SIG	700	$\pm 4,1$	-3,3	+3,2	-0,3	+0,2	-1,4	+1,5	-0,1	+0,1	$\pm 0$
MC	800	$\pm 4,9$	-6,6	+4,9	-1,4	+1,1	-1,5	+2,1	-0,0	+0,3	$\pm 5,5$
SIG	800	$\pm 3,2$	-3,5	+3,7	-0,8	+0,5	-1,4	+1,5	-1,4	+1,4	$\pm 0$
MC	900	$\pm 13,0$	-8,2	+9,2	-4,5	+2,7	-1,5	+7,4	-1,7	+1,6	$\pm 5,4$
SIG	900	$c \pm 4,1$	-4,3	+1,4	-2,2	+0,5	-1,4	+1,4	-1,0	+0,6	$\pm 0$
MC	1000	$\pm 13,0$	-8,2	+9,2	-4,5	+2,7	-1,5	+7,4	-1,7	+1,6	$\pm 5,4$
SIG	1000	$\pm 3,5$	-3,1	+4,1	-0,7	+0,2	-1,4	+1,4	-1,2	+1,2	$\pm 0$

TABLE 4.7 – Incertitudes en % pour les différentes masses de LQ au niveau de sélection optimal.

Une faible valeur du  $CL_{s+b}$  favorise l'hypothèse du fond uniquement par rapport à l'hypothèse signal + fond. En effet, elle correspond à une faible probabilité dans le cas signal plus fond que le test statistique  $X$  soit inférieur à la valeur observée. Ainsi, si  $CL_{s+b} < 0.05$ , le signal sera exclu à 95% de niveau de confiance. Inversement, des valeurs de  $CL_b$  très proches de 1 permettent de mettre en doute l'hypothèse de fond uniquement. Il exprime la probabilité que les événements de fond seulement donnent une valeur de  $X$  inférieure à celle observée. Il peut donc être utilisé pour quantifier la confiance sur une potentielle découverte.

Cette approche fréquentiste classique présente cependant un problème dans le cas d'une faible sensibilité de l'expérience et d'une fluctuation négative du fond. Dans ce cas là, même si on attend peu ou pas de signal, l'hypothèse signal plus fond peut être exclue à 95% à cause de la fluctuation du fond. Pour éviter ce phénomène, on va normaliser  $CL_{s+b}$  par  $CL_b$ . On définit alors l'approche fréquentiste modifiée par le nouveau niveau de confiance  $CL_s$  comme suit :

$$CL_s = \frac{CL_{s+b}}{CL_b} \quad (4.13)$$

Avec cette nouvelle valeur, si l'expérience est peu sensible au signal et que le fond subit une fluctuation négative, même si le  $CL_{s+b}$  devient très petit, le  $CL_s$  ne permettra pas l'exclusion du modèle car on aura aussi  $CL_b \ll 1$ . On considère que l'hypothèse signal plus fond est exclue à 95% si  $CL_s \leq 0.05$ . Dans notre analyse, on exclut jusqu'à la masse  $m_{LQ}^{lim}$  la plus petite masse telle que  $(CL_s(m_{LQ}) \geq 0.05)$

Il reste maintenant à définir le test statistique  $X$  adéquat pour notre analyse. De nombreux tests peuvent être utilisés comme le  $\chi^2$  ou la méthode des moindres carrés mais un seul permet d'optimiser le calcul de  $CL_s$ , il s'agit du rapport des vraisemblances. En effet, d'après le théorème de Neyman-Pearson [106], le choix du rapport des vraisemblances est le meilleur dans le cadre d'hypothèses complètement spécifiées et quand les deux hypothèses présentent la même densité de probabilité.

#### 4.4.5.2 Rapport de vraisemblance

Le rapport de vraisemblance est un test statistique qui correspond au résultat d'une expérience dans deux hypothèses différentes. Soit les données sont constituées du fond seul, soit elles contiennent du signal et du bruit de fond. Soient  $s$  et  $b$  le nombre d'événements attendus pour le signal et le fond et  $n$  celui dans les données. Le rapport de vraisemblance s'écrit comme le rapport des densités de probabilités correspondant aux deux hypothèses (équation (4.14)).

$$Q = \frac{\mathcal{L}(n, s + b)}{\mathcal{L}(n, b)} \quad (4.14)$$

avec  $\mathcal{L}(n, x) = \frac{e^{-x}}{n!} x^n$  la probabilité de Poisson d'observer  $n$  événements quand une valeur moyenne de  $x$  est attendue. Pour faciliter les calculs, on définit le logarithme du rapport de vraisemblance comme suit :

$$\ln(Q) = n \ln\left(1 + \frac{s}{b}\right) - s \quad (4.15)$$

Si l'expérience de comptage porte sur  $N$  variables aléatoires indépendantes  $i$ , la formule du rapport de vraisemblance se généralise à l'équation (4.16)

$$Q = \prod_{i=1}^N \frac{\mathcal{L}(n_i, s_i + b_i)}{\mathcal{L}(n_i, b_i)} \quad (4.16)$$

Les  $N$  variables indépendantes peuvent être soit différentes observables soit différents intervalles pour la même observable. Le rapport de vraisemblance logarithmique devient alors :

$$\ln(Q) = \sum_{i=1}^N n_i \ln\left(1 + \frac{s_i}{b_i}\right) - s_i \quad (4.17)$$

Dans tous les cas, le test statistique utilisé pour déterminer les niveaux de confiance est  $-2 \ln(Q)$ .

#### 4.4.5.3 Optimisation des coupures

Pour chaque valeur de masse de LQ on va donc déterminer les coupures en  $\cancel{E}_T$ ,  $S_T$  et  $N_{jets}$  qui donnent la meilleure limite attendue sur la section efficace de production. On fait varier  $\cancel{E}_T$  par pas de 25 GeV,  $S_t$  par pas de 100 GeV et on considérera les multiplicités en jets suivantes : plus de 4 jets de 40 GeV, plus de 5 jets de 40 GeV et plus de 6 jets de 40 GeV. Les résultats de cette optimisation sont résumés dans le tableau 4.8. On y présente, pour chaque valeur de masse, la meilleure coupure trouvée ainsi que les nombres d'événements qui correspondent à ce niveau de sélection.

Pour vérifier la qualité de nos accords données/MC à ce niveau de sélection, on vérifie certaines distributions représentatives. Les figures 4.16 montrent que pour les coupures optimales correspondant à une masse de LQ de 300 GeV, les accords sont bons même avec peu d'événements.

#### 4.4.6 Masse limite et interprétation des résultats

La figure 4.17 et le tableau 4.9 montrent les limites attendues et observées obtenues après optimisation ainsi que la section efficace de production des LQ avec ses incertitudes. La limite supérieure sur la masse des LQ pour cette analyse top  $\tau$  est obtenue en faisant l'intersection de la limite de la section efficace observée à 95 % par la courbe inférieure de la section efficace théorique. Ainsi dans le cadre de ce modèle, on peut exclure à 95 % de confiance des LQ de masse inférieure à 350 GeV.

### 4.5 Recherche de Leptoquarks de 3 ème génération se désintégrant en top $\nu$

#### 4.5.1 État final et canal d'analyse

Ici, on considère le rapport de branchement des LQ à 100 % en quark top et neutrino (cas  $\beta = 0$ ). Notre état final est donc exclusivement  $LQLQ \rightarrow t\bar{\nu}\nu$ . La

$M_{LQ}$	$N_{jets}^{40GeV}$	$S_t(GeV)$	$\cancel{E}_T$ (GeV)	$N_{signal}$	$\Delta N_{signal}^{stat}$	$\Delta N_{signal}^{syst}$	$N_{fond}$	$\Delta N_{fond}^{stat}$	$\Delta N_{fond}^{syst}$	$N_{donnees}^{ev}$
200	$\geq 4$	$\geq 700$	$\geq 125$	210,4	$\pm 53,7$	$\pm 55,0$	1076,4	$\pm 11,4$	$\pm 117,1$	1024
300	$\geq 4$	$\geq 900$	$\geq 100$	104,4	$\pm 12,5$	$\pm 11,9$	520,7	$\pm 8,1$	$\pm 53,7$	505
400	$\geq 4$	$\geq 900$	$\geq 150$	41,5	$\pm 3,3$	$\pm 3,4$	281,5	$\pm 6,0$	$\pm 30,6$	278
500	$\geq 4$	$\geq 900$	$\geq 200$	14,0	$\pm 0,9$	$\pm 0,8$	153,8	$\pm 4,6$	$\pm 16,4$	152
600	$\geq 4$	$\geq 1000$	$\geq 200$	7,3	$\pm 0,4$	$\pm 0,5$	95,3	$\pm 3,6$	$\pm 8,9$	101
700	$\geq 4$	$\geq 1100$	$\geq 200$	3,5	$\pm 0,1$	$\pm 0,2$	61,9	$\pm 3,1$	$\pm 6,8$	63
800	$\geq 4$	$\geq 1100$	$\geq 200$	2,0	$\pm 0,1$	$\pm 0,1$	61,9	$\pm 3,1$	$\pm 6,8$	63
900	$\geq 4$	$\geq 1200$	$\geq 325$	0,5	$\pm 0,0$	$\pm 0,0$	12,9	$\pm 1,7$	$\pm 2,1$	5
1000	$\geq 4$	$\geq 1200$	$\geq 325$	0,2	$\pm 0,0$	$\pm 0,0$	12,9	$\pm 1,7$	$\pm 2,1$	5

TABLE 4.8 – Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top  $\tau$ ).

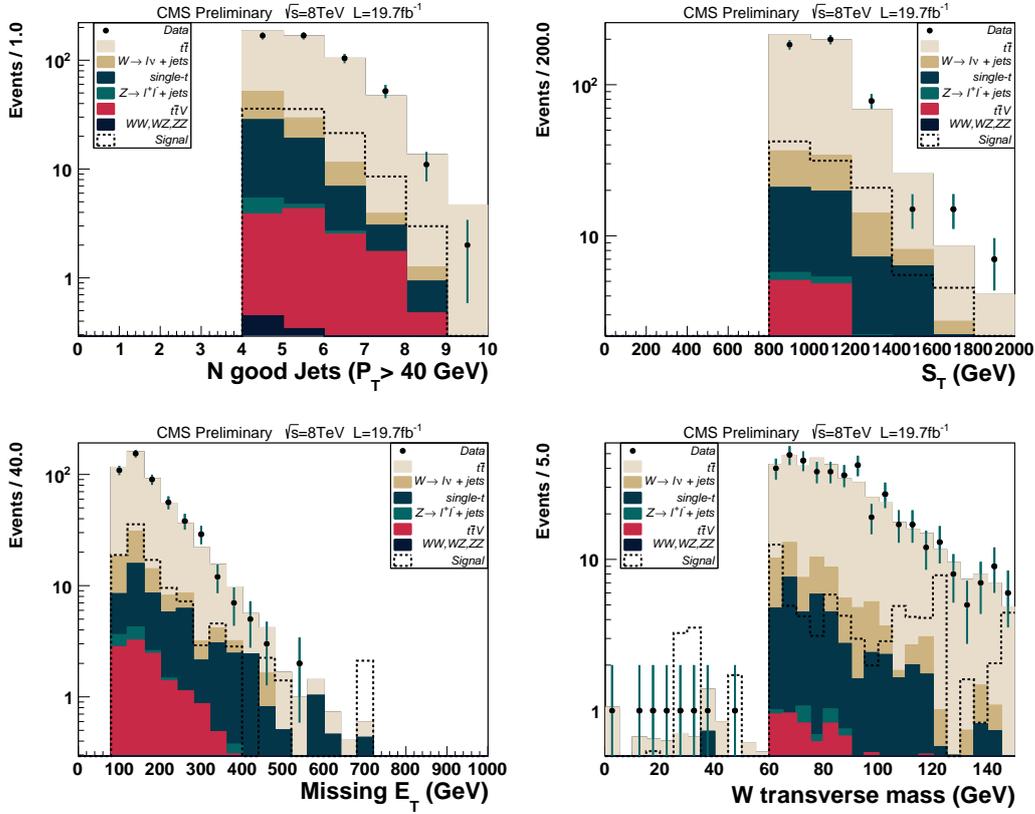


Fig. 4.16 – Multiplicité en jets de  $p_t > 40\text{GeV}$  (en haut à gauche)  $\cancel{E}_T$  (en bas à gauche) et  $S_t$  (à droite) la sélection optimale correspondant à  $m_{LQ} = 30$  GeV.

Masse LQ	limite observée (pb)	limite attendue (pb)	$\sigma$ théorique (pb)
200	13,8	16,6	17,4
300	1,51	1,74	1,89
400	0,458	0,475	0,342
500	0,203	0,206	0,082
600	0,0959	0,0832	0,0235
700	0,0504	0,0428	0,0076
800	0,02837	0,02676	0,00269
900	0,02230	0,01742	0,00101
1000	0,01311	0,01252	0,00040

TABLE 4.9 – Meilleures limites attendues et observées pour les différentes masses de LQ (LQ en top  $\tau$ ).

$\cancel{E}_T$  va donc jouer un rôle déterminant dans cette analyse. Pour le signal, une grande partie de l'énergie de la collision sera emportée par les neutrinos et restera non détectée. La désintégration de la paire  $t\bar{t}$  associée donnera en effet les même

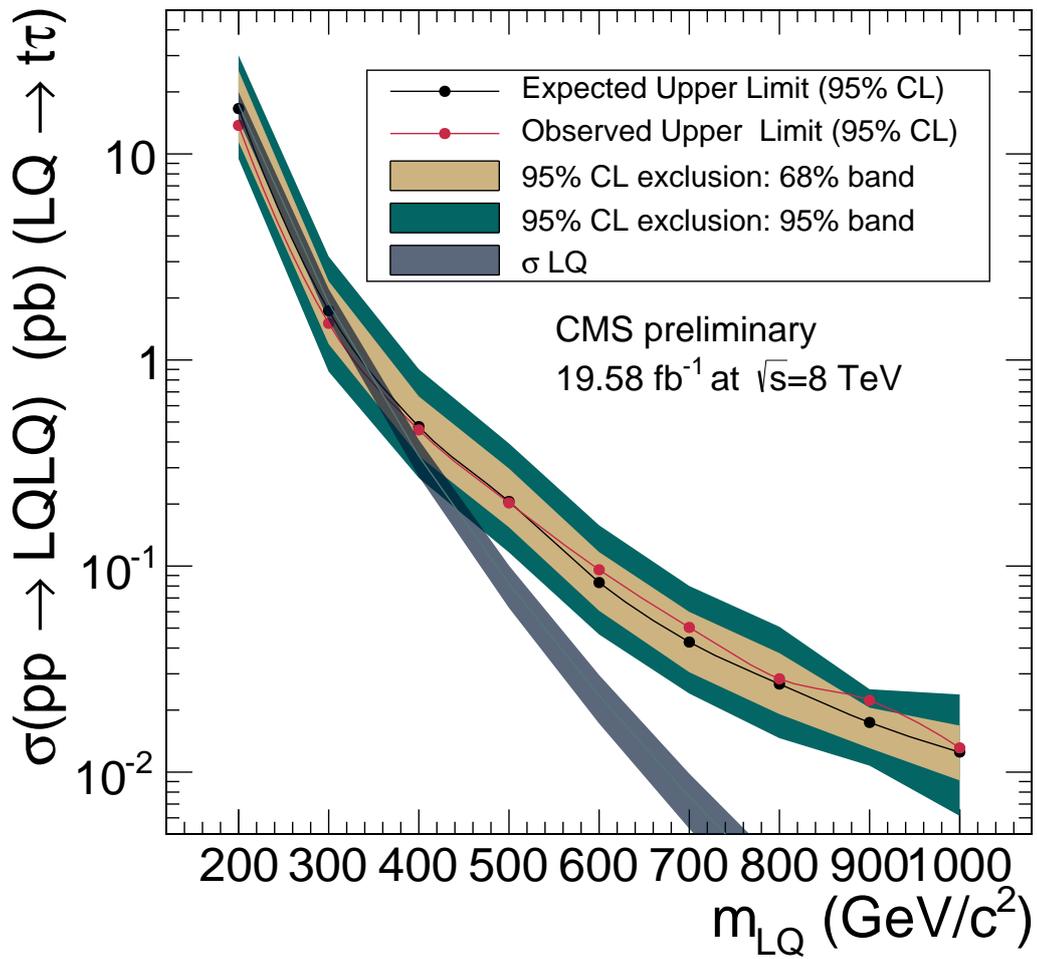


Fig. 4.17 – Limites attendues et observées pour les différentes masses de LQ ainsi que les bandes d'exclusion à 95% et à 68% (LQ en top  $\tau$ ).

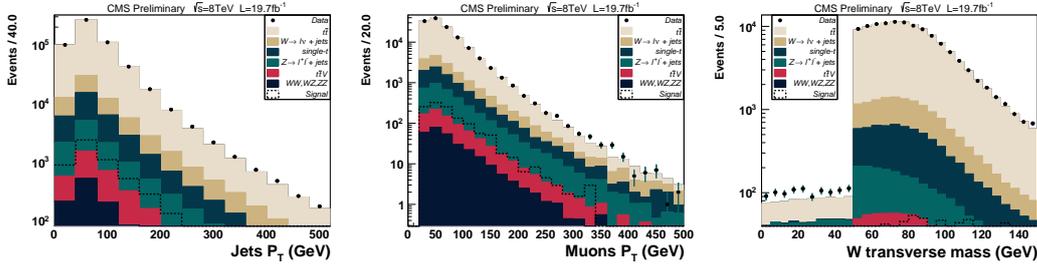


Fig. 4.18 –  $P_t$  des bons jets (à gauche) des muons (au milieu) et  $W_{mt}$  (à droite). Le signal correspond ici à une masse de LQ de 500 GeV et une désintégration à 100% en top  $\nu$ .

états finals que le fond du modèle standard. On attend donc ni plus d'activité ni des leptons plus énergétiques dans le signal que dans le fond  $t\bar{t}$ . L'état final de la désintégration des LQ en top  $\nu$  en collisionneur hadronique est donc très similaire à celui du fond  $pp \rightarrow t\bar{t}^3$  sauf qu'il prévoit une  $\cancel{E}_T$  beaucoup plus importante. Il est à noter que la désintégration préférentielle se fait en top  $\nu_\tau$  à cause de la conservation de la saveur dans l'hypothèse de non mélange entre générations. Cependant les désintégrations en top  $\nu_e$  et en top  $\nu_\mu$  sont autorisées par notre modèle (voir 1.4.1) et sont de toute façon indiscernables au niveau analyse. La somme des trois canaux de désintégration sera donc explorée en même temps et de façon inclusive.

#### 4.5.2 Pré-sélection des événements

L'état final étant très proche de celui de l'analyse précédente, on conserve un niveau de pré-sélection similaire à celui de l'analyse top  $\tau$ . C'est à dire :

- au moins un bon muon ( $p_t \geq 26$  GeV) ;
- au moins un jet de b (CSVM medium) ;
- 4 bons jets avec les mêmes seuils en  $p_t$  que la sélection top (45, 45, 35 et 30 GeV) ;
- $W_{mt} \geq 50$  GeV ;
- masse dimuon supérieure à 50 GeV.

Les seuls différences sont le léger relâchement sur le seuil en  $p_t$  du 4<sup>ème</sup> jet (30 GeV au lieu de 40GeV) et de la coupure en masse transverse du W (50 au lieu de 60 GeV). Ces petits changements ont été apportés pour obtenir une meilleure signification à ce niveau de pré-sélection (voir 4.4.2). Les accords données/MC n'en sont d'ailleurs pas altérés comme on peut le voir sur la figure 4.18. Pour notre signal, on se restreint donc au cas où un des deux quarks tops de la paire  $t\bar{t}$  donnera un W qui se désintégrera en un muon et un neutrino. Et où l'autre se désintégrera hadroniquement en deux jets. Dans ce cas, on obtient bien pour le signal : un muon (issu d'un W), 4 jets dont au moins un b (2 b et 2 issu de l'autre W) et de la  $\cancel{E}_T$  (provenant des neutrinos).

3. nombreux jets dont au moins 2 b jets et des leptons très énergétiques.

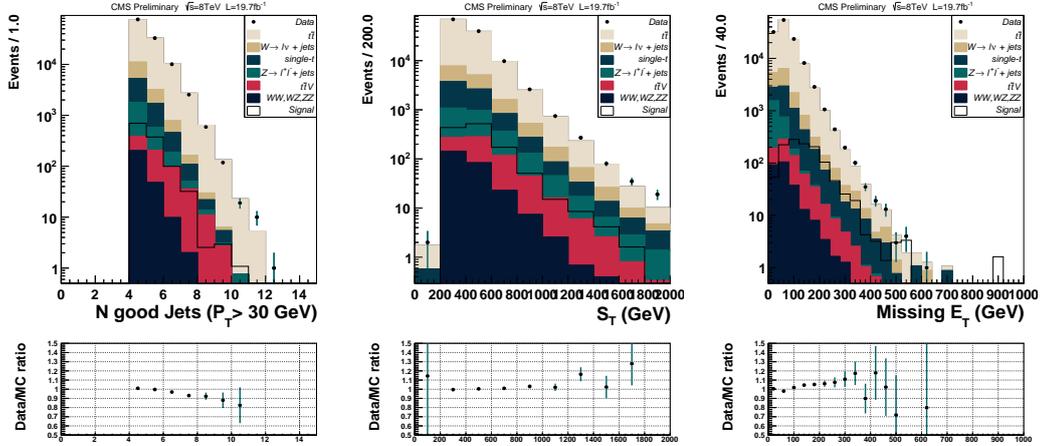


Fig. 4.19 – Multiplicité en jets de  $p_t > 40\text{GeV}$  (à gauche)  $S_t$  (au milieu) et  $\cancel{E}_T$  (à droite) pour la pré-sélection optimale. Le signal correspond ici à une masse de LQ de 500 GeV et une désintégration à 100% en top  $\nu$ .

### 4.5.3 Optimisation finale et incertitudes

Les trois variables sur lesquelles on effectue l’optimisation restent les mêmes (voir figure 4.19). Comme attendu,  $S_T$  et le nombre de bons jets sont beaucoup moins (voir pas) discriminant par rapport à l’analyse en top  $\tau$ . On décide cependant de les conserver dans la procédure d’optimisation car ils permettent de sonder indépendamment des zones différentes de l’espace des phases. En revanche, la  $\cancel{E}_T$  se détache bien pour le signal par rapport au fond du MS et elle va constituer notre principal levier dans cette analyse. La grille de coupures reste la même que pour l’analyse en top  $\tau$ , des pas de 25 GeV en  $\cancel{E}_T$  et des pas de 100 GeV en  $S_T$ .

Les résultats de cette optimisation sont présentés dans le tableau 4.10. Comme on s’y attendait, le nombre de jets n’intervient plus du tout dans l’optimisation. En revanche,  $S_T$  conserve son influence pour des faibles masses de LQ. La  $\cancel{E}_T$  est la variable la plus discriminante, notamment pour les grandes masses de LQ pour lesquelles l’algorithme va systématiquement chercher des événements avec au moins 400 GeV de  $\cancel{E}_T$ .

Les incertitudes restent de même nature et de même ordre de grandeur que dans l’analyse top  $\tau$  (partie 4.4). Elles sont présentées pour chaque point de masse de LQ au niveau de sélection optimal dans le tableau 4.11.

### 4.5.4 Masse limite et interprétation des résultats

La figure 4.20 et le tableau 4.12 montrent les limites attendues et observées obtenues après optimisation ainsi que la section efficace de production des LQ avec ses incertitudes. La limite supérieure sur la masse des LQ pour cette analyse top  $\nu$  est obtenue en faisant l’intersection de la limite de la section efficace observée à 95 % par la courbe inférieure de la section efficace théorique. Ainsi dans le cadre de ce

$M_{LQ}$ (GeV)	$N_{jets}$	$S_t$ (GeV)	$\cancel{E}_T$ (GeV)	$N_{signal}$	$\Delta N_{signal}^{stat}$	$\Delta N_{signal}^{syst}$	$N_{fond}$	$\Delta N_{fond}^{stat}$	$\Delta N_{fond}^{syst}$	$N_{donnees}^{ev}$
200	$\geq 4$	$\geq 900$	$\geq 150$	192.2	$\pm 51.3$	$\pm 42.8$	395.8	$\pm 7.3$	$\pm 44.5$	405
300	$\geq 4$	$\geq 400$	$\geq 275$	63.6	$\pm 9.7$	$\pm 2.6$	368.8	$\pm 7.5$	$\pm 38.7$	392
400	$\geq 4$	$\geq 300$	$\geq 350$	20.2	$\pm 2.3$	$\pm 1.1$	94.2	$\pm 4.1$	$\pm 11.7$	85
500	$\geq 4$	$\geq 300$	$\geq 350$	12.4	$\pm 0.9$	$\pm 0.7$	94.2	$\pm 4.1$	$\pm 11.7$	85
600	$\geq 4$	$\geq 200$	$\geq 375$	5.2	$\pm 0.3$	$\pm 0.3$	61.5	$\pm 3.4$	$\pm 7.3$	63
700	$\geq 4$	$\geq 500$	$\geq 475$	1.4	$\pm 0.1$	$\pm 0.1$	11.7	$\pm 1.5$	$\pm 1.7$	10
800	$\geq 4$	$\geq 200$	$\geq 475$	0.7	$\pm 0.0$	$\pm 0.0$	11.8	$\pm 1.5$	$\pm 1.7$	10
900	$\geq 4$	$\geq 200$	$\geq 475$	0.3	$\pm 0.0$	$\pm 0.0$	11.8	$\pm 1.5$	$\pm 1.7$	10
1000	$\geq 4$	$\geq 300$	$\geq 500$	0.2	$\pm 0.0$	$\pm 0.0$	8.7	$\pm 1.4$	$\pm 0.9$	7

TABLE 4.10 – Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top  $\nu$ ).

type	mLQ (GeV)	stat	JES	JER	trig eff	PU	X sec
MC	200	$\pm 1.8$	-5.9 +6.8	-1.1 +2.2	-1.6 +1.9	-0.1 +0.1	$\pm 5.4$
SIG	200	$\pm 26.7$	-15.2 +0.0	-7.6 +0.0	-1.4 +1.4	-4.1 +5.5	$\pm 0$
MC	300	$\pm 2.0$	-6.1 +6.2	-1.5 +1.5	-1.6 +1.9	-0.2 +0.1	$\pm 5.4$
SIG	300	$\pm 15.3$	-0.1 +2.4	-0.0 +0.1	-1.6 +1.5	-1.8 +1.2	$\pm 0$
MC	400	$\pm 4.3$	-7.4 +7.8	-1.1 +2.0	-1.6 +1.9	-0.5 +0.4	$\pm 5.4$
SIG	400	$\pm 11.5$	-3.5 +2.2	-0.8 +1.4	-1.6 +1.5	-1.3 +0.8	$\pm 0$
MC	500	$\pm 4.3$	-7.4 +7.8	-1.1 +2.0	-1.6 +1.9	-0.5 +0.4	$\pm 5.4$
SIG	500	$\pm 7.2$	-3.6 +3.6	-0.6 +0.9	-1.6 +1.5	-0.2 +0.3	$\pm 0$
MC	600	$\pm 5.5$	-7.4 +7.1	-0.3 +1.4	-1.6 +2.1	-0.3 +0.2	$\pm 5.4$
SIG	600	$\pm 6.0$	-3.3 +2.9	-0.2 +1.2	-1.6 +1.6	-1.1 +0.5	$\pm 0$
MC	700	$\pm 12.7$	-6.5 +9.4	-3.1 +0.7	-1.5 +1.3	-1.1 +0.6	$\pm 5.4$
SIG	700	$\pm 6.6$	-7.2 +4.3	-1.6 +0.1	-1.6 +1.5	-1.2 +1.4	$\pm 0$
MC	800	$\pm 12.7$	-6.5 +9.4	-3.1 +0.7	-1.5 +1.3	-1.1 +0.6	$\pm 5.4$
SIG	800	$\pm 5.4$	-4.5 +2.5	-0.2 +1.3	-1.6 +1.5	-1.7 +2.1	$\pm 0$
MC	900	$\pm 12.7$	-6.5 +9.4	-3.1 +0.7	-1.5 +1.3	-1.1 +0.6	$\pm 5.4$
SIG	900	$\pm 5.0$	-2.0 +2.1	-0.5 +0.0	-1.6 +1.5	-0.1 +0.1	$\pm 0$
MC	1000	$\pm 15.8$	-4.6 +6.2	-0.2 +2.2	-1.5 +1.2	-1.9 +1.8	$\pm 5.4$
SIG	1000	$\pm 4.5$	-3.7 +2.2	-0.3 +1.2	-1.6 +1.5	-1.7 +1.6	$\pm 0$

TABLE 4.11 – Incertitudes en % pour les différentes masses de LQ (LQ en top  $\nu$ ).

Masse LQ (GeV)	limite observée (pb)	limite attendue (pb)	$\sigma$ théorique (pb)
200	9.0	8.1	17.4
300	2.80	2.21	1.89000
400	0.371	0.445	0.342
500	0.138	0.174	0.082
600	0.0999	0.0931	0.0235
700	0.0421	0.0470	0.0076
800	0.02841	0.03230	0.00269
900	0.02455	0.02775	0.00101
1000	0.01685	0.01945	0.00040

TABLE 4.12 – Limites attendues et observées pour différentes masses de LQ (LQ en top  $\nu$ ).

modèle, on peut exclure à 95 % de confiance des LQ de masses inférieures à 240 GeV.

## 4.6 Recherche de Leptoquarks de 3 ème génération se désintégrant en top $\mu$

### 4.6.1 État final et canal d'analyse

Ici, on considère le rapport de branchement des LQ à 100 % en quark top et un muon. Notre état final est donc exclusivement  $LQLQ \rightarrow t\bar{t}\mu^+\mu^-$ . L'état final sera donc 2 muons et les produits de désintégrations habituels de  $t\bar{t}$ . On s'attend donc à ce que le nombre de muons soit plus important dans le signal que dans le fond principal. La grande différence par rapport aux deux analyses précédentes est l'absence de neutrinos dans l'état final de désintégration des LQ. Ainsi, la  $\cancel{E}_T$  ne sera plus discriminante par rapport au fond du MS. Il faudra donc trouver d'autres variables permettant de séparer le signal du fond. Dans un collisionneur, la désintégration des LQ en top  $\mu$  est donc caractérisée par des jets dont au moins 2 b jets, des leptons chargés très énergétiques plus nombreux dont au moins 2 muons. Il faut noter que ce canal de désintégration comporte un mélange entre deux générations de leptons (2ème et 3ème). Il est cependant autorisé par la théorie, comme nous l'avons précisé dans l'introduction de notre modèle effectif de LQ ( voir partie 1.4.1). Compte tenu de la présence des deux muons supplémentaires dans l'état final, il permet par ailleurs une bonne séparation du signal par rapport aux fonds du MS.

### 4.6.2 Pré-sélection des événements

Pour cette analyse top  $\mu$ , on part donc d'un constat simple, le nombre de leptons chargés est très discriminant pour séparer le signal du bruit de fond MS. La figure 4.21 représente le nombre de bons leptons au niveau de sélection de l'analyse top

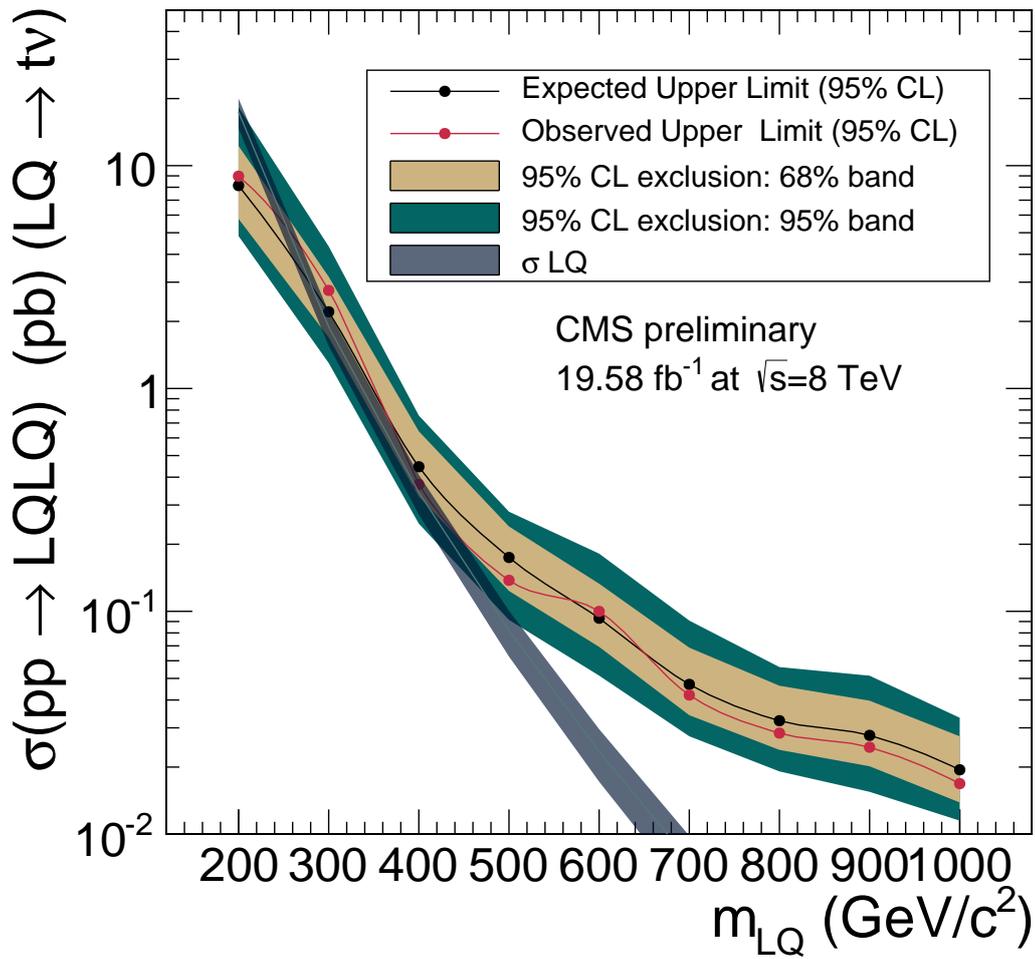


Fig. 4.20 – Limites attendues et observées en fonction de la masse de LQ ainsi que les bandes d'exclusion à 95% et à 68% de confiance (LQ en top  $\nu$ ).

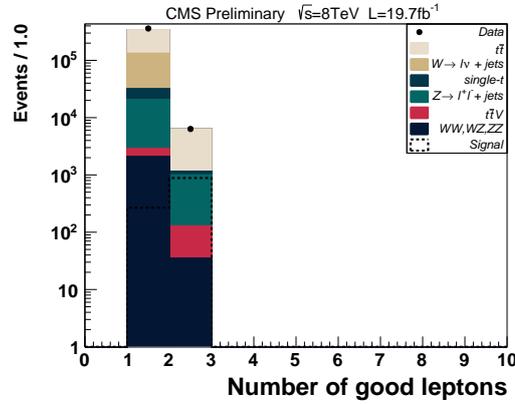


Fig. 4.21 – Nombre de bons leptons correspondant à la sélection de l'analyse top  $\tau$  (sans la coupure sur  $W_{mt}$ ) mais avec un signal top  $\mu$  pour  $m_{LQ} = 500$  GeV

$\tau$  (mais sans coupure sur la masse transverse du W) en superposant un signal top  $\mu$  correspondant à  $m_{LQ} = 500$  GeV. On s'aperçoit que de demander deux leptons dans l'état final va permettre d'éliminer une grande part du fond par rapport au signal et donc d'augmenter grandement la sensibilité de cette analyse.

On calque notre sélection d'objets sur celle de l'analyse top  $\tau$  décrite dans la partie 4.4. On garde ainsi la même sélection pour les jets et on continue à utiliser le module fournissant les efficacités de trigger. Comme on va se placer dans une configuration avec au moins deux leptons dans l'état final, la coupure sur la masse transverse du W pour exactement un bon muon n'est plus requise. Elle permettait en effet de supprimer les événements QCD dans les données car ils n'étaient pas pris en compte dans la simulation. Mais on vérifiera que le fait de demander 2 leptons nous permet de parvenir au même résultat. A la place, on ajoute une coupure sur la masse invariante des deux muons afin de réduire le fond  $Z \rightarrow \mu\mu$ . Ainsi, quand exactement deux bons muons sont identifiés, si la masse reconstruite des deux muons est comprise entre 81 GeV et 101 GeV on ne conserve pas l'événement. Enfin, pour augmenter la statistique, on relâche un peu la coupure en  $p_t$  sur les muons en demandant des muons de 20 GeV au lieu de 26 GeV (selon les recommandations du groupe top pour une analyse dilepton). On conserve par ailleurs la demande d'au moins un jet de b par événement pour supprimer les fonds MS autres que  $t\bar{t}$ .

La figure 4.22 montre (en haut) les caractéristiques des muons ( $p_t, \eta$ ). Elle permet de voir que même en relâchant la coupure sur le  $p_t$  des muons on garde un bon accord données/MC. En bas à gauche de cette même figure, le diagramme de la masse transverse du W permet de s'assurer que les problèmes liés à la QCD n'existent plus en demandant au moins deux bons leptons. En effet on ne retrouve plus l'excès d'événements QCD dans les données à basse masse transverse. Enfin, le dernier diagramme montre la masse invariante dimuon et la coupure correspondant au pic du boson Z. Le signal étant largement distribué en dehors du pic, cette coupure sera très discriminante par rapport au fond du MS  $Z \rightarrow \mu\mu$ .

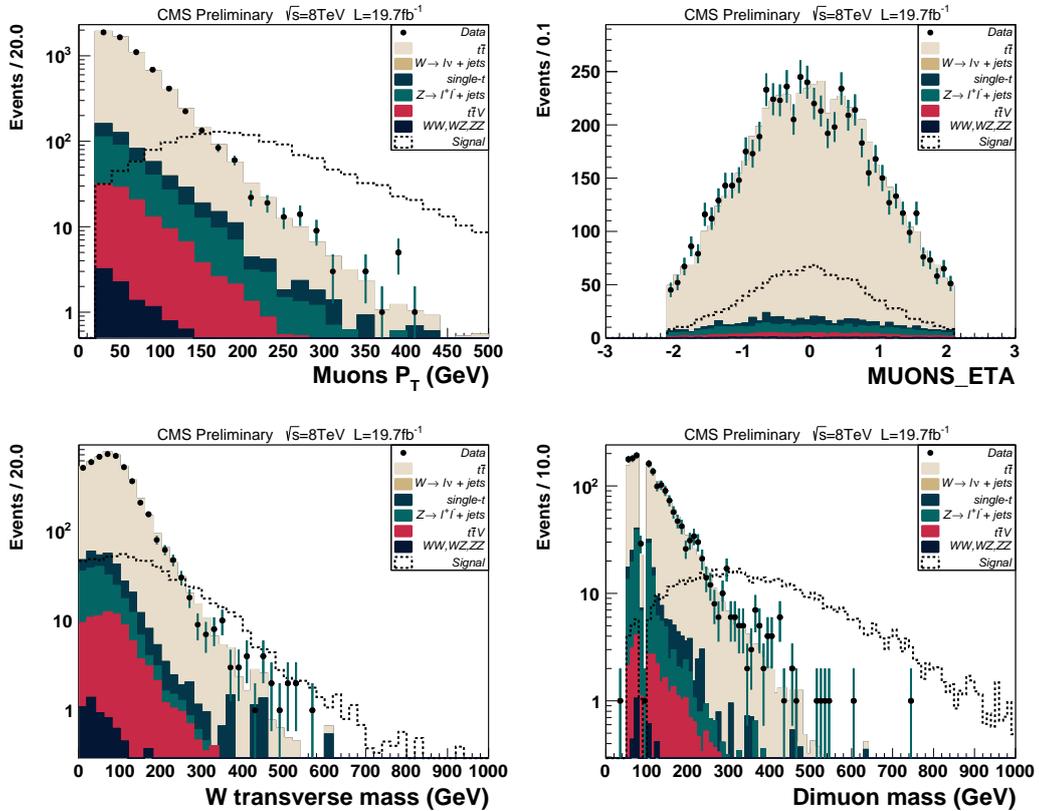


Fig. 4.22 – En haut, impulsion transverse et pseudo rapidité des muons. En bas, la masse transverse du W (à gauche) et la masse dimuon (à droite)

Notre sélection requérant 2 leptons (muons ou électrons), il faut aussi s'assurer que les caractéristiques des électrons sont bien simulées. La figure 4.23 (en haut), présente le  $p_t$  et  $\eta$  des électrons et permet de contrôler un bon accord données/MC. En bas de cette même figure, le nombre de bons muons (à gauche) et de bons électrons (à droite) sont comparés. Notre statistique finale sera principalement constituée des deux premiers intervalles de chaque histogramme. Des événements à deux bons muons (environ 1500) et des événements avec un bon muon et un bon électron (environ 3000). On voit bien sur ces diagrammes l'intérêt de la stratégie dileptons (plutôt que dimuons par exemple). Elle permet en effet d'obtenir une plus grande proportion de signal par rapport au fonds du MS tout en gardant un nombre d'événements conséquent pour éviter de trop grandes erreurs statistiques.

#### 4.6.3 Optimisation finale et incertitudes

Dans cette dernière analyse, la  $\cancel{E}_T$ , n'est plus discriminante, en effet, il n'y a plus de neutrino dans l'état final de désintégration des LQ.  $S_T$  reste une variable discriminante (voir figure 4.24 à gauche). Mais quand on regarde un peu plus précisément, on s'aperçoit que le  $p_t$  du muon le plus énergétique d'une part et

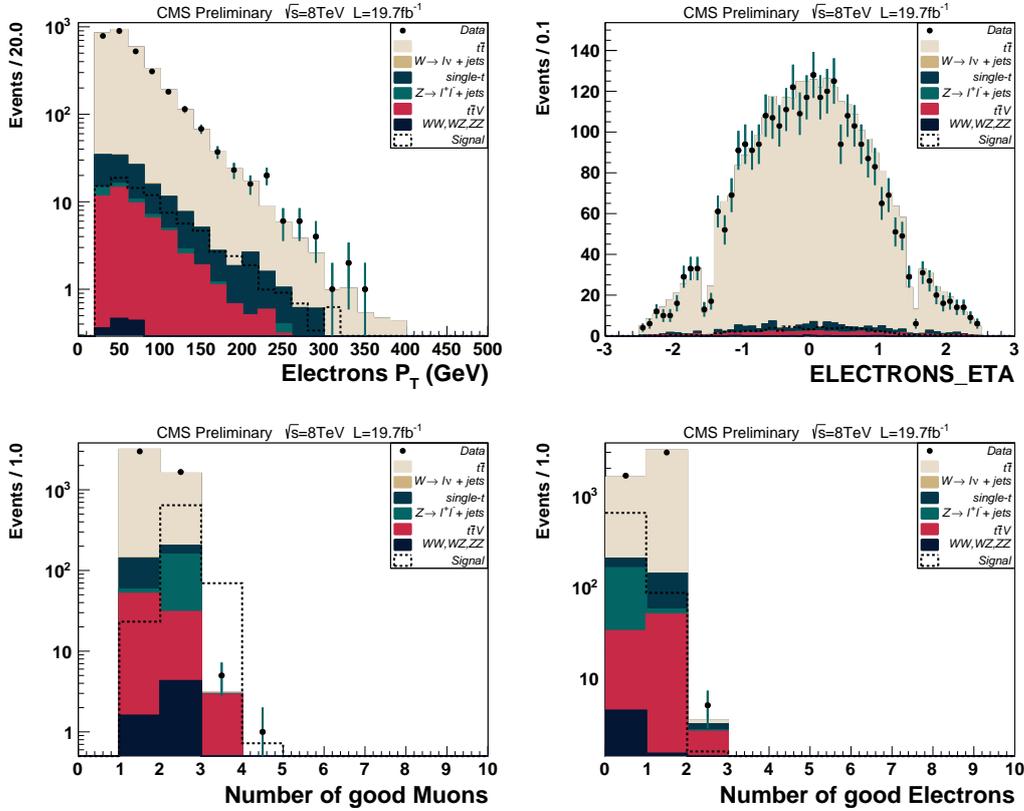


Fig. 4.23 – En haut,  $p_t$  et  $\eta$  des électrons ; En bas, nombre de bon muons et d'électrons.

la somme scalaire des  $p_t$  des jets d'autre part sont aussi très discriminants. Non seulement ce sont des grandeurs plus simples que  $S_t$ , mais encore, elles permettent de bien séparer le signal du bruit de fond dans des zones distinctes de l'espace des phases. On choisit donc de réaliser l'optimisation sur ces quantités plus simples et bien modélisées (voir figure 4.24 à droite). On garde par ailleurs le nombre de jets dans notre algorithme d'optimisation. La procédure d'optimisation reste la même que dans la partie 4.4, toujours basée sur la meilleure limite obtenue selon la méthode  $Cl_s$  modifiée (voir partie 4.4.5.1). La grille de coupures est maintenant constituée de pas de 20 GeV pour le  $p_t$  du muon le plus énergétique et de pas de 100 GeV en  $H_T$ .

Les résultats de cette optimisation sont présentés dans le tableau 4.13. Compte tenu de la sélection des objets, le  $p_t$  du muon le plus énergétique sélectionné sera toujours au moins égal à 20 GeV et  $H_T$  sera toujours supérieur à la somme des seuils en  $p_t$  des quatre premiers jets ( $45+45+35+30=155$  GeV).

Les incertitudes restent de même ordre que dans l'analyse top  $\tau$  (partie 4.4). Elles sont présentées pour chaque valeur de masse de LQ au niveau de la sélection optimale dans le tableau 4.14.

$M_{LQ}$ (GeV)	$N_{jets}$	$H_t$ (GeV)	$\mu_{P_t}$ (GeV)	$N_{signal}$	$\Delta N_{signal}^{stat}$	$\Delta N_{signal}^{syst}$	$N_{fond}$	$\Delta N_{fond}^{stat}$	$\Delta N_{fond}^{syst}$	$N_{donnees}^{ev}$
200	$\geq 6$	$\geq 155$	$\geq 20$	7049.8	$\pm 307.7$	$\pm 679.7$	731.6	$\pm 6.2$	$\pm 90.9$	670
300	$\geq 6$	$\geq 300$	$\geq 80$	3902.2	$\pm 72.1$	$\pm 253.6$	147.5	$\pm 2.9$	$\pm 16.5$	141
400	$\geq 5$	$\geq 200$	$\geq 160$	1373.4	$\pm 17.2$	$\pm 63.6$	59.7	$\pm 1.9$	$\pm 5.8$	61
500	$\geq 5$	$\geq 200$	$\geq 240$	272.8	$\pm 3.8$	$\pm 11.8$	9.9	$\pm 0.9$	$\pm 0.8$	7
600	$\geq 5$	$\geq 155$	$\geq 240$	142.5	$\pm 1.4$	$\pm 4.8$	9.9	$\pm 0.9$	$\pm 0.8$	7
700	$\geq 5$	$\geq 155$	$\geq 240$	65.6	$\pm 0.5$	$\pm 2.2$	9.9	$\pm 0.9$	$\pm 0.8$	7
900	$\geq 5$	$\geq 600$	$\geq 240$	9.6	$\pm 0.1$	$\pm 0.3$	5.2	$\pm 0.7$	$\pm 0.6$	6
1000	$\geq 4$	$\geq 1000$	$\geq 240$	2.7	$\pm 0.0$	$\pm 0.1$	1.8	$\pm 0.6$	$\pm 0.3$	0

TABLE 4.13 – Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top  $\mu$ ).

type	mLQ (GeV)	stat	JES	JER	trig eff	PU	X sec
MC	200	$\pm 0.8$	-7.8 +7.6	-1.1 +0.9	-1.3 +1.4	-0.1 +0.1	$\pm 5.5$
SIG	200	$\pm 4.4$	-6.2 +6.7	-1.2 +0.4	-0.9 +0.9	-1.8 +1.9	$\pm 0$
MC	300	$\pm 2.0$	-6.8 +6.8	-1.3 +0.5	-1.2 +1.1	-0.4 +0.3	$\pm 5.5$
SIG	300	$\pm 1.8$	-4.5 +4.5	-0.7 +0.4	-0.7 +0.7	-0.7 +0.9	$\pm 0$
MC	400	$\pm 3.1$	-5.6 +5.0	-0.1 +0.5	-1.1 +1.1	-0.3 +0.1	$\pm 5.5$
SIG	400	$\pm 1.2$	-3.2 +2.1	-0.5 +0.7	-0.7 +0.7	-0.1 +0.3	$\pm 0$
MC	500	$\pm 8.7$	-4.0 +2.9	-1.7 +1.6	-1.0 +1.0	-1.2 +0.9	$\pm 5.4$
SIG	500	$\pm 1.4$	-3.0 +2.9	-0.0 +0.3	-0.7 +0.7	-0.7 +0.8	$\pm 0$
MC	600	$\pm 8.7$	-4.0 +2.9	-1.7 +1.6	-1.0 +1.0	-1.2 +0.9	$\pm 5.4$
SIG	600	$\pm 0.9$	-2.3 +1.9	-0.4 +0.4	-0.7 +0.7	-0.7 +0.7	$\pm 0$
MC	700	$\pm 8.7$	-4.0 +2.9	-1.7 +1.6	-1.0 +1.0	-1.2 +0.9	$\pm 5.4$
SIG	700	$\pm 0.7$	-1.7 +2.3	-0.7 +0.3	-0.7 +0.7	-0.4 +0.6	$\pm 0$
MC	900	$\pm 13.7$	-3.1 +6.2	-1.7 +4.2	-1.0 +1.0	-4.3 +4.0	$\pm 5.4$
SIG	900	$\pm 0.7$	-2.4 +2.1	-0.0 +0.3	-0.7 +0.7	-0.0 +0.2	$\pm 0$
MC	1000	$\pm 33.0$	-3.4 +5.2	-2.6 +0.0	-1.0 +1.0	-12.6 +11.7	$\pm 5.5$
SIG	1000	$\pm 0.9$	-2.9 +2.7	-0.2 +0.3	-0.7 +0.7	-0.5 +0.6	$\pm 0$

TABLE 4.14 – Incertitudes en % pour les différentes masses de LQ (LQ en top  $\mu$ ).

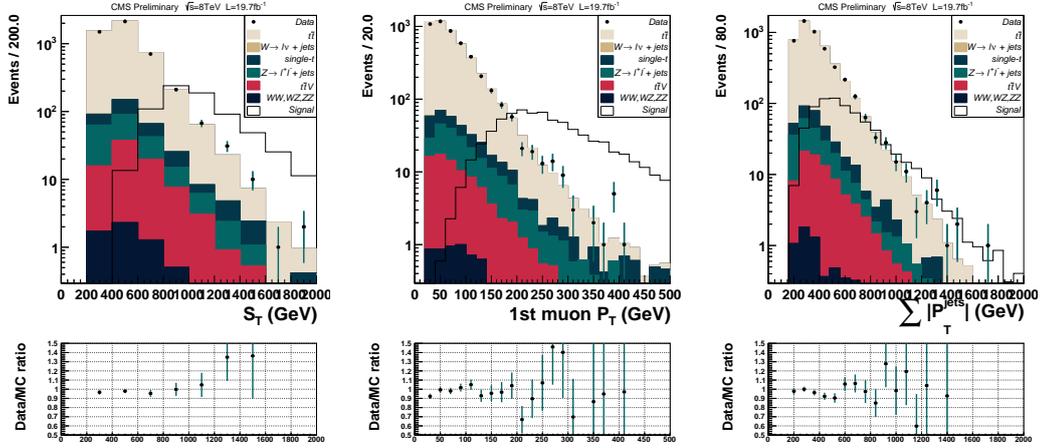


Fig. 4.24 –  $S_t$  (à gauche),  $p_t$  du muon le plus énergétique (au milieu) et  $H_t$  correspondant à la sélection dileptons avec un signal superposé  $LQ \rightarrow t\mu$  ( $m_L Q = 500\text{GeV}$ )

Masse LQ (GeV)	limite observée (pb)	limite attendue (pb)	$\sigma$ théorique (pb)
200	0.2	0.3	17.4
300	0.02	0.02	1.89
400	0.005	0.005	0.342
500	0.002	0.002	0.082
600	0.0010	0.0012	0.0235
700	0.0008	0.0009	0.0076
800	0.00065	0.00075	0.00269
900	0.00057	0.00065	0.00101
1000	0.00043	0.00058	0.00040

TABLE 4.15 – Limites attendues et observées pour les différentes masses de LQ (LQ en top  $\mu$ ).

#### 4.6.4 Masse limite et interprétation des résultats

La figure 4.25 et le tableau 4.15 montrent les limites attendues et observées obtenues après optimisation ainsi que la section efficace de production des LQ avec ses incertitudes. La limite supérieure sur la masse des LQ pour cette analyse top  $\mu$  est obtenue en faisant l'intersection de la limite de la section efficace observée à 95 % de confiance par la courbe inférieure de la section efficace théorique. Ainsi, dans le cadre de ce modèle, on peut exclure à 95 % de confiance des LQ pour des masses inférieures à 900 GeV.

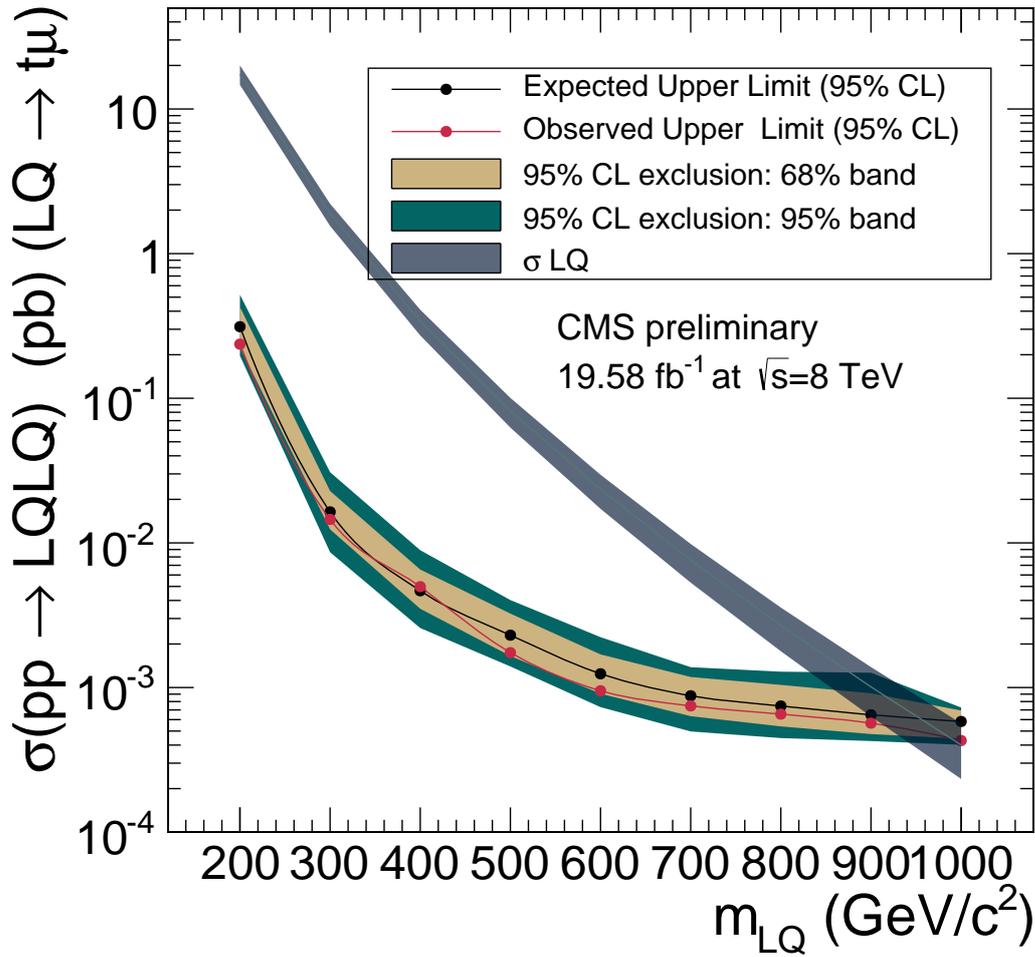


Fig. 4.25 – Limites attendues et observées pour les différentes masses de LQ ainsi que les bandes d'exclusion à 95% et à 68% (LQ en top  $\mu$ ).

## 4.7 Conclusions et perspectives

A masse égale, la section efficace de production de paire de LQ est environ 10 fois plus faible que celle d'une paire  $t\bar{t}$ . Quand la désintégration des deux LQ est accompagnée de la production d'une paire  $t\bar{t}$ , les états finals sont de plus extrêmement proches. Les analyses concernant la désintégration de LQ de troisième génération en quark top représentent donc un véritable challenge en collisionneur hadronique. Voilà pourquoi, elles n'avaient jamais été menées à bien jusque là. Avec la totalité du run 2012 à 8 TeV une fenêtre d'analyse s'est ouverte. Encore fallait-il dépasser les contraintes techniques liées à l'augmentation de la luminosité.

Après analyse, les leptoquarks de troisième génération se désintégrant en top  $\tau$  et top  $\nu$  ont été respectivement exclus jusqu'à des masses de 300 et 240 GeV. Ces limites restent relativement basses car, à 8 TeV, les signaux correspondants restent très proches des fonds du modèle standard. Même les variables les plus discriminantes ne permettent pas de séparer facilement le signal du fond du MS avec la statistique du run 2012. Les pentes de la section efficace de production et de la limite observée sont d'ailleurs très proches dans la zone d'exclusion ce qui rend ces études très sensibles aux fluctuations statistiques. Ces analyses, sont donc à la limite de ce que le LHC peut explorer pour l'instant et elles ont représenté un véritable défi en terme d'analyse de données et d'optimisation. Avec la reprise du fonctionnement en 2015, le passage à 13 TeV et l'augmentation de la statistique, les limites pour ces deux premières analyses seront sans aucun doute grandement améliorées. Sous condition cependant que les défis techniques que présente ce nouveau mode de fonctionnement soient maîtrisés (augmentation de l'empilement, mise au point des nouveaux chemins de déclenchement...).

Pour la dernière analyse en top  $\mu$ , des masses de LQ ont par contre été exclues jusqu'à 900 GeV. Le fait de demander au moins deux leptons permet en effet une forte suppression du fond du MS par rapport au signal. On est ici dans le cas d'une analyse beaucoup plus facile. Une fois qu'on a déterminé le bon niveau de présélection, l'optimisation des coupures donne des résultats impressionnants en terme de limites. Les pentes de la section efficace de production et de la limite observée sont ici bien distinctes au niveau de l'intersection ce qui donne une limite sur la masse de LQ sans équivoque.

Au final, les analyses de ces trois canaux de désintégration ont permis de poser des limites inédites sur les LQ de troisième génération dans les états finals comportant au moins un muon.



# Bibliographie

- [1] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics : a practical introduction ; 4th ed.* CRC Press, Boca Raton, FL, 2013. (Cit  en page 2.)
- [2] E. Bauer, *Introduction   la th orie des groupes et   ses applications   la physique quantique.* Annales de l'Institut Henri Poincar . Les Presses universitaires de France, 1933.  
<http://books.google.fr/books?id=ajc-AQAAIAAJ>. (Cit  en page 3.)
- [3] C. Chevalley, *Theory of Lie Groups*. No. vol. 1 in Princeton landmarks in mathematics and physics. Princeton University Press, 1999.  
<http://books.google.fr/books?id=t-41JV63sRIC>. (Cit  en page 3.)
- [4] S. Thornton and J. Marion, *Classical Dynamics of Particles and Systems*. Brooks/Cole, Cengage Learning, 2004.  
<http://books.google.fr/books?id=H0qLQgAACAAJ>. (Cit  en page 3.)
- [5] D. Tong, « Lectures on Dynamics and Relativity ».  
<http://www.damtp.cam.ac.uk/user/tong/relativity.html>. (Cit  en page 5.)
- [6] R. Ellis, W. Stirling, and B. Webber, *QCD and Collider Physics*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2003. <http://books.google.fr/books?id=TqrPVoS6s0UC>. (Cit  en page 8.)
- [7] G. P. Salam, « Elements of QCD for hadron colliders », [arXiv:1011.5131](https://arxiv.org/abs/1011.5131) [hep-ph]. (Cit  en page 8.)
- [8] A. Pich, « The Standard model of electroweak interactions », [arXiv:0705.4264](https://arxiv.org/abs/0705.4264) [hep-ph]. (Cit  en page 8.)
- [9] A. Djouadi, « The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model », *Phys.Rept.* **459** (2008) 1–241, [arXiv:hep-ph/0503173](https://arxiv.org/abs/hep-ph/0503173) [hep-ph]. (Cit  en page 8.)
- [10] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, « Experimental Test of Parity Conservation in Beta Decay », *Phys. Rev.* **105** (Feb, 1957) 1413–1415.  
<http://link.aps.org/doi/10.1103/PhysRev.105.1413>. (Cit  en page 9.)
- [11] M. Kobayashi and T. Maskawa, « CP Violation in the Renormalizable Theory of Weak Interaction », *Prog. Theor. Phys.* **49** (1973) 652–657. (Cit  en page 10.)
- [12] Particle Data Group Collaboration, J. B. et al., « Review of Particle Physics », *Phys. Rev. D* **86** (Jul, 2012) 010001.  
<http://link.aps.org/doi/10.1103/PhysRevD.86.010001>. (Cit  en pages 13 et 101.)

- [13] S. F. King and C. Luhn, « Neutrino Mass and Mixing with Discrete Symmetry », *Rept.Prog.Phys.* **76** (2013) 056201, [arXiv:1301.1340](https://arxiv.org/abs/1301.1340) [hep-ph]. (Cité en page 15.)
- [14] W. Buchmüller and D. Wyler, « Constraints on SU(5)-type leptoquarks », *Physics Letters B* **177** no. 3–4, (1986) 377 – 382. <http://www.sciencedirect.com/science/article/pii/0370269386907719>. (Cité en page 16.)
- [15] W. Buchmüller, R. Ruckl, and D. Wyler, « Leptoquarks in Lepton - Quark Collisions », *Phys.Lett.* **B191** (1987) 442–448. (Cité en pages 16 et 17.)
- [16] J. Blumlein and R. Ruckl, « Production of scalar and vector leptoquarks in  $e^+ e^-$  annihilation », *Phys.Lett.* **B304** (1993) 337–346. (Cité en page 20.)
- [17] S. Davidson, D. C. Bailey, and B. A. Campbell, « Model independent constraints on leptoquarks from rare processes », *Z.Phys.* **C61** (1994) 613–644, [arXiv:hep-ph/9309310](https://arxiv.org/abs/hep-ph/9309310) [hep-ph]. (Cité en page 21.)
- [18] CMS Collaboration Collaboration, S. Chatrchyan *et al.*, « Search for pair production of first- and second-generation scalar leptoquarks in  $pp$  collisions at  $\sqrt{s} = 7$  TeV », *Phys.Rev.* **D86** (2012) 052013, [arXiv:1207.5406](https://arxiv.org/abs/1207.5406) [hep-ex]. (Cité en page 21.)
- [19] ATLAS Collaboration Collaboration, G. Aad *et al.*, « Search for pair production of first or second generation leptoquarks in proton-proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector at the LHC », *Phys.Rev.* **D83** (2011) 112006, [arXiv:1104.4481](https://arxiv.org/abs/1104.4481) [hep-ex]. (Cité en page 21.)
- [20] M. Leurer, « A Comprehensive study of leptoquark bounds », *Phys.Rev.* **D49** (1994) 333–342, [arXiv:hep-ph/9309266](https://arxiv.org/abs/hep-ph/9309266) [hep-ph]. (Cité en page 21.)
- [21] D0 Collaboration Collaboration, A. et Al., « Search for Third Generation Scalar Leptoquarks Decaying into  $\tau b$  », *Phys. Rev. Lett.* **101** (Dec, 2008) 241802. <http://link.aps.org/doi/10.1103/PhysRevLett.101.241802>. (Cité en page 21.)
- [22] CMS Collaboration Collaboration, S. Chatrchyan *et al.*, « Search for pair production of third-generation leptoquarks and top squarks in  $pp$  collisions at  $\sqrt{s} = 7$  TeV », *Phys.Rev.Lett.* **110** (2013) 081801, [arXiv:1210.5629](https://arxiv.org/abs/1210.5629) [hep-ex]. (Cité en page 21.)
- [23] ATLAS Collaboration Collaboration, G. Aad *et al.*, « Search for third generation scalar leptoquarks in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector », *JHEP* **1306** (2013) 033, [arXiv:1303.0526](https://arxiv.org/abs/1303.0526) [hep-ex]. (Cité en page 21.)
- [24] CMS Collaboration Collaboration, S. Chatrchyan *et al.*, « Search for third-generation leptoquarks and scalar bottom quarks in  $pp$  collisions at  $\sqrt{s} = 7$  TeV », *JHEP* **1212** (2012) 055, [arXiv:1210.5627](https://arxiv.org/abs/1210.5627) [hep-ex]. (Cité en page 21.)

- [25] ALICE Collaboration, « The ALICE experiment at the CERN LHC », *Journal of Instrumentation* **3** no. 08, (2008) S08002. <http://stacks.iop.org/1748-0221/3/i=08/a=S08002>. (Cité en pages 24 et 29.)
- [26] ATLAS Collaboration, « The ATLAS Experiment at the CERN Large Hadron Collider », *Journal of Instrumentation* **3** no. 08, (2008) S08003. <http://stacks.iop.org/1748-0221/3/i=08/a=S08003>. (Cité en pages 24 et 29.)
- [27] T. C. Collaboration *Journal of Instrumentation* **3** no. 08, (2008) S08004. <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>. (Cité en pages 24, 29 et 33.)
- [28] T. L. Collaboration *Journal of Instrumentation* **3** no. 08, (2008) S08005. <http://stacks.iop.org/1748-0221/3/i=08/a=S08005>. (Cité en pages 24 et 29.)
- [29] N. Cartiglia, « Measurement of the proton-proton total cross section at 2, 7, 8 and 57 TeV », [arXiv:1303.2927](https://arxiv.org/abs/1303.2927) [hep-ex]. (Cité en page 25.)
- [30] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, « LHC Design Report v.3 : the LHC Injector Chain », tech. rep., Geneva, 2004. (Cité en page 25.)
- [31] CMS Collaboration, « CMS public luminosity results ». <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. (Cité en pages 28 et 141.)
- [32] L. collaboration. <http://lhcb.web.cern.ch/lhcb/TDR/TDR.htm>. (Cité en page 29.)
- [33] LHCb Collaboration Collaboration, R. Aaij *et al.*, « Strong constraints on the rare decays  $B_s \rightarrow \mu^+ \mu^-$  and  $B^0 \rightarrow \mu^+ \mu^-$  », *Phys.Rev.Lett.* **108** (2012) 231801, [arXiv:1203.4493](https://arxiv.org/abs/1203.4493) [hep-ex]. (Cité en page 29.)
- [34] LHCb collaboration Collaboration, R. Aaij *et al.*, « First observation of CP violation in the decays of Bs mesons », *Phys.Rev.Lett.* **110** (2013) 221601, [arXiv:1304.6173](https://arxiv.org/abs/1304.6173) [hep-ex]. (Cité en page 29.)
- [35] T. L. Collaboration, « The LHCf detector at the CERN Large Hadron Collider », *Journal of Instrumentation* **3** no. 08, (2008) S08006. <http://stacks.iop.org/1748-0221/3/i=08/a=S08006>. (Cité en page 30.)
- [36] T. T. Collaboration, « The TOTEM Experiment at the CERN Large Hadron Collider », *Journal of Instrumentation* **3** no. 08, (2008) S08007. <http://stacks.iop.org/1748-0221/3/i=08/a=S08007>. (Cité en page 30.)
- [37] B. et al., *Total cross-section, elastic scattering and diffraction dissociation at the Large Hadron Collider at CERN : TOTEM Technical Design Report*. Technical Design Report TOTEM. CERN, Geneva, 2004. (Cité en page 30.)
- [38] T. M. collaboration Tech. Rep. CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, Jun, 2009. (Cité en page 30.)

- [39] T. C. Collaboration, *CMS Physics : Technical Design Report Volume 1 : Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. There is an error on cover due to a technical problem for some items. (Cit  en page 33.)
- [40] T. C. Collaboration, « CMS Physics : Technical Design Report Volume 2 : Physics Performance », *J. Phys. G* **34** no. CERN-LHCC-2006-021. CMS-TDR-8-2, (2007) 995–1579. 669 p. revised version submitted on 2006-09-22 17 :44 :47. (Cit  en page 33.)
- [41] D. Green, *The Physics of Particle Detectors*. Cambridge University Press, 2000. (Cit  en page 33.)
- [42] *The CMS magnet project : Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997. (Cit  en page 33.)
- [43] CMS Collaboration Collaboration, S. Taroni, « Performance of the current CMS pixel detector », Tech. Rep. CMS-CR-2013-396. CERN-CMS-CR-2013-396, CERN, Geneva, Nov, 2013. (Cit  en pages 35 et 141.)
- [44] for the CMS collaboration Collaboration, V. Veszpremi, « Operation and performance of the CMS tracker », Tech. Rep. arXiv :1402.0675, Feb, 2014. Comments : accepted for publication in Journal of Instrumentation. (Cit  en pages 36 et 141.)
- [45] CMS Collaboration Collaboration, C. La Licata, « Calibration and Monitoring of the CMS Electromagnetic Calorimeter », Tech. Rep. CMS-CR-2013-264. CERN-CMS-CR-2013-264, CERN, Geneva, Sep, 2013. (Cit  en pages 39 et 141.)
- [46] CMS HCAL Collaboration Collaboration, C. collaboration, « Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters », Tech. Rep. CMS-NOTE-2008-010, CERN, Geneva, Mar, 2008. (Cit  en page 39.)
- [47] CMS HCAL Collaboration Collaboration, C. collaboration, « Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges », Tech. Rep. CMS-NOTE-2006-138, CERN, Geneva, May, 2007. (Cit  en page 39.)
- [48] CMS Collaboration Collaboration, « HCAL performance from first collisions data », (Cit  en page 40.)
- [49] CMS Collaboration Collaboration, D. Benedetti, « Performance of the particle flow algorithm in CMS », Tech. Rep. CMS-CR-2010-127, CERN, Geneva, Jul, 2010. (Cit  en page 48.)
- [50] CMS Collaboration Collaboration, « Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET », Tech. Rep. CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr, 2009. (Cit  en page 48.)
- [51] W. Adam, B. Mangano, T. Speer, and T. Todorov, « Track Reconstruction in the CMS tracker », Tech. Rep. CMS-NOTE-2006-041, CERN, Geneva, Dec, 2006. (Cit  en page 49.)

- [52] E. Meschi, T. Monteiro, C. Seez, and P. Vikas, « Electron Reconstruction in the CMS Electromagnetic Calorimeter », Tech. Rep. CMS-NOTE-2001-034, CERN, Geneva, Jun, 2001. (Cité en page 50.)
- [53] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, « Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC », *Journal of Physics G : Nuclear and Particle Physics* **31** no. 9, (2005) N9. <http://stacks.iop.org/0954-3899/31/i=9/a=N01>. (Cité en page 50.)
- [54] H. Bethe and W. Heitler, « On the Stopping of fast particles and on the creation of positive electrons », *Proc.Roy.Soc.Lond.* **A146** (1934) 83–112. (Cité en page 50.)
- [55] CMS Collaboration Collaboration, S. Chatrchyan *et al.*, « Performance of CMS muon reconstruction in  $pp$  collision events at  $\sqrt{s} = 7$  TeV », *JINST* **7** (2012) P10002, [arXiv:1206.4071](https://arxiv.org/abs/1206.4071) [physics.ins-det]. (Cité en page 51.)
- [56] T. C. collaboration, « Determination of jet energy calibration and transverse momentum resolution in CMS », *Journal of Instrumentation* **6** no. 11, (2011) P11002. <http://stacks.iop.org/1748-0221/6/i=11/a=P11002>. (Cité en page 53.)
- [57] T. C. Collaboration, *CMS computing : Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 2005. Submitted on 31 May 2005. (Cité en page 56.)
- [58] F. James, « Monte Carlo theory and practice », *Reports on Progress in Physics* **43** no. 9, (1980) 1145. <http://stacks.iop.org/0034-4885/43/i=9/a=002>. (Cité en pages 60, 61 et 62.)
- [59] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, « MadGraph 5 : Going Beyond », *JHEP* **1106** (2011) 128, [arXiv:1106.0522](https://arxiv.org/abs/1106.0522) [hep-ph]. (Cité en pages 62 et 72.)
- [60] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, « ALPGEN, a generator for hard multiparton processes in hadronic collisions », *JHEP* **0307** (2003) 001, [arXiv:hep-ph/0206293](https://arxiv.org/abs/hep-ph/0206293) [hep-ph]. (Cité en page 62.)
- [61] T. Sjostrand, L. Lonnblad, and S. Mrenna, « PYTHIA 6.2 : Physics and manual », [arXiv:hep-ph/0108264](https://arxiv.org/abs/hep-ph/0108264) [hep-ph]. (Cité en pages 62, 68, 73 et 88.)
- [62] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, *et al.*, « General-purpose event generators for LHC physics », *Phys.Rept.* **504** (2011) 145–233, [arXiv:1101.2599](https://arxiv.org/abs/1101.2599) [hep-ph]. (Cité en page 64.)
- [63] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, « New Generation of Parton Distributions with Uncertainties from Global QCD Analysis », *Journal of High Energy Physics* **2002** no. 07, (2002) 012. <http://stacks.iop.org/1126-6708/2002/i=07/a=012>. (Cité en pages 64, 66 et 142.)

- [64] J. D. Bjorken, « Asymptotic Sum Rules at Infinite Momentum », *Phys. Rev.* **179** (Mar, 1969) 1547–1553.  
<http://link.aps.org/doi/10.1103/PhysRev.179.1547>. (Cit  en page 64.)
- [65] R. P. Feynman, « Very High-Energy Collisions of Hadrons », *Phys. Rev. Lett.* **23** (Dec, 1969) 1415–1417.  
<http://link.aps.org/doi/10.1103/PhysRevLett.23.1415>. (Cit  en page 64.)
- [66] J. C. Collins, « What exactly is a parton density? », *Acta Phys.Polon.* **B34** (2003) 3103, [arXiv:hep-ph/0304122](https://arxiv.org/abs/hep-ph/0304122) [hep-ph]. (Cit  en page 64.)
- [67] S. I. Alekhin, J. Bl mlein, H. B. B ttcher, L. Del Debbio, S. Forte, A. Glazov, A. Guffanti, J. Huston, G. Ingelman, J. I. Latorre, S. Moch, A. Piccione, J. Pumplin, V. Ravindran, J. Rojo, G. P. Salam, R. S. Thorne, J. A. M. Vermaseren, and A. Vogt, « DGLAP evolution and parton fits », . (Cit  en page 64.)
- [68] A. Martin, W. Stirling, R. Thorne, and G. Watt, « Parton distributions for the LHC », *Eur.Phys.J.* **C63** (2009) 189–285, [arXiv:0901.0002](https://arxiv.org/abs/0901.0002) [hep-ph]. (Cit  en page 66.)
- [69] J. C. Collins, D. E. Soper, and G. F. Sterman, « Factorization for Short Distance Hadron - Hadron Scattering », *Nucl.Phys.* **B261** (1985) 104. (Cit  en page 66.)
- [70] R. Kleiss and R. Pittau, « Weight optimization in multichannel Monte Carlo », *Comput.Phys.Commun.* **83** (1994) 141–146, [arXiv:hep-ph/9405257](https://arxiv.org/abs/hep-ph/9405257) [hep-ph]. (Cit  en page 67.)
- [71] G. Altarelli and G. Parisi, « Asymptotic Freedom in Parton Language », *Nucl.Phys.* **B126** (1977) 298. (Cit  en page 68.)
- [72] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, « Herwig++ Physics and Manual », *Eur.Phys.J.* **C58** (2008) 639–707, [arXiv:0803.0883](https://arxiv.org/abs/0803.0883) [hep-ph]. (Cit  en page 68.)
- [73] S. Catani, F. Krauss, R. Kuhn, and B. Webber, « QCD matrix elements + parton showers », *JHEP* **0111** (2001) 063, [arXiv:hep-ph/0109231](https://arxiv.org/abs/hep-ph/0109231) [hep-ph]. (Cit  en page 68.)
- [74] F. Krauss, « Matrix Elements and Parton Showers in Hadronic Interactions », *Journal of High Energy Physics* **2002** no. 08, (2002) 015.  
<http://stacks.iop.org/1126-6708/2002/i=08/a=015>. (Cit  en page 68.)
- [75] ATLAS Collaboration Collaboration, G. Aad *et al.*, « Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC », *New J.Phys.* **13** (2011) 053033, [arXiv:1012.5104](https://arxiv.org/abs/1012.5104) [hep-ex]. (Cit  en pages 69 et 142.)
- [76] T. Sj strand and M. van Zijl, « A multiple-interaction model for the event structure in hadron collisions », *Phys. Rev. D* **36** (Oct, 1987) 2019–2041.

- <http://link.aps.org/doi/10.1103/PhysRevD.36.2019>. (Cité en page 69.)
- [77] G. Gustafson, « Multiple scattering, underlying event, and minimum bias », [arXiv:0712.1941](https://arxiv.org/abs/0712.1941) [hep-ph]. (Cité en page 69.)
- [78] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, « Parton Fragmentation and String Dynamics », *Phys.Rept.* **97** (1983) 31–145. (Cité en page 70.)
- [79] J.-C. Winter, F. Krauss, and G. Soff, « A modified cluster-hadronisation model », *The European Physical Journal C - Particles and Fields* **36** no. 3, (2004) 381–395. <http://dx.doi.org/10.1140/epjc/s2004-01960-8>. (Cité en page 70.)
- [80] P. Richardson, « Spin correlations in Monte Carlo simulations », *JHEP* **0111** (2001) 029, [arXiv:hep-ph/0110108](https://arxiv.org/abs/hep-ph/0110108) [hep-ph]. (Cité en page 70.)
- [81] S. A. et Al., « Geant4—a simulation toolkit », *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** no. 3, (2003) 250 – 303. <http://www.sciencedirect.com/science/article/pii/S0168900203013688>. (Cité en page 71.)
- [82] S. Abdullin, P. Azzi, F. Beaudette, P. Janot, A. Perrotta, and the CMS Collaboration, « The Fast Simulation of the CMS Detector at LHC », *Journal of Physics : Conference Series* **331** no. 3, (2011) 032049. <http://stacks.iop.org/1742-6596/331/i=3/a=032049>. (Cité en page 71.)
- [83] J. Alwall, A. Ballestrero, P. Bartalini, S. Belov, E. Boos, *et al.*, « A Standard format for Les Houches event files », *Comput.Phys.Commun.* **176** (2007) 300–304, [arXiv:hep-ph/0609017](https://arxiv.org/abs/hep-ph/0609017) [hep-ph]. (Cité en page 72.)
- [84] E. Boos, M. Dobbs, W. Giele, I. Hinchliffe, J. Huston, *et al.*, « Generic user process interface for event generators », [arXiv:hep-ph/0109068](https://arxiv.org/abs/hep-ph/0109068) [hep-ph]. (Cité en page 72.)
- [85] T. Stelzer and W. Long, « Automatic generation of tree level helicity amplitudes », *Comput.Phys.Commun.* **81** (1994) 357–371, [arXiv:hep-ph/9401258](https://arxiv.org/abs/hep-ph/9401258) [hep-ph]. (Cité en page 72.)
- [86] S. Ovnyn, X. Rouby, and V. Lemaitre, « DELPHES, a framework for fast simulation of a generic collider experiment », [arXiv:0903.2225](https://arxiv.org/abs/0903.2225) [hep-ph]. (Cité en page 73.)
- [87] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, *et al.*, « UFO - The Universal FeynRules Output », *Comput.Phys.Commun.* **183** (2012) 1201–1214, [arXiv:1108.2040](https://arxiv.org/abs/1108.2040) [hep-ph]. (Cité en page 73.)
- [88] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, *et al.*, « Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions », *Eur.Phys.J.* **C53** (2008) 473–500, [arXiv:0706.2569](https://arxiv.org/abs/0706.2569) [hep-ph]. (Cité en page 74.)

- [89] N. Lavesson and L. Lonnblad, « Merging parton showers and matrix elements : Back to basics », *JHEP* **0804** (2008) 085, [arXiv:0712.2966](https://arxiv.org/abs/0712.2966) [hep-ph]. (Cité en page 74.)
- [90] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, « Matching matrix elements and shower evolution for top-quark production in hadronic collisions », *JHEP* **0701** (2007) 013, [arXiv:hep-ph/0611129](https://arxiv.org/abs/hep-ph/0611129) [hep-ph]. (Cité en page 74.)
- [91] S. Mrenna and P. Richardson, « Matching matrix elements and parton showers with HERWIG and PYTHIA », *JHEP* **0405** (2004) 040, [arXiv:hep-ph/0312274](https://arxiv.org/abs/hep-ph/0312274) [hep-ph]. (Cité en page 74.)
- [92] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, « Longitudinally invariant  $K_t$  clustering algorithms for hadron hadron collisions », *Nucl.Phys.* **B406** (1993) 187–224. (Cité en page 74.)
- [93] A. Belyaev, C. Leroy, R. Mehdiev, and A. Pukhov, « Leptoquark single and pair production at LHC with CalcHEP/CompHEP in the complete model », *JHEP* **0509** (2005) 005, [arXiv:hep-ph/0502067](https://arxiv.org/abs/hep-ph/0502067) [hep-ph]. (Cité en page 84.)
- [94] J. Blumlein, E. Boos, and A. Kryukov, « Leptoquark pair production cross-sections at hadron colliders », [arXiv:hep-ph/9811271](https://arxiv.org/abs/hep-ph/9811271) [hep-ph]. (Cité en page 85.)
- [95] M. Kramer, T. Plehn, M. Spira, and P. Zerwas, « Pair production of scalar leptoquarks at the CERN LHC », *Phys.Rev.* **D71** (2005) 057503, [arXiv:hep-ph/0411038](https://arxiv.org/abs/hep-ph/0411038) [hep-ph]. (Cité en pages 85 et 86.)
- [96] W. Beenakker, R. Hopker, and M. Spira, « PROSPINO : A Program for the production of supersymmetric particles in next-to-leading order QCD », [arXiv:hep-ph/9611232](https://arxiv.org/abs/hep-ph/9611232) [hep-ph]. (Cité en page 86.)
- [97] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, *et al.*, « New generation of parton distributions with uncertainties from global QCD analysis », *JHEP* **0207** (2002) 012, [arXiv:hep-ph/0201195](https://arxiv.org/abs/hep-ph/0201195) [hep-ph]. (Cité en pages 86 et 88.)
- [98] N. Kidonakis, « NNLL threshold resummation for top-pair and single-top production », [arXiv:1210.7813](https://arxiv.org/abs/1210.7813) [hep-ph]. (Cité en page 93.)
- [99] CMS Collaboration, « Pt(top-quark) based reweighting of ttbar MC ». <https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting>. (Cité en page 94.)
- [100] CMS Collaboration, « MuSclFit : Muon momentum scale calibration and momentum resolution fit ». <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonScaleCalib>. (Cité en page 95.)
- [101] CMS Collaboration, « CMS official 8TeV cross sections ». <https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSectionsat8TeV>. (Cité en pages 106 et 139.)

- 
- [102] CMS Collaboration, « CMS official 7TeVcross sections ». <https://twiki.cern.ch/twiki/bin/viewauth/CMS/StandardModelCrossSections>. (Cité en pages 106 et 139.)
- [103] CMS Collaboration Collaboration, « CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update », Tech. Rep. CMS-PAS-LUM-13-001, CERN, Geneva, 2013. (Cité en page 107.)
- [104] A. L. Read, « Modified frequentist analysis of search results (the  $CL_s$  method) », (Cité en page 107.)
- [105] T. Junk, « Confidence level computation for combining searches with small statistics », *Nucl.Instrum.Meth.* **A434** (1999) 435–443, [arXiv:hep-ex/9902006](https://arxiv.org/abs/hep-ex/9902006) [hep-ex]. (Cité en page 107.)
- [106] J. Neyman and E. S. Pearson, « On the Problem of the Most Efficient Tests of Statistical Hypotheses », *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231** (1933) 289–337. <http://www.jstor.org/stable/91247>. (Cité en page 109.)



# Liste des tableaux

1.1	Les champs fermioniques du Modèle standard et les nombres quantiques associés. $I$ et $I_3$ représentent respectivement l'isospin faible et sa troisième composante ; $Q$ est la charge électrique. . . . .	14
1.2	Les différentes possibilités de leptoquarks et leur nombre quantiques associés. Les colonnes $U(1)_Y$ , $SU(2)_L$ et $SU(3)_c$ représentent respectivement les représentations d'hypercharge, d'isospin et de la QCD. .	18
2.1	Configuration du LHC pour les premières années de fonctionnement.	27
2.2	Luminosité délivrée par le LHC aux 4 expériences principales pendant les 3 premières années de fonctionnement en collisions p-p et Pb-Pb.	30
4.1	Les colonnes 2 et 3 présentent les sections efficaces NLO (en pb) accompagnées des incertitudes dues aux pdf pour les différents échantillons générés. Les colonnes 3 et 4 indiquent les sections efficaces NLO (en pb) calculées en faisant varier l'échelle de renormalisation/factorisation.	86
4.2	Les échantillons simulés du modèle standard utilisés avec leurs sections efficaces, le nombre d'événements générés et la luminosité équivalente.	91
4.3	Les échantillons de données utilisés ainsi que les chemins de déclenchement et les luminosités intégrées correspondantes. . . . .	92
4.4	Nombre d'événements et incertitudes statistiques associées qui passent la sélection dimuons. . . . .	96
4.5	Nombre d'événements et erreurs statistiques associées qui passent la sélection $W+4\text{jets } M_T(W) \geq 60$ . . . . .	105
4.6	Sections efficaces et incertitudes théoriques associées [101] (et [102] quand les incertitudes à 8 TeV n'étaient pas disponibles). . . . .	106
4.7	Incertitudes en % pour les différentes masses de LQ au niveau de sélection optimal. . . . .	108
4.8	Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top $\tau$ ). . . . .	111
4.9	Meilleures limites attendues et observées pour les différentes masses de LQ (LQ en top $\tau$ ). . . . .	112
4.10	Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top $\nu$ ). . . . .	116
4.11	Incertitudes en % pour les différentes masses de LQ (LQ en top $\nu$ ). .	117
4.12	Limites attendues et observées pour différentes masses de LQ (LQ en top $\nu$ ). . . . .	118
4.13	Coupures optimales et nombres d'événements correspondants pour les différentes masses de LQ (LQ en top $\mu$ ). . . . .	123
4.14	Incertitudes en % pour les différentes masses de LQ (LQ en top $\mu$ ). .	124

4.15 Limites attendues et observées pour les différentes masses de LQ (LQ en top $\mu$ ). . . . .	125
--	-----

# Table des figures

1.1	Le potentiel du champs de Higgs. . . . .	7
1.2	Une désintégration possible du proton par l'intermédiaire d'un LQ. . .	17
1.3	Vertex d'interaction gluon avec LQ $\bar{L}Q$ (à gauche) gluon gluon avec LQ $LQ$ (à droite). . . . .	20
2.1	Vue en coupe du LHC ©CERN. . . . .	24
2.2	Schéma du complexe d'accélérateurs du CERN ©CERN. . . . .	26
2.3	Luminosité intégrée enregistrée par l'expérience CMS au cours des trois premières années de fonctionnement [31]. . . . .	28
2.4	Un exemple d'événement à haute multiplicité dans CMS avec les traces associées à chaque vertex. . . . .	28
2.5	Représentation de l'expérience CMS. . . . .	31
2.6	Le système de coordonnées de l'expérience CMS. . . . .	32
2.7	Vue en coupe du détecteur CMS avec l'intensité et les lignes de champ du champ magnétique. . . . .	33
2.8	Coupe longitudinale du trajectographe de l'expérience CMS. . . . .	34
2.9	Quantité de matière dans le trajectographe, exprimée en longueurs de radiation en fonction de $\eta$ [43]. . . . .	35
2.10	Résolution du vertex primaire dans le plan transverse (à gauche) et le long de l'axe des faisceaux (à droite) en fonction du nombre de traces rattachées au vertex [44]. . . . .	36
2.11	Vue longitudinale des parties tonneau, bouchon ainsi que du détecteur de pied de gerbes constituant le calorimètre électromagnétique de CMS. . . . .	37
2.12	à gauche : Réponse relative du ECAL aux rayons laser (447nm en 2012) moyenné sur tous les cristaux pour les périodes de prises de données 2011 et 2012 à droite : Résolution en énergie relative du ECAL par super clusters en $\eta$ obtenue pour des événements $Z \rightarrow ee$ [45]. . . . .	39
2.13	Vue longitudinale du calorimètre Hadronique dans son ensemble, HB, HE, HO et HF. . . . .	40
2.14	Réponse du HCAL pour des traces de 9 à 11 GeV pour 3 régions en $\eta$ . A gauche $ \eta  < 1,1$ le tonneau ; au milieu $1,7 <  \eta  < 2,2$ les bouchons ; A droite $1,1 <  \eta  < 1,7$ la région de transition. . . . .	41
2.15	Coupe longitudinale d'un quart de l'expérience CMS. . . . .	42
2.16	Coupe transverse de l'expérience CMS. . . . .	43
2.17	Vue en coupe des chambres à muons présentant l'agencement des tubes à dérives (DT), des chambres à pistes cathodiques (CSC) et des chambres à plaques résistives (RPC). . . . .	44

2.18	A gauche : vue schématique en coupe de l'agencement des tubes à dérives. A droite : section d'un tube à dérive avec les lignes de dérives et les isochrones. . . . .	45
2.19	Efficacité du déclenchement local des RPC en fonction du $p_t$ (à gauche) et de la pseudorapidité $\eta$ (à droite) quand un segment de trace est identifié dans les DT. . . . .	45
2.20	Architecture du premier niveau de déclenchement pour CMS. . . . .	46
2.21	Principe de fonctionnement de l'algorithme de particle flow. Dans le détecteur, les particules sont vues comme des traces et des dépôts d'énergies. . . . .	48
2.22	Efficacité de reconstruction et facteurs correctifs (data/MC) des muons en fonction du $p_t$ (à droite), de la pseudorapidité (au milieu) et du nombre de vertex (à droite) sur l'ensemble des données 2012. . . . .	52
2.23	Incertitudes sur les corrections d'échelle d'énergie d'un jet de 30 GeV (à gauche) et de 100 GeV (à droite) en fonction de la pseudorapidité $\eta$ . . . . .	54
2.24	Distribution de $\cancel{E}_T$ pour des événements $Z \rightarrow \mu\mu$ (à gauche), $Z \rightarrow ee$ (au milieu) et avec photons (à droite). . . . .	55
2.25	Organisation des différents niveaux de tiers de la grille de calcul CMS. . . . .	56
3.1	Comparaison de la chaîne de génération des événements pour les données et la simulation. . . . .	63
3.2	Représentation d'un événement Monte-Carlo produit par un générateur. L'interaction dure (bulle rouge au centre) est réalisée par l'interaction de deux partons provenant de protons incidents (en vert), puis suivie par la désintégration de 3 partons produits (en rouge). Des radiations ISR (en bleu) et FSR (en rouge) sont ajoutées et une interaction secondaire avec le reste des partons incidents a lieu (bulle violette). On distingue ensuite l'hadronisation des partons produits (bulles vertes), puis la désintégration des hadrons instables (vert foncé). Des radiations de photons sont ajoutées à différentes étapes (en jaune). . . . .	65
3.3	Fonction de densité partonique CTEQ6M [63] pour $Q=2$ GeV (gauche) et $Q=100$ GeV (droite) [63]. . . . .	66
3.4	Comparaison des multiplicités en particules chargées entre les données et différents scénarios pour les simulations dans l'expérience ATLAS [75]. . . . .	69
3.5	Événements MC générés avec MADGRAPH avant et après matching pendant la campagne 2011 à 7 TeV pour des processus $Z$ +jets, $W$ +jets et $t\bar{t}$ . . . . .	77
3.6	Distributions de validation partonique. . . . .	78
3.7	Distributions de validation hadronique. . . . .	78
3.8	Diagrammes DJR pour l'échantillon $t\bar{t}$ réalisé avec MADGRAPH et hadronisé avec PYTHIA pour un $Q_{MS}$ approprié (60 GeV). . . . .	80

3.9	Diagrammes DJR pour l'échantillon $t\bar{t}b\bar{b}$ réalisé avec MADGRAPH et hadronisé avec PYTHIA pour un mauvais $Q_{MS}$ (10 GeV). . . . .	81
4.1	Diagrammes de feynman de production de LQ célibataire au premier ordre en voie s (à gauche) et voie t (à droite). . . . .	84
4.2	Diagrammes de Feynman de production de paires de LQ au premier ordre par annihilation quark anti-quark (a), par voie t (b) et par fusion de gluons (c, d, e, f). . . . .	85
4.3	Section efficace théorique de production de paire de LQ à 8 TeV au LHC. Les incertitudes hautes et basses liées à l'échelle de renormalisation et aux PDF sont respectivement représentées par les bandes de couleurs beige et turquoise. . . . .	87
4.4	Impulsion transverse au niveau généré des LQ (en haut à gauche), des muons (en haut à droite), des jets (en bas à droite) et $\cancel{E}_T$ (en bas à gauche) (dans le cas $LQ \rightarrow t\mu$ . . . . .	89
4.5	Distribution du nombre de vertex dans les données et la simulation après repondération pour une sélection dimuon. . . . .	94
4.6	Masse sans correction MUSCLE fit (à gauche) et avec (à droite) du système dimuons. . . . .	96
4.7	Impulsion transverse des muons sans correction MUSCLEFIT (à gauche) et avec (à droite). . . . .	97
4.8	$\eta$ (en haut et $\phi$ (à bas) des bons muons sans correction MUSCLEFIT (à gauche) et avec (à droite). . . . .	98
4.9	Nombre de jets de plus de 40 GeV (à gauche) et impulsion transverse de tous les bons jets (à droite). . . . .	99
4.10	Distribution de la masse transverse du W avant coupure (à gauche) et après (à droite). . . . .	100
4.11	Distribution de la somme des impulsions transverses des muons et des jets avant coupure en masse transverse du candidat W à 60 GeV (à gauche) et après (à droite). Le signal correspond à des LQ de 300 GeV	101
4.12	Impulsion transverse des muons (en haut) et $\cancel{E}_T$ (en bas) avec les échantillons W+jets et Z+jets inclusif (à gauche) et exclusif (à droite).	102
4.13	Caractéristiques ( $P_t$ en haut et $\eta$ en bas) du muon (à gauche) et du quatrième jet (à droite) utilisés pour calculer l'efficacité du trigger. .	103
4.14	Multiplicité en jets de $p_t > 40\text{GeV}$ (en haut à gauche) $\cancel{E}_T$ (en bas à gauche) et $S_t$ (à droite). . . . .	104
4.15	Multiplicité en jets de $p_t > 40\text{GeV}$ (à gauche) $S_t$ (au milieu) et $\cancel{E}_T$ (à droite) pour la pré-sélection optimale. Le signal correspond ici à une masse de LQ de 400 GeV et une désintégration à 100% en top $\tau$ .	106
4.16	Multiplicité en jets de $p_t > 40\text{GeV}$ (en haut à gauche) $\cancel{E}_T$ (en bas à gauche) et $S_t$ (à droite) la selection optimale correspondant à $m_{LQ} = 30\text{ GeV}$ . . . . .	112
4.17	Limites attendues et observées pour les différentes masses de LQ ainsi que les bandes d'exclusion à 95% et à 68% (LQ en top $\tau$ ). . . . .	113

4.18	$P_t$ des bons jets (à gauche) des muons (au milieu) et $W_{mt}$ (à droite). Le signal correspond ici à une masse de LQ de 500 GeV et une désintégration à 100% en top $\nu$ . . . . .	114
4.19	Multiplicité en jets de $p_t > 40\text{GeV}$ (à gauche) $S_t$ (au milieu) et $\cancel{E}_T$ (à droite) pour la pré-sélection optimale. Le signal correspond ici à une masse de LQ de 500 GeV et une désintégration à 100% en top $\nu$ . . . . .	115
4.20	Limites attendues et observées en fonction de la masse de LQ ainsi que les bandes d'exclusion à 95% et à 68% de confiance (LQ en top $\nu$ ). . . . .	119
4.21	Nombre de bons leptons correspondant à la sélection de l'analyse top $\tau$ (sans la coupure sur $W_{mt}$ ) mais avec un signal top $\mu$ pour $m_{LQ} = 500\text{ GeV}$ . . . . .	120
4.22	En haut, impulsion transverse et pseudo rapidité des muons. En bas, la masse transverse du W (à gauche) et la masse dimuon (à droite) . . . . .	121
4.23	En haut, $p_t$ et $\eta$ des électrons; En bas, nombre de bon muons et d'électrons. . . . .	122
4.24	$S_t$ (à gauche), $p_t$ du muon le plus énergétique (au milieu) et $H_t$ correspondant à la sélection dileptons avec un signal superposé $LQ \rightarrow t\mu$ ( $m_{LQ} = 500\text{GeV}$ ) . . . . .	125
4.25	Limites attendues et observées pour les différentes masses de LQ ainsi que les bandes d'exclusion à 95% et à 68% (LQ en top $\mu$ ). . . . .	126

## Recherche de Leptoquarks de troisième génération dans l'expérience CMS au LHC

---

### Résumé :

L'expérience CMS (Compact Muon Solenoid), construite sur l'anneau du LHC (Large Hadron Collider) a enregistré une quantité colossale de données provenant des collisions proton-proton de 2009 à 2012 à des énergies dans le centre de masse de 7 TeV et 8 TeV. Cette expérience est consacrée aux mesures de précision des paramètres du modèle standard, à l'étude du boson de Higgs ainsi qu'à la mise à l'épreuve de théories au-delà du modèle standard. Un certain nombre de ces théories prévoient l'existence de particules appelées leptoquarks, de nouveaux bosons se couplant à la fois aux leptons et aux quarks.

La première partie de cette thèse est consacrée à la présentation du Modèle Standard des particules et à une introduction théorique aux leptoquarks. La seconde présente brièvement le fonctionnement du LHC et de l'expérience CMS. La troisième partie est dédiée au travail effectué dans le groupe générateur de CMS au cours de la première année de ma thèse. Cette mission consistait principalement à générer et à valider massivement des événements Monte-Carlo avec le logiciel Madgraph pour l'ensemble de la collaboration CMS. Enfin, la dernière partie porte sur l'analyse de données et la recherche de paires de leptoquarks scalaires de troisième génération. Trois canaux de désintégration des LQ sont envisagés : en quark top et lepton tau, en quark top et neutrino et enfin en quark top et muon. Cette étude est basée sur l'ensemble des données enregistrées par l'expérience CMS au cours de l'année 2012. Ce qui représente une luminosité intégrée totale de 19,7 fb. Différentes sélections basées sur au moins un lepton et de nombreux jets sont examinées afin de définir les régions présentant les meilleures efficacités de signal par rapport au fond du Modèle Standard. Enfin, une optimisation des coupures sur des variables très discriminantes est réalisée afin d'obtenir les meilleures limites à 95% de confiance sur les sections efficaces de production des leptoquarks.

Aucun excès significatif par rapport aux prévisions du modèle standard n'a pu être mis en évidence ce qui a permis de poser des limites sur les sections efficaces de production des leptoquarks. Les leptoquarks de troisième génération sont exclus jusqu'à des masses de 320 GeV à 95 % de niveau de confiance pour le canal top tau, jusqu'à 240 GeV pour le canal top nu et jusqu'à 910 GeV pour le canal top mu.

---

**Discipline :** Physique des particules.

---

**Mots clés :** CERN, LHC, expérience CMS, physique des hautes énergies, Modèle Standard, collision proton-proton, leptoquarks, exclusion, limites à 95% de confiance.

---

NOM : CHASSERAT		DATE de SOUTENANCE		
Prénoms : Julien		5 mars 2014		
TITRE : Recherche de leptoquarks de troisième génération dans l'expérience CMS au LHC				
Numéro d'ordre : 33-2014				
DIPLOME DE DOCT.	DOCTEUR- INGENIEUR	DOCTORAT D'ETAT	DOCTORAT DE 3e CYCLE	Spécialité : physique des particules
X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
RESUME :				
<p>L'expérience CMS (Compact Muon Solenoid), construite sur l'anneau du LHC (Large Hadron Collider) a enregistré une quantité colossale de données provenant des collisions proton-proton de 2009 à 2012 à des énergies dans le centre de masse de 7 TeV et 8 TeV. Cette expérience est consacrée aux mesures de précision des paramètres du modèle standard, à l'étude du boson de Higgs ainsi qu'à la mise à l'épreuve de théories au-delà du modèle standard. Un certain nombre de ces théories prévoient l'existence de particules appelées leptoquarks, de nouveaux bosons se couplant à la fois aux leptons et au quarks.</p> <p>La première partie de cette thèse est consacrée à la présentation du Modèle Standard des particules et à une introduction théorique aux leptoquarks. La seconde explique brièvement le fonctionnement du LHC et de l'expérience CMS. La troisième partie est dédiée au travail effectué dans le groupe générateur de CMS au cours de la première année de ma thèse. Cette mission consistait principalement à générer et à valider massivement des événements Monte-Carlo avec le logiciel Madgraph pour l'ensemble de la collaboration CMS. Enfin, la dernière partie porte sur l'analyse de données et la recherche de paires de leptoquarks scalaires de troisième génération. Trois canaux de désintégration des LQ sont envisagés : en quark top et lepton tau, en quark top et neutrino et enfin en quark top et muon. Cette étude est basée sur l'ensemble des données enregistrées par l'expérience CMS au cours de l'année 2012. Ce qui représente une luminosité intégrée totale de 19,7 fb. Différentes sélections basées sur au moins un lepton et de nombreux jets sont examinées afin de définir les régions présentant les meilleures efficacités de signal par rapport au fond du Modèle Standard. Enfin, une optimisation des coupures sur des variables très discriminantes est réalisée afin d'obtenir les meilleures limites à 95 % de confiance sur les sections efficaces de production des leptoquarks.</p> <p>Aucun excès significatif par rapport aux prévisions du modèle standard n'a pu être mis en évidence ce qui a permis de poser des limites sur les sections efficaces de production des leptoquarks. Les leptoquarks de troisième génération sont exclus jusqu'à des masses de 320 GeV à 95 % de niveau de confiance pour le canal top tau, jusqu'à 240 GeV pour le canal top nu et jusqu'à 910 GeV pour le canal top mu.</p>				
MOTS-CLES : CERN, LHC, expérience CMS, physique des hautes énergies, Modèle Standard, collision proton-proton, leptoquarks, exclusion, limites à 95% de confiance				
Directeur de recherches : Patrice Verdier				
Composition du jury : I. Laktineh, M. Jaffré, C. Royon, S. Davidson, S. Muanza, P. Verdier			144 pages	