THÈSE

présentée pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ LOUIS PASTEUR STRASBOURG I

Spécialité : Physique des Particules

présentée par :

Vincent SICCARDI

Etude du couplage top-W-b par la mesure de l'hélicité du boson W dans l'expérience DØ

Soutenue le 11 juillet 2008 devant la commission d'examen :

M.	Jorgen D'HONDT	Rapporteur externe
M.	Ulrich GOERLACH	Rapporteur interne
M.	Eric KAJFASZ	Examinateur
M.	Dominique PALLIN	Rapporteur externe
Mme.	Isabelle RIPP-BAUDOT	Directeur de thèse

Remerciements

"J'y gagne, dit le renard, à cause de la couleur du blé."
Antoine de Saint-Exupéry, *Le Petit Prince.*

Voilà sans doute la partie la plus difficile d'une rédaction de thèse... les remerciements, où comment faire passer en quelques mots l'importance qu'ont eu tant de personnes durant toutes ces années... et surtout... n'oublier personne... Alors autant commencer par tout ceux que j'oublierai : merci à vous, pour tout !

Merci donc, en tout premier lieu, à toi Isabelle Ripp-Baudot. Merci Isa de m'avoir soutenu, de t'être occupé de moi toutes ces années. Merci de m'avoir pris sous ton aile, depuis ce stage de fin d'études il y a... bien longtemps déjà. Je doute retrouver un jour un "chef" avec qui parler BD tout en le conduisant sous une tempête de neige manger un bon gros hamburger américain...
Je tiens également à remercier sincèrement Jorgen D'Hondt, Dominique Pallin et Ulrich Goerlach pour avoir accepté d'être rapporteurs et jury de cette thèse. Merci de l'intérêt que vous y avez porté. Et bien sûr un grand merci à Eric Kajfasz de m'avoir fait le grand plaisir de compléter ce jury !
Quelques mois de cette thèse furent passés à Chicago, sous la direction d'Andrzej Zieminski, une personne d'une gentillesse et d'une prévenance exceptionnelles. Parti bien trop vite, je suis heureux d'avoir pu croiser sa route, et lui dédie ce mémoire.

Un grand merci à tout mes "collègues" (comme on dit chez moi) du couloir, de "DØ Strasbourg", et autres partenaires de cantine et de cafés, dans le désordre complet, Arnaud Gay, Sébastien Greder (de retour, enfin ! Dépêche toi de trouver le Higgs, tant que je signe encore !), Anne-Catherine Lebihan (j'te verrai plus danser sur les tables...ni parler au frigo !), Denis Gelé, Daniel Bloch (un grand merci pour ton soutien, et ta maîtrise du b-tagging !), Benoit Clément (mon voisin des premières années, qui code ses idées plus vite que son ombre), Jean-Laurent Agram, Jérôme Baudot, Walter Geist, Pierre Van Hove, Joaquim Speck (l'hélicité du W compte sur toi !). Une pensée toute particulière pour François Charles, qui nous a quitté bien trop tôt... Merci aussi à Jean-Pierre Engel, Jean-Pierre Froberger, Roger Strub et Jean Schuller...

Merci à tous les copains thésards de DØ, merci d'avoir rendu si vivantes nos réunions DØ-France, et merci surtout d'avoir rendu la survie en milieu hostile finalement si agréable... Fermilab, sous la neige, sous la chaleur suffocante, l'Outback, les Dorms, le User Center et ces pizzas-bières-billard-ping-pong, le Wilson Hall, mes amis les Bisons, les Gooses, la security de Fermilab ("Thomas, tu penses que la sécurité a mis la sirène pour moi ?"), Chicago, The Bean, le Golden Mile, le lac Michigan, mes potes les singes du Lincoln Park Zoo, l'Art Institut, les musées, les Sears et Hancock Tower, The Loop et... et... et... et tous ces souvenirs, engrangés seul, ou avec vous, qui ont fait de ces Lieux, de cette Ville, ma seconde maison, et dont les souvenirs resteront pour toujours gravés dans ma mémoire... Dans le désordre total, merci à vous, Marine, Thomas (que de souvenirs avec vous deux, lors de l'école d'été du CERN !), Sam (on se fait un mexicain ?), Bertrand, Marion, Christophe (tu m'préviens si tu trouves un "Hooters" à Barcelona ?), Nikolas, Anne-Fleur (la Lemon Head Team toujours vivra... je compte sur toi pour propager la bonne parole ! Bienvenue voisine en terre étrangère !), Florent (enfin un gars d'Clermont qu'on pouvait voir !), Florent, Vincent, Anne-Marie, Aurélien, Jean-Roch, Jérémie, Philippe, Fabrice, François, Alexandre, Emmanuel, Pierre-Simon, Mathieu, Fabrice, Jovan, et ceux que j'espère ne pas avoir oubliés... quand vous voulez pour un New York Strip very very well done au Rock Bottom, un cheesecake à Chicago ou des saucisses au fromage (no peanuts no garlic) ! Et bravo à vous d'avoir supporté mon humour... faudra qu'on se refasse un UNO Bob l'éponge d'ailleurs, j'ai une revanche à prendre ! Merci aussi à vous les permanents, Eric, Slava (et sa Vodka à 90 ° qui aura marqué tout le monde), Gerald, Steve (laissez le vous menez écouter du blues à Chicago !), Jean-François, Frédéric, Christophe, Gregorio, Pierre, Yannick, Gérard, Jan, Patrice, Arnaud, Laurent et j'en passe...

Ces quatre années de thèse ont aussi été quatre années d'enseignement à l'ULP... une facette du métier particulièrement enrichissante... et source de nombreuses rencontres qui m'auront marquées (non, pas les étudiantes -_-')... Patrick Lévêque, Mounir Maaloum, Olivier Bengone, Hugo Faivre et bien sûr tout particulièrement Cécile Jollet... J'ai eu l'occasion de côtoyer bien d'autres personnes, au cours des remplacements divers et variés, bon courage à vous tous pour supporter les promotions à venir ! Et merci aussi aux étudiants de m'avoir supporté !

Et puis, il y a les copains du labo et du DEA : Tatjana (je ne peux décemment imprimer pour la postérité ton petit surnom !), Ludo (3 gamins en 1 thèse, qui dit mieux ?), Nathalie, Didier (A quand la prochaine soirée pizza-bière ? Tu verras, un jour je passerai une medium ! On peut être fier, mec, soutenir en juillet, c'est pas donné à tout le monde ! Bon courage pour la suite !), Goulven, Vincent, Marie-Delphine, Auguste, Alex, Hélène, Virgile, Mathieu, Arnaud (tu vas la finir ouais ?), Khadija, Christina, Eric, Jonathan... Merci pour ces soirées Barbeuk, ces cafés, ces discussions au détour d'un couloir, ces stands à la Fête de la Science et autres plaintes de fin de thèse et râlages sur ce cher CIES... Merci pour toutes ces amitiées liées. Nos chemins divergent, mais j'espère bien

vous croiser encore !

Enfin, il y a ceux que j'en oublierai presque les avoir connus par le labo, vue l'importance qu'ils ont pris : Sam et Delphine, Anto et Isabelle, Carine et bien sûr, Jérémy... Heureusement que t'es arrivé ! Et désespère pas, tu en auras un jour, des données ! Et évidemment, la Maman de Noisette, merci à toi pour tout, le reste, et plus encore...

Une pensée également pour tout ces amis, rencontrés lors des quelques années "PS", et qui, petit à petit, ont quitté les lieux... mais avec qui le contact est toujours resté, aussi fort... Lise (merci d'avoir toujours été là quand il le fallait...), Rémy, Damien, Tristan (on part où les gars la prochaine fois ? Pérou ? Mongolie ?), Amandine, Nanthilde, Elodie et Jonathan (bon courage à vous deux au pays du Burger King !), Mathieu, Matthieu, Thibaud, Marion, Thibaut, Jean-luc, Bruno, Fedoua, Adrian, Nico, Benoir, Bengod, Scribus, Nicolas, Laure, Bastien, Stan, Gogo, Tudi, Aurélie, Loïc, Antoine, Clémence... Trop de choses à dire sur chacun. J'espère encore plus à dire des années à venir...

Et mes amis d'enfance, Julien, Thierry, Thomas... on s'est vu trop peu durant toutes ses années, espérons que la tendance s'inverse ! Me souvenir de Martigues me fait penser que peut-être, sans Dominique Michel, de thèse il n'y aurait jamais eu ne serait-ce que l'idée !

La fin de thèse a sans doute été "facilitée" par ma reprise du karaté... merci aux gens cotoyés au club, merci à Gérard, Jacquy et Ghoerghe, et à Cécile pour m'avoir poussé à m'y remettre !

Merci également à Merrick et à tous les gens croisés durant ces 5 ans d'aventures... merci en particulier à toi Carole, pour tout... une pensée pour Lameth, Arayel, Arno, Arnould, Dwight, Blade, Az', Arnould, Thor Ayerdhal, Thoorriin, Leohan, Marcel, Couala et tout les autres. Khrô le bô, Nightwish et Lhaurana vous saluent.

Et bien évidemment, je remercie du fond du cœur mes parents et ma ("p'tite") sœur pour leur soutien indéfectible et leur présence constante à mes côtés... je regrette d'être trop peu souvent rentré à la maison, fils indigne que je suis !

Table des matières

Introduction **1**

1 Cadre théorique **3**
1.1 Le Modèle Standard de la physique des particules . . . 3
1.1.1 Les particules de matière . . . 4
1.1.2 Les interactions . . . 4
1.1.2.1 Description des interactions dans le Modèle Standard . . 5
1.1.2.2 L'interaction forte . . . 6
1.1.2.3 L'interaction électrofaible . . . 7
1.1.3 Le mécanisme de Higgs . . . 8
1.2 Au-delà du Modèle Standard . . . 11
1.2.1 Succès et insuffisances du Modèle Standard . . . 11
1.2.2 Les théories au-delà du Modèle Standard et le quark top . . . 14
1.3 Le quark top . . . 16
1.3.1 Caractéristiques du quark top . . . 16
1.3.2 Production du quark top au Tevatron . . . 18
1.3.2.1 Production forte du quark top . . . 18
1.3.2.2 Production faible du quark top . . . 18
1.3.3 Désintégration du quark top et d'une paire de quarks top . . . 21

2 Le Tevatron et l'expérience DØ **23**
2.1 Le complexe de production et d'accélération du Tevatron . . . 24
2.1.1 Production du faisceau de protons . . . 24
2.1.1.1 Le Pré-accélérateur . . . 24
2.1.1.2 Le Linac . . . 24
2.1.1.3 Le Booster . . . 26
2.1.2 L'Injecteur Principal . . . 26
2.1.3 Production du faisceau d'antiprotons . . . 26
2.1.3.1 La Station-Cible . . . 26
2.1.3.2 Le Debuncher . . . 27
2.1.3.3 L'Accumulateur . . . 27
2.1.3.4 Le Recycleur . . . 27

2.1.4 Le Tevatron . . . 28
2.2 Le détecteur DØ . . . 30
2.2.1 Le système de coordonnées . . . 30
2.2.2 Le système de reconstruction des traces et des vertex . . . 33
2.2.2.1 Le détecteur de vertex en silicium SMT (*Silicon Microstrip Tracker)* . . . 33
2.2.2.2 Le détecteur de traces à fibres scintillantes CFT (*Central Fiber Tracker*) . . . 36
2.2.2.3 L'aimant solénoïdal . . . 37
2.2.2.4 Performances du trajectographe . . . 38
2.2.2.5 Les détecteurs de pieds de gerbes . . . 38
2.2.3 Le calorimètre . . . 39
2.2.3.1 Granularité . . . 42
2.2.3.2 Performances du calorimètre . . . 44
2.2.3.3 Le Détecteur Inter-Cryostat . . . 44
2.2.4 Le spectromètre à muons . . . 45
2.2.4.1 Les aimants toroïdaux . . . 45
2.2.4.2 Le détecteur central . . . 45
2.2.4.3 Les détecteurs avant et arrière : FAMUS . . . 49
2.2.4.4 Le blindage . . . 50
2.2.5 La reconstruction des muons dans le spectromètre central . . . 50
2.2.5.1 La détermination de la distance de dérive dans les PDT 51
2.2.5.2 Détermination des T_0 et procédures de contrôle de l'alignement . . . 55
2.2.6 Le détecteur de protons avant . . . 61
2.2.7 La luminosité . . . 61
2.3 Le système de déclenchement . . . 65
2.3.1 Le niveau 1 . . . 65
2.3.2 Le niveau 2 . . . 66
2.3.3 Le niveau 3 . . . 67
2.4 Reconstruction des données . . . 69
2.5 Simulation des événements . . . 69
2.5.1 Générateurs . . . 70
2.5.2 Simulation du détecteur et des interactions au Tevatron . . . 70
2.6 Reconstruction et identification hors ligne des objets . . . 71
2.6.1 Les traces . . . 71
2.6.2 Le vertex primaire . . . 72
2.6.2.1 L'algorithme *2-Pass Primary Vertex* . . . 73
2.6.2.2 L'algorithme *Adaptive Primary Vertex* . . . 73
2.6.3 Les jets . . . 73
2.6.3.1 Reconstruction . . . 74
2.6.3.2 Identification . . . 75

2.6.3.3 Correction . . . 75
2.6.3.4 Correction spécifique à la simulation : JSSR . . . 76
2.6.4 Les objets électromagnétiques . . . 77
2.6.4.1 Reconstruction des objets électromagnétiques . . . 77
2.6.4.2 Identification et discrimination . . . 77
2.6.4.3 Définition des électrons *tight* . . . 78
2.6.5 Muons . . . 79
2.6.5.1 Reconstruction des muons . . . 79
2.6.5.2 Identification des muons . . . 79
2.6.5.3 Définition des muons *tight* . . . 81
2.6.6 L'énergie transverse manquante . . . 81

3 L'identification des jets de quarks beaux 83
3.1 Identification des jets de quarks beaux par l'algorithme JLIP . . . 84
3.1.1 Lots de données et de simulation utilisés pour la calibration . . . 84
3.1.1.1 Données et simulations p17 . . . 84
3.1.1.2 Données et simulations p20 . . . 85
3.1.2 Sélection des événements et des objets . . . 86
3.1.2.1 Critères de déclenchement . . . 86
3.1.2.2 Vertex primaire . . . 86
3.1.2.3 Muons . . . 86
3.1.2.4 Sélection des jets et des traces et notion de *taggabilité* . . . 87
3.1.2.5 Réjection des V^0 . . . 88
3.1.3 Définition du paramètre d'impact d'une trace . . . 90
3.1.4 Calibration de l'erreur sur le paramètre d'impact . . . 93
3.1.5 Description de l'algorithme d'identification JLIP . . . 95
3.1.6 Détermination des fonctions de résolution . . . 96
3.1.7 Estimation des performances de l'identification des jets beaux par l'algorithme JLIP . . . 97
3.1.7.1 Efficacité d'identification dans la simulation Monte-Carlo 99
3.1.7.2 Efficacité d'identification dans les données . . . 99
3.1.7.3 Estimation du taux de mauvaise identification dans les données . . . 101
3.2 Identification des jets de quarks beaux par un réseau de neurones . . . 118
3.2.1 Les algorithmes SVT et CSIP . . . 118
3.2.1.1 L'algorithme d'identification SVT : Secondary Vertex Tag . . . 118
3.2.1.2 L'algorithme d'identification CSIP : Counting Signed Impact Parameter . . . 119
3.2.2 Identification des jets de quarks beaux par un réseau de neurones 120
3.2.3 Evolution future de l'algorithme d'identification des jets beaux dans DØ . . . 121

4 Sélection des paires de quarks top **127**
4.1 Signal et bruit de fond . . . 128
4.1.1 Le signal . . . 128
4.1.2 Les processus de bruit de fond . . . 129
4.1.2.1 Le bruit de fond multijets QCD . . . 129
4.1.2.2 Le bruit de fond W+jets . . . 130
4.1.2.3 Les autres bruits de fond . . . 131
4.2 Sélection des événements . . . 131
4.2.1 Les données . . . 131
4.2.2 Qualité des événements . . . 132
4.2.3 Critères de déclenchement . . . 133
4.2.3.1 Canal électron+jets . . . 133
4.2.3.2 Canal muon+jets . . . 135
4.2.3.3 Simulation du déclenchement . . . 136
4.2.4 Sélection du Vertex Primaire . . . 136
4.2.5 Sélection des jets . . . 137
4.2.6 Sélection des leptons . . . 137
4.2.6.1 Canal e+jets . . . 137
4.2.6.2 Canal mu+jets . . . 138
4.2.7 Energie transverse manquante . . . 138
4.2.8 Coupures triangulaires . . . 138
4.2.9 Identification des jets de quarks beaux . . . 139
4.2.10 L'ajustement cinématique contraint . . . 140
4.2.11 Bilan de la sélection . . . 142
4.3 Détermination des contributions du signal $t\bar{t}$ et des bruits de fond . . . 142
4.3.1 Contribution du bruit de fond QCD multijets . . . 143
4.3.1.1 Principe . . . 143
4.3.1.2 Calcul des erreurs associées . . . 145
4.3.1.3 Détermination des efficacités ϵ_{sig} et ϵ_{qcd} . . . 145
4.3.1.4 Mesure de la proportion de QCD . . . 146
4.3.2 Fonction discriminante . . . 146
4.3.2.1 Définition du discriminant . . . 146
4.3.2.2 Variables étudiées pour construire le discriminant . . . 148
4.3.2.3 Choix des variables du discriminant . . . 150
4.3.2.4 Point de fonctionnement du discriminant et du réseau de neurones . . . 158
4.3.3 Contribution du bruit de fond W+jets . . . 159
4.3.4 Contrôle de la description des données par la simulation . . . 164

5 Mesure de l'hélicité du boson W 169
5.1 Estimateurs des fractions d'hélicité . . . 169
5.1.1 Structure du vertex top-W-b et hélicité du W . . . 169
5.1.2 Les différents estimateurs . . . 170
5.2 Mesure de l'hélicité du boson W . . . 174
5.2.1 Principe . . . 174
5.2.2 Calibration de la méthode . . . 177
5.2.3 Mesure dans les données . . . 180
5.2.3.1 Fraction d'hélicité longitudinale f_0 . . . 180
5.2.3.2 Fraction d'hélicité droite f_+ . . . 180
5.3 Etude des erreurs systématiques . . . 183
5.3.1 Principe de la mesure des erreurs . . . 183
5.3.2 Estimation des erreurs systématiques . . . 185
5.3.2.1 Sources d'erreurs systématiques . . . 185
5.3.2.2 Bilan des erreurs systématiques . . . 188
5.4 Estimation d'une limite supérieure sur la fraction d'hélicité droite . . . 189
5.5 Conclusion . . . 189

Conclusion 193

Bibliographie 195

Introduction

Le travail présenté dans ce mémoire de thèse s'intéresse aux propriétés du quark top, plus particulièrement à la structure de son couplage au boson vecteur W^+ de l'interaction faible chargée. Cette étude est basée sur l'analyse des collisions proton-antiproton qui ont été délivrées par le Tevatron du Fermilab entre 2002 et 2006 et enregistrées par l'expérience DØ.

Le Tevatron est actuellement le collisionneur le plus puissant en fonctionnement et le seul capable de produire des quarks top. Il a connu une première période de fonctionnement, de 1992 à 1996 (*Run I*), durant laquelle une luminositée intégrée de 120 pb^{-1} a été délivrée par expérience avec une énergie dans le centre de masse de 1.8 TeV. Ces données ont permis la découverte en 1995 du dernier des six quarks connus, le quark top. Depuis 2001, une deuxième phase de fonctionnement est en cours, appelée le *Run II*, pour laquelle des améliorations substantielles ont été apportées au collisionneur et à ses deux expériences. La luminosité instantanée a été augmentée d'un ordre de grandeur et l'énergie de collision est passée à 1.96 TeV, permettant de produire plusieurs dizaines de fois plus de quarks top qu'au *Run I*. Après avoir confirmé l'existence de ce quark, le Run II permet maintenant d'aborder sa caractérisation précise. Sa masse très élevée singularise ce quark qui pourrait ouvrir une fenêtre sur des processus physiques nouveaux. Le collisionneur proton-proton LHC du CERN devrait rentrer en fonction avant la fin de cette année (2008) et délivrer rapidement des collisions à 14 TeV. Avec une section efficace de production de quarks top cent fois plus élevée que le Tevatron, ce nouveau collisionneur sera une usine à top. Plusieurs millions de paires de quarks top seront ainsi produits. Le quark top jouera le rôle de chandelle standard dans les tous premiers jours du LHC, pour devenir ensuite rapidement le principal bruit de fond de nombreuses analyses mais peut-être aussi le révélateur de nouveaux processus.

Le premier chapitre de ce mémoire de thèse introduit les principaux éléments du modèle standard de la physique des particules pour expliquer les motivations de la mesure que nous avons effectuée. Les propriétés du quark top sont ensuite présentées ainsi que ses modes de production au Tevatron.

La description du dispositif expérimental fait l'objet du deuxième chapitre : le complexe d'accélération des protons et des antiprotons, les composantes du détecteur DØ et enfin quelques explications concernant la reconstruction et l'identification des objets à partir des données enregistrées par le détecteur.

Le troisième chapitre est consacré à l'algorithme d'identification des jets issus de l'hadronisation d'un quark b. L'identification des jets beaux permet de sélectionner efficacement des quarks top. Le groupe DØ de l'IPHC Strasbourg, au sein duquel j'ai effectué mon travail de thèse, a développé un tel algorithme et j'ai participé à la procédure de validation permettant son utilisation pour l'ensemble des analyses menées dans DØ.

Les deux derniers chapitres présentent la mesure de l'hélicité du boson W issu de la désintégration du top que j'ai effectuée avec 1 fb^{-1} de données. Le quatrième chapitre traite la sélection des collisions susceptibles de contenir des quarks top puis la quantification des différentes contributions de signal et de bruit de fond à ces candidats sélectionnés. Enfin le cinquième chapitre détaille la procédure menant à la mesure, sa validation et l'étude des incertitudes statistiques et systématiques.

1
Cadre théorique

Dans ce chapitre nous replaçons notre travail expérimental dans son contexte théorique pour exposer les motivations menant à l'étude du couplage tWb au Tevatron. Dans une première section nous présentons le modèle standard de la physique des particules puis nous nous focalisons dans les sections suivantes sur la physique au-delà du modèle standard et le quark top. Nous discutons les propriétés du quark top, ses voies de production au Tevatron et son mode de désintégration.

1.1 Le Modèle Standard de la physique des particules

Le cadre théorique du modèle standard est la théorie quantique des champs qui permet de décrire les interactions fondamentales entre les constituants élémentaires de la matière en respectant les principes de la relativité restreinte et ceux de la mécanique quantique. Ses prédictions ont été testées avec succès avec une précision de l'ordre du ‰. Le développement du formalisme du modèle standard a commencé dans les années 1960 avec les travaux de A. Salam [1], S.L. Glashow [2] et S. Weinberg [3], qui ont unifié les interactions électromagnétique et faible. Au même moment l'hypothèse des quarks est proposée par M. Gell-Mann [4], Y. Ne'eman [5] et G. Zweig [6]. Les quarks sont des particules interagissant par interaction forte et portant un nouveau nombre quantique appelé la couleur. Dans les années 1970 une théorie de jauge est développée par D.J. Gross, F. Wilczek [7] et H.D. Politzer [8] pour décrire l'interaction entre les quarks, la ChromoDynamique Quantique (QCD). La combinaison de tous ces formalismes débouche sur le modèle standard qui est une théorie de jauge basée sur l'invariance sous les symétries $SU(2)_L \times U(1)_Y \times SU(3)_C$.

1.1.1 Les particules de matière

On désigne par particules élémentaires les particules les plus fondamentales connues à ce jour. Les particules de matière suivent une statistique de Fermi et obéissent au principe d'exclusion de Pauli. On distingue les quarks, sensibles à toutes les interactions et les leptons qui ne subissent pas l'interaction forte. La désintégration β et la désintégration du neutron par interaction faible chargée ont suggéré que le neutrino électronique et l'électron sont deux états d'une même particule et par suite leur classification en doublets d'isospin faible. Puis suite à l'observation par C.S. Wu [9] de la violation de la parité par l'interaction faible, une distinction entre composantes de chiralité gauche (indice L) et droite (indice R) a été introduite. Dans le cadre du modèle standard, seuls les fermions gauches et les anti-fermions droits sont sensibles à l'interaction faible. Nous reviendrons sur cette propriété très importante pour le travail présenté dans ce mémoire de thèse dans la section 1.1.2.3.

Dans le modèle standard les neutrinos sont décrits par des particules de masse strictement nulle, dont la chiralité se confond avec l'hélicité. L'expérience de Goldhaber [10] a montré que seuls les neutrinos d'hélicité négative sont observés dans la nature. Nous savons depuis l'observation des oscillations des neutrinos en 1998 [11] que les neutrinos ont une masse. Leurs masses ne sont pas encore mesurées mais sont certainement très faibles et le modèle standard reste correct au premier ordre.

Expérimentalement on a observé 12 fermions et 12 anti-fermions fondamentaux. Ces fermions constituent trois familles de deux doublets d'isospin. Ces familles sont formellement complètement identiques. La seule différence entre elles est la masse des particules qui va croissant de la première à la troisième famille.

première famille :	$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	e^-_R	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	u_R	d_R
deuxième famille :	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	μ^-_R	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	c_R	s_R
troisième famille :	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	τ^-_R	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	t_R	b_R

La matière qui nous constitue et nous entoure est formée uniquement de particules appartenant à la première famille : les quarks u et d composent les nucléons, qui eux-mêmes composent les atomes quand on ajoute les électrons. Les autres particules sont produites par des interactions entre particules de la première famille dans les collisionneurs de particules et les rayons cosmiques.

1.1.2 Les interactions

Les interactions entre particules de matière sont dues à l'échange de particules qui obéissent à la statistique de Bose, on les appelle des bosons. Elles sont au nombre de quatre : la faible, véhiculée par les bosons massifs W^+, W^- et Z^0, la forte, véhiculée par 8 gluons de masse nulle, l'électromagnétique, véhiculée par le photon et enfin

la gravitation qui serait portée par un graviton de masse nulle. Tous ces bosons sont vecteurs (de spin 1), sauf le graviton qui aurait un spin 2. Ce dernier boson n'a pas encore été observé et la gravitation n'est pas décrite dans le cadre du modèle standard de la physique des particules parce qu'il n'y a pas de théorie quantique de la gravitation. Cette non-description de la gravitation ne pose pas de problème à l'échelle des particules élémentaires pour laquelle cette interaction est totalement négligeable.

Les théories physiques sont basées sur la considération d'invariance des processus sous certaines transformations, par exemple des transformations continues de translation dans le temps, dans l'espace et de rotation. De ces symétries découlent des conservations de grandeurs physiques, respectivement l'énergie, la quantité de mouvement et le spin. De la même manière dans la théorie décrivant la nature à l'échelle des particules élémentaires nous voulons que les lois physiques restent inchangées sous l'introduction d'une phase. En théorie des champs, de telles transformations agissant uniformément dans tout l'espace sont appellées des transformations de jauge globales. Les transformations dépendant des coordonnées d'espace-temps sont, elles, qualifiées de transformations de jauge locales. Nous allons voir en prenant l'exemple de l'électrodynamique quantique (QED) comment les interactions apparaissent dans le modèle standard dès lors que l'on exige l'invariance du lagrangien sous des transformations de jauges locales.

1.1.2.1 Description des interactions dans le Modèle Standard

Considérons une particule libre de spin $\frac{1}{2}$, de masse m et de charge électrique e. Elle est décrite par un spineur de Dirac $\psi(x)$ et son lagrangien s'écrit :

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x)$$

En appliquant la transformation globale de jauge suivante :

$$\psi(x) \rightarrow e^{i\alpha}\psi(x)$$

le lagrangien reste invariant, en effet les phases s'annulent :

$$\bar{\psi}(x) \rightarrow e^{-i\alpha}\bar{\psi}(x)$$

$$\partial_\mu\psi(x) \rightarrow e^{i\alpha}\partial_\mu\psi(x)$$

Il en découle la conservation globale de la charge électrique.

Maintenant si l'on applique une transformation locale de jauge, dans laquelle la phase $\alpha = \alpha(x)$ dépend des coordonnées d'espace-temps x^μ, des nouveaux termes apparaissent dans le lagrangien du fait de la dérivée ∂_μ :

$$\partial_\mu\psi(x) \rightarrow e^{i\alpha(x)}\partial_\mu\psi(x) + i(\partial_\mu\alpha(x))e^{i\alpha(x)}\psi(x)$$

Pour garder le lagrangien invariant sous une transformation locale de jauge, il faut compenser ces termes en remplaçant la dérivée ∂_μ par une dérivée covariante D_μ, de telle manière qu'on ait la transformation suivante :

$$D_\mu \psi(x) \rightarrow e^{i\alpha(x)} D_\mu \psi(x)$$

Ceci est obtenu avec la dérivée modifiée : $D_\mu = \partial_\mu - i\alpha(x)A_\mu$.

Ainsi l'invariance de jauge locale du lagrangien est obtenue en introduisant un nouveau champ vectoriel A_μ, associé au champ électromagnétique. La transformation considérée ici est décrite par le groupe de symétrie $U(1)$ des transformations unitaires à une dimension, qui comporte un seul générateur correspondant au boson de jauge de masse nulle et de spin 1, le photon. Il apparaît un couplage entre le fermion, qui était libre au départ, et le champ électromagnétique, décrit par le terme $e\bar{\psi}(x)\gamma^\mu A_\mu \psi(x)$ avec $\alpha(x) = e$.

1.1.2.2 L'interaction forte

La découverte des baryons Δ^{++}, Δ^- et Ω, composés des trois quarks de même saveur (respectivement *uuu*, *ddd* et *sss*) entraîne l'introduction d'un nouveau nombre quantique pour différencier les quarks et respecter ainsi le Principe d'exclusion de Pauli qui s'applique aux fermions. Ce nombre quantique supplémentaire est appelé la charge de couleur. Il y a trois couleurs, qui sont par convention appelées Rouge, Vert et Bleu.

La conservation de la couleur par l'interaction forte est décrite par la ChromoDynamique Quantique (QCD) et le groupe associé est $SU(3)$, qui possède 8 générateurs. La dérivée covariante D_μ introduite pour conserver l'invariance locale sous $SU(3)_C$ (l'indice C rappelle la conservation de la couleur) fait apparaître 8 bosons de jauge, les gluons G^a_μ :

$$D_\mu = \partial_\mu + ig_s \lambda_a G^a_\mu$$

où λ_a sont les 8 générateurs de $SU(3)$ et g_s est la constante de l'interaction forte.

La constante de couplage fort $\alpha_s = g_s/4\pi$ évolue avec l'énergie :

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{4\pi}\ b_0\ \ln\left(\frac{Q^2}{\mu^2}\right)}$$

où Q^2 est le moment transféré, μ est une échelle d'énergie fixée, N_f vaut le nombre de saveurs, N_C le nombre de couleurs et $b_0 = -\frac{2}{3}N_f + \frac{11}{3}N_C = 7$. Cette évolution permet de décrire deux caractéristiques de l'interaction forte :

- **la liberté asymptotique :** à courte distance, ou encore à grand moment transféré, la constante de couplage tend vers 0. L'intensité de l'interaction forte diminue lorsque la distance entre deux quarks diminue et deux quarks infiniment proches sont asymptotiquement libres.

– **le confinement :** à l'inverse, lorsque la distance entre deux quarks augmente, ou encore lorsque le moment transféré devient faible, la constante de couplage devient grande. Dans ce cas la théorie n'est plus calculable par un développement perturbatif en puissances de la constante de couplage, décrit par les diagrammes de Feynman. L'intensité de l'interaction forte augmente lorsque la distance entre deux quarks augmente. Ainsi lorsque deux quarks s'éloignent, une nouvelle paire quark-antiquark est produite à partir du vide de QCD et ces nouveaux quarks vont se lier avec les deux quarks de départ pour former des hadrons, objets sans couleur. Ce phénomène d'hadronisation est à l'origine des jets de particules observés dans les détecteurs.

1.1.2.3 L'interaction électrofaible

Toutes les observations expérimentales ont montré jusqu'à présent que le boson W ne se couple qu'aux fermions de chiralité gauche. Pour un fermion décrit par le spineur $\psi(x)$, les projections de chiralité gauche $\psi_L(x)$ et droite $\psi_R(x)$ sont définies par :

$$\psi(x) = \psi_L(x) + \psi_R(x), \text{ avec } \psi_L(x) = \frac{1-\gamma_5}{2}\psi(x) \text{ et } \psi_R(x) = \frac{1+\gamma_5}{2}\psi(x)$$

L'interaction faible chargée est alors décrite dans le modèle standard par une structure purement V-A (*via* le projecteur $1-\gamma_5$), les autres contributions (scalaire, pseudo-scalaire et tensorielle) n'étant pas prises en considération. Nous avons vu dans la section 1.1.1 que les fermions élémentaires sont regroupés en doublet gauche d'isospin faible suite à la symétrie observée expérimentalement entre deux membres d'un même doublet. On postule alors l'invariance du lagrangien de l'interaction faible dans la distinction entre un membre d'isospin faible $+\frac{1}{2}$ et un membre d'isospin faible $-\frac{1}{2}$. Cette invariance est une invariance de jauge $SU(2)$ locale (rotation d'un fermion de type *up* à un fermion de type *down*), à trois générateurs $\overrightarrow{T}$. De manière similaire à ce que nous avons montré pour $U(1)$, il faut introduire une dérivée covariante pour garantir l'invariance du lagrangien sous $SU(2)$, ce qui va introduire trois nouveaux bosons de jauge $\overrightarrow{W}_\mu$ de masse nulle et un couplage entre ces bosons de jauge et les fermions.

Au vu des grandes similitudes entre le formalisme de l'interaction faible et l'interaction électromagnétique, ces deux interactions sont formellement unifiées en postulant une seule invariance de jauge pour les deux types de forces. Le choix minimal, appelé le modèle standard des interactions électrofaibles, consiste à postuler une invariance du lagrangien sous une transformation locale du groupe $SU(2)_L \times U(1)_Y$. Avec la dérivée covariante, quatre champs de jauge de masse nulle $\overrightarrow{W}_\mu$ et B_μ sont introduits :

$$D_\mu = \partial_\mu + ig\overrightarrow{T}\cdot\overrightarrow{W}_\mu + ig'\frac{1}{2}B_\mu$$

où $\overrightarrow{T}$ correspond aux trois générateurs de $SU(2)_L$, g et g' sont respectivement les constantes de couplage des interactions faible et électromagnétique.

Si l'on essaie d'introduire directement un terme de masse de type $m^2 B_\mu B^\mu$ pour trois des quatre bosons de jauge, de manière à les faire correspondre aux bosons W^+, W^- et Z^0 de l'interaction faible, on brise l'invariance de jauge $SU(2)_L \times U(1)_Y$.

1.1.3 Le mécanisme de Higgs

Le terme de masse d'un fermion $\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$ dans le Lagrangien est de la forme $-m\bar{\psi}\psi = -m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$. De manière similaire au problème que nous venons de soulever pour les bosons de jauge de l'interaction faible, un tel terme n'est pas invariant sous la symétrie $SU(2)_L \times U(1)_Y$. Pour que le lagrangien décrive ce qui est observé expérimentalement, c'est-à-dire que l'interaction faible ne voit que les fermions de chiralité gauche et les anti-fermions droits, il faut donc éliminer de tels termes de masse. Or les bosons vecteurs de l'interaction faible et les fermions sont massifs, c'est une propriété avérée expérimentalement, il est donc nécessaire de trouver un mécanisme pour introduire la masse des particules élémentaires.

En 1964 P.W. Higgs [12] d'une part et F. Englert, R. Brout [13] d'autre part, proposent d'introduire un champ supplémentaire $\phi(x)$ dans le lagrangien décrivant l'interaction électrofaible. Ce champ est en auto-interaction *via* un potentiel $V(\phi)$ tel que représenté sur la figure 1.1 :

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

Ce potentiel est symétrique et avec $\lambda > 0$ et $\mu^2 < 0$ il possède un ensemble de solutions dégénérées non nulles correspondant à des états fondamentaux équivalents à une phase près :

$$\phi^+\phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}$$

où $v = \sqrt{\frac{\mu^2}{\lambda}}$ est appelée la *vev*, acronyme de *vacuum expectation value*. Le choix d'un minimum particulier brise la symétrie, alors que le potentiel la conserve. Un tel mécanisme est appelée une brisure spontanée de la symétrie.

Le champ $\phi(x)$ le plus simple que l'on puisse introduire pour le mécanisme de Higgs est scalaire, formant un doublet de $SU(2)_L$ de champs complexes :

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{bmatrix}$$

Le lagrangien correspondant à ce champ est :

$$\mathcal{L}_{Higgs} = [D_\mu\phi]^+[D^\mu\phi] - V(\phi)$$

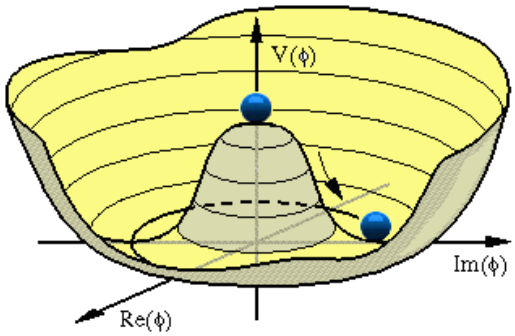

FIG. 1.1: Forme du potentiel de Higgs $V(\phi) = \mu^2\phi^+\phi + \lambda(\phi^+\phi)^2$ avec $\lambda > 0$ et $\mu^2 < 0$.

On choisit comme état fondamental $\phi_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$. On choisit d'autre part que le membre inférieur du doublet ϕ ait une charge nulle ce qui permet d'éviter un couplage avec le photon, avec pour conséquence ultérieure une masse nulle pour ce boson. Le membre supérieur est en conséquence chargé (pour avoir la même hypercharge Y que le membre inférieur du doublet) mais c'est sans importance puisque sa valeur moyenne dans le vide est nulle.

Pour engendrer les particules on effectue des petites oscillation autour du fondamental, paramétrisées par :

$$\phi(x) = \exp\left(-i\overrightarrow{\sigma}\cdot\overrightarrow{\xi}(x)\right)\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

Les trois champs réels $\overrightarrow{\xi}(x)$ n'ont pas de signification physique, ils peuvent être éliminés par un changement de jauge approprié. On les appelle des bosons de Goldstone. C'est par ce choix de jauge que les trois bosons W^+, W^- et Z^0 acquièrent une masse. Il ne reste qu'un seul degré de liberté, le champ $H(x)$ correspondant au boson scalaire et neutre de Higgs.

Les termes de couplage entre le boson de Higgs et les bosons de l'interaction électrofaible dans le lagrangien apparaissent *via* la dérivée covariante :

$$\mathcal{L}^{cin}_{Higgs} = [D_\mu\phi]^+[D^\mu\phi] = \begin{pmatrix} 0 & \frac{v+H(x)}{\sqrt{2}} \end{pmatrix}\left|\partial_\mu + ig\overrightarrow{T}\cdot\overrightarrow{W}_\mu + ig'\frac{1}{2}B_\mu\right|^2\begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix}$$

Les termes quadratiques $\frac{1}{8}v^2g^2(W_1^{\mu 2} + W_2^{\mu 2}) + \frac{1}{8}v^2(g'^2B^\mu - gW_3^\mu)^2$ sont analogues à des termes de masse mais comportent un terme non diagonal en $W_3^\mu B_\mu$. Contrairement à W_1 et W_2, les champs W_3 et B ne sont pas états propres de masse. On peut réécrire la partie concernée sous la forme matricielle suivante :

$$\mathcal{L}^{W_3,B} = \frac{1}{2}\left[W_{3\mu}B_\mu\right]\begin{bmatrix} g^2\frac{v^2}{4} & -gg'\frac{v^2}{4} \\ -gg'\frac{v^2}{4} & g'^2\frac{v^2}{4} \end{bmatrix}\begin{bmatrix} W_3^\mu \\ B^\mu \end{bmatrix}$$

Les états propres de cette matrice de masse sont les bosons vecteurs A_μ de l'interaction électromagnétique et Z_μ de l'interaction faible neutre. Les valeurs propres de cette matrice sont les masses de ces deux bosons de jauge. Finalement les masses des bosons de jauge et du boson de Higgs sont données par :

$$M_W = \frac{1}{2}gv$$

$$M_Z = \frac{v\sqrt{g^2+g'^2}}{2}$$

$$M_A = 0$$

$$M_H = \sqrt{-2\mu^2}$$

La moyenne dans le vide du doublet de Higgs est connue expérimentalement : $v \simeq 246$ GeV, en utilisant la valeur de la constante de Fermi G_F, que l'on peut relier à la constante de couplage g en se placant dans la limite où l'énergie mise en jeu est faible par rapport à M_W (interaction de contact) :

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{1}{2v^2}$$

Par contre le paramètre μ est inconnu et par conséquence la masse du boson de Higgs n'est pas prédite par le Modèle Standard.

Enfin pour introduire un terme de masse pour les fermions fondamentaux, on s'inspire du mécanisme de Higgs en rajoutant au lagrangien des termes de couplage, appelés couplages de Yukawa, entre le doublet scalaire de Higgs et les fermions :

$$\mathcal{L}_{Yukawa} = -g_Y(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi^+\psi_L)$$

où ψ_L est un doublet gauche de fermions, ψ_R est un singulet droit, et g_Y sont les 9 couplages de Yukawa des 3 leptons chargés et des 6 quarks. Ces couplages de Yukawa sont reliés à la masse m_f des fermions : $g_Y = m_f\sqrt{2}/v$. Notons que le couplage des fermions au boson de Higgs est proportionnel à la masse des fermions (alors que nous avons vu que le couplage du Higgs aux bosons de jauge est proportionnel à la masse au carré du boson). Par conséquence le quark top a un couplage fort au Higgs, proche de l'unité.

Dans le cas des quarks, la situation est un peu plus compliquée que dans le cas des leptons. Les quarks de type *up* et ceux de type *down* sont massifs, alors que les membres

de type *up* des doublets de leptons, les neutrinos, sont décrits comme des particules sans masse dans le modèle standard. Les états propres de masse des quarks (se couplant au boson de Higgs) ne coïncident pas avec les états propres de $SU(2)_L$ (se couplant aux bosons W^+, W^- et Z). On introduit alors une matrice de changement de base, appelée la matrice de Cabibbo-Kobayashi-Maskawa (CKM), qui est une matrice unitaire 3x3 décrite par 4 paramètres (3 angles réels et une phase imaginaire).

1.2 Au-delà du Modèle Standard

1.2.1 Succès et insuffisances du Modèle Standard

Les observables du modèle standard sont calculées par un développement perturbatif en puissance de la constante de couplage lorsque sa valeur est inférieure à 1. L'ordre du développement perturbatif d'un diagramme de Feynman correspond au nombre de vertex. Les contributions sont inversement proportionnelles à l'ordre. Il y a l'ordre 0, appelé l'ordre de l'arbre, puis les ordres supérieurs faisant intervenir des boucles de particules virtuelles émises et réabsorbées. La quadri-impulsion des particules dans ces boucles n'est pas contrainte et il apparaît des intégrales divergentes lors du calcul des diagrammes aux ordres supérieurs à l'arbre. Ces divergences sont éliminées par un procédé mathématique appelé la renormalisation. La constante de couplage est remplacée par une constante de couplage effectif qui dépend de l'échelle d'énergie considérée. L'évolution des constantes de couplage avec l'énergie des interactions décrites dans le cadre du Modèle Standard est montrée sur la figure 1.2.

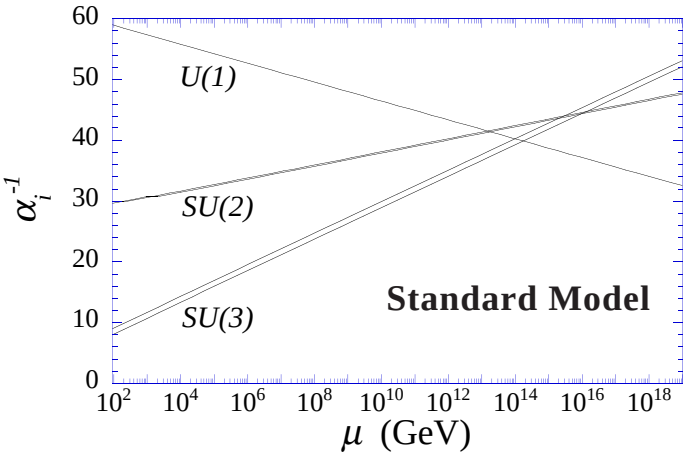

FIG. 1.2: Evolution de l'inverse des constantes de couplage avec l'énergie pour les interactions forte, faible et électromagnétique décrites dans le cadre du Modèle Standard.

Toutes les données expérimentales montrent un fabuleux accord avec les prédictions calculées dans le cadre du modèle standard et la cohérence interne de ce dernier est robuste. Cette robustesse est illustrée par la figure 1.3 qui montre l'écart (rapporté à son erreur) entre la mesure et la prédiction faite dans le cadre du modèle standard, pour les observables électrofaibles.

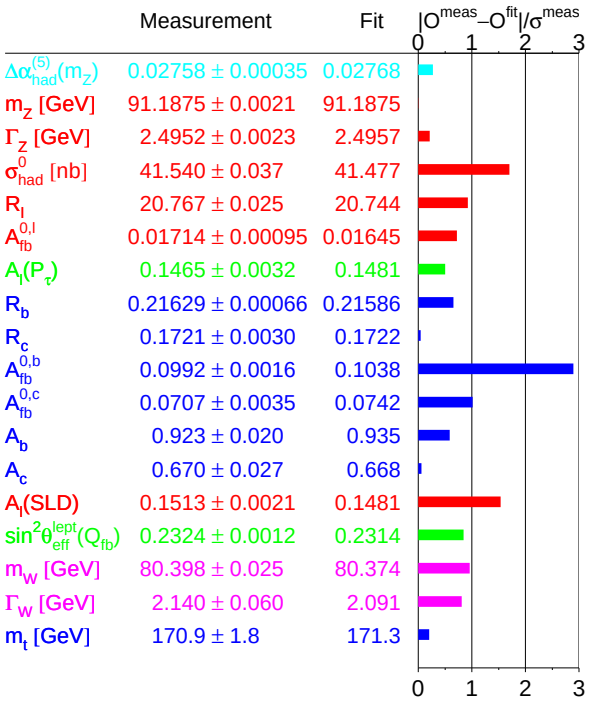

FIG. 1.3: Ecart (rapporté à son erreur) entre la mesure et la prédiction faite dans le cadre du Modèle Standard, pour les observables électrofaibles [14].

Cette théorie souffre néanmoins d'un certain nombre d'insuffisances conceptuelles et d'aspects peu satisfaisants, dont en voici certains :

- il faut inclure la gravitation (décrite par la relativité générale) dont les effets ne sont plus négligeables à l'échelle de Planck (10^{19} GeV) ;
- les constantes de couplage des interactions forte, faible et électromagnétique convergent au même ordre de grandeur à haute énergie, mais l'unification n'est pas exacte ;
- les corrections radiatives aux masses des particules scalaires élémentaires (comme le boson de Higgs) divergent quadratiquement. Pour éliminer ces divergences il est nécessaire d'ajuster avec une très grande précision la valeur des paramètres. Cette précision doit être de plus croissante avec l'ordre des diagrammes de Feynman considérés. Ce type d'ajustement peu satisfaisant est désigné comme le problème de naturalité (ou encore de hiérarchie) ;

- il y a 19 paramètres libres dans le modèle standard : les 3 constantes de couplage des interactions faible, forte et électromagnétique, les 9 masses des fermions massifs, les 2 masses des bosons W et Z, les quatre paramètres de la matrice de Cabibbo-Kobayashi-Maskawa et la masse du boson de Higgs ;
- la structure en 3 générations des particules élémentaires de matière n'est pas expliquée. Ce chiffre de 3 est déduit du nombre de familles de neutrinos légers, qui est une mesure expérimentale obtenue à partir de la largeur du Z ainsi qu'à partir de l'abondance des éléments synthétisés avant l'appartion des étoiles dans l'Univers. L'unitarité de la matrice de Cabibbo-Kobayashi-Maskawa qui décrit les transitions entre les quarks de type *up* et les quarks de type *down* par interaction faible chargée est aussi en accord avec 3 familles de quarks. Cependant rien dans la théorie ne permet de prédire ce chiffre ;
- la différence de masse des particules d'une famille à l'autre n'est pas expliquée, notamment avec une amplitude aussi importante : par exemple entre le quark u de masse de l'ordre du MeV/c^2 et le quark top de masse 172.6 GeV/c^2. Pourquoi le boson de Higgs engendre-t-il des masses différentes pour les 3 familles alors qu'aucun nombre quantique ne semble les distinguer ?

A ces points nous pouvons ajouter des observations expérimentales qui sortent du cadre du Modèle Standard :

- l'observation de l'oscillation de saveur des neutrinos en 1998 [11] montre que ceux-ci sont massifs, alors qu'ils sont décrits comme des particules de masse strictement nulle dans le Modèle Standard ;
- en 1998 aussi, les expériences basées sur l'observation de supernovae de type Ia ont mis en évidence une composante majoritaire d'énergie noire dans l'énergie de l'Univers, contribuant actuellement à hauteur de 70 % [15]. Pour l'instant les physiciens ne connaissent pas la nature de cette énergie noire ;
- la présence de matière noire exotique dans l'Univers à toutes les échelles, non lumineuse contrairement aux étoiles, sensibles à l'interaction gravitationnelle et non baryonique, est aussi bien avérée maintenant.

Finalement les particules appartenant aux trois familles de fermions connus et présentés précédemment, qui sont celles décrites par le Modèle Standard, ne représentent que moins de 4 % de ce qui constitue notre Univers. Le Modèle Standard est une théorie effective, à l'instar de la mécanique classique, valable dans un domaine d'énergie correspondant au minimum à l'énergie du LEP et du Tevatron, c'est-à-dire l'échelle de la centaine de GeV. Pour une énergie plus élevée qui n'est pas encore déterminée, cette théorie n'est plus suffisante et doit être remplacée par une théorie plus complète. Plusieurs idées existent pour étendre (voire remplacer) le Modèle Standard. Ces théories au-delà du Modèle Standard dépendent d'un grand nombre de paramètres inconnus et

pour l'instant les mesures expérimentales permettent de contraindre ces paramètres et ces nouvelles théories mais n'arrivent généralement pas à les exclure complètement.

1.2.2 Les théories au-delà du Modèle Standard et le quark top

La théorie au-delà du Modèle Standard la plus en vogue est la supersymétrie. Son développement remonte aux années 1970 (travaux de J. Wess et B. Zumino [16]). Son principe de base est l'élaboration d'une symétrie entre les fermions et les bosons. Pour chaque boson de spin 1 du Modèle Standard, il existe un partenaire supersymétrique fermionique (appelé jaugino), de même pour chaque fermion du Modèle Standard, il existe un partenaire supersymétrique bosonique de spin 0 (les sfermions). Cette supersymétrie permet de régler *ad hoc* le problème de naturalité exposé dans la section précédente 1.2.1. En effet il apparaît une nouvelle contribution bosonique pour chaque contribution fermionique et réciproquement, et ces contributions étant égales mais de signes opposés, elles se compensent. En réalité, comme il n'a jamais été observé de partenaires supersymétriques pour les particules connues, il faut que les masses des partenaires susy diffèrent de celles des particules standards. Autrement dit il faut que la supersymétrie ne soit pas une symétrie exacte. Mentionnons enfin que la supersymétrie permet aussi la convergence des constantes de couplage des interactions faible, forte et électromagnétique, contrairement au Modèle Standard et que cette théorie propose, sous certaines conditions, un candidat pour la matière noire non baryonique.

Il existe d'autres propositions de physique au-delà du Modèle Standard actuel, comme les théories de technicouleur, de dimensions supplémentaires, de grande unification englobant les symétries du Modèle Standard dans une symétrie plus large ($SO(10)$, $E(6)$, ...). Ces autres théories permettent elles aussi de pallier certains défauts du Modèle Standard. Elles prédisent également des nouveaux processus et des nouvelles particules, dont aucun n'a été mis en évidence pour l'instant. Ces théories ne sont pas forcément exclusives, comme par exemple les dimensions supplémentaires et la supersymétrie.

La mise en fonction imminente du collisionneur le plus puissant jamais construit, le collisionneur pp LHC au CERN, laisse espérer une observation prochaine de nouveaux processus physiques. De tels processus pourraient être mis en évidence dans le secteur du quark top, soit par la mesure précise de ses propriétés soit par la recherche directe d'un nouveau mode de production ou de désintégration du top, faisant intervenir des particules inconnues. Ces processus au-delà du Modèle Standard sont déjà activement recherchés en physique du quark top au Tevatron, seule machine permettant de produire jusqu'à présent ce quark très lourd. Cette physique sera poursuivie ensuite au LHC avec une section efficace de production de paires de quarks top 100 fois supérieure et une luminosité un ordre de grandeur plus élevée qu'au Tevatron, permettant la production et l'étude d'un très grand nombre de quarks top.

Cependant, une fois un nouveau processus mis en évidence, nous pouvons nous attendre à ce qu'il soit délicat de conclure quelle théorie décrit ce nouveau signal et quelles

théories sont éliminées par cette observation. Dans ce contexte aussi la mesure précise des propriétés du quark top peut aider pour l'interprétation.

Nous pouvons citer plusieurs analyses de physique du top contribuant à tester des théories au-delà du Modèle Standard, dont une grande partie sont déjà effectuées au Tevatron :

- les mesures des sections efficaces de production du quark top par interaction forte ($\sigma_{t\bar{t}}$) et par interaction faible (σ_{st}). Un désaccord entre prédiction et mesure pour $\sigma_{t\bar{t}}$ pourrait être le signe d'une désintégration du top en autre chose que $t \to Wb$, par exemple $t \to H^+b$ où H^+ est un boson de Higgs supersymétrique, ou bien un mode de production exotique du top. Concernant le mode de production faible, la sensibilité à la nouvelle physique est différente pour la voie s d'une part dont le mode alternatif de production peut par exemple être un nouveau boson lourd W', un boson de Higgs supersymétrique chargé H^+ ou un mode excité de Kaluza-Klein, et d'autre part la voie t dont le mode alternatif de production peut par exemple être *via* un changement de saveur par un boson neutre (Z, γ ou g) ;
- la mesure de l'élément V_{tb} de la matrice CKM, soit au travers de la mesure de la production de paires $t\bar{t}$ en fonction du nombre de jets de quarks b identifiés, soit par la mesure de la section efficace de production électrofaible d'un quark top unique. Une déviation par rapport à la production peut être due par exemple à une quatrième génération de quarks, une désintégration exotique du quark top ou un bruit de fond exotique ;
- l'étude du couplage tWb par la mesure de l'hélicité du boson W, qui correspond au travail entrepris durant cette thèse. Un désaccord entre la valeur prédite dans le cadre du modèle standard des fractions des trois hélicités possibles pour le W (-1, +1 et 0) et la mesure expérimentale sera observé si la structure de l'interaction faible chargée n'est pas de type V-A. La contribution de type V+A au vertex tWb a déjà été limitée à quelques pourcent dès la première mesure de $b \to s\gamma$ par l'expérience CLEO [17]. Il s'agit cependant d'une contrainte indirecte dépendant notamment de contributions de diagrammes Pingouins sur lesquels l'incertitude théorique est grande. Le quark top permet de tester directement cette propriété fondamentale de l'interaction faible à une échelle d'énergie supérieure à toutes les mesures faites jusqu'à présent. Les contributions anormales au vertex tWb peuvent être testées en mesurant soit le taux de production d'un quark top unique via l'interaction faible, soit l'hélicité du boson W issu de la désintégration d'un top produit par paire via l'interaction forte. Les informations issues de toutes ces mesures sont complémentaires et sont pertinentes pour contraindre des couplages anormaux effectifs différents ;
- la recherche de production résonante de paires de quarks top, lesquelles seraient issues de la désintégration d'un nouveau boson lourd $Z' \to t\bar{t}$ dont l'existence est prédite par de nombreuses théories au-delà du modèle standard ;

- la recherche de désintégrations rares dans le cadre du modèle standard comme $t \to Zc$ et $t \to Z\gamma$, dont un taux trop élevé trahirait un changement de saveur de quark par courant neutre (FCNC) ;
- enfin il faut s'assurer que le quark top est bien le compagnon d'isospin du quark b, en mesurant sa charge électrique et son spin.

1.3 Le quark top

1.3.1 Caractéristiques du quark top

Le quark top est le dernier quark découvert, par les expériences CDF et DØ au Tevatron en 1995 [18]. Sa masse est singulièrement élevée pour une particule élémentaire, puisqu'elle équivaut à celle d'un noyau de tungstène. Cette découverte est survenue après une recherche de longue durée, en effet son existence était soupconnée dès la mise en évidence du quark b en 1977 pour compléter le doublet de $SU(2)_L$. Sa masse a pu être mesurée indirectement dès 1990 [19] en comparant les valeurs des observables électrofaibles (essentiellement la proportion R_b de $Z \to b\bar{b}$) mesurées auprès des accélérateurs e^+e^- LEP et SLC et les valeurs prédites de ces observables. En effet les boucles fermioniques qu'il faut prendre en compte dans le calcul comportent des contributions dominantes proportionnelles au carré de la masse du top.

Jusqu'à présent le quark top n'a pu être étudié qu'au Tevatron, seul collisionneur assez puissant pour le produire sur sa couche de masse. La masse du quark top est mesurée très précisément par les expériences CDF et DØ auprès du Tevatron, la valeur mondiale est $m_{top} = 172.6 \pm 0.8(\text{stat}) \pm 1.1(\text{syst})$ GeV/c^2 [20]. Les mesures à l'origine de cette moyenne mettent en oeuvre des techniques élaborées, utilisant tous les états finals possibles de la paire de quarks top. Les plus précises comparent notamment les données événement par événement aux éléments de matrice calculés à l'ordre dominant (LO) en fonction des observables (les quadri-impulsions des objets). Pour s'affranchir de la source dominante d'incertitude systématique qui est la dépendance de la mesure en fonction de l'échelle d'énergie des jets (JES), un ajustement simultané de la masse du quark top et de JES est effectué. La précision de 1.4 GeV/c^2 atteinte sur la masse du top est particulièrement impressionnante et résulte d'un important effort dans cette optique des deux collaborations CDF et DØ. Le collisionneur LHC qui entrera en fonction bientôt prévoit d'atteindre une précision de 1 GeV/c^2 avec une luminosité de 10 fb^{-1} (1 année après le démarrage du LHC). La mesure de la masse du quark top est déjà limitée par les incertitudes systématiques, d'où la difficulté pour améliorer la précision actuelle auprès d'un collisionneur hadronique.

L'importance accordée à la précision sur la masse du top est essentiellement due à son rôle dans la contrainte induite sur la masse du boson de Higgs *via* les corrections quantiques (la masse du top, comme toutes les masses de fermions, intervient au carré

dans les corrections radiatives à l'ordre de la boucle). La figure 1.4 montre la contrainte apportée sur la masse du Higgs par les mesures précises de la masse du boson W et du quark top. Actuellement le facteur limitant la contrainte sur la masse du Higgs est la précision sur la mesure de la masse du W. Cette dernière doit atteindre 0.07 fois la précision sur la masse du quark top pour avoir une contribution similaire à la contrainte sur la masse du Higgs. La valeur actuelle est $M_W = 80.398 \pm 0.025$ GeV/c^2.

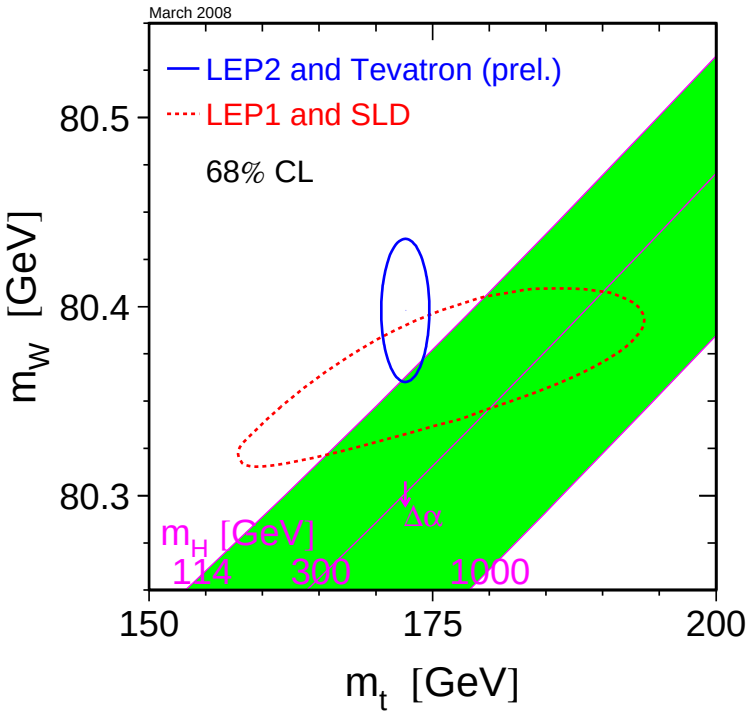

FIG. 1.4: Contraites sur la masse du boson de Higgs issues des mesures directes des masses du quark top et du boson W (ellipse bleue) et des mesures indirectes dans le secteur élecrofaible (contour rouge).

Les autres grandeurs à mesurer pour caractériser le quark top sont sa largeur, son spin et sa charge. Le Tevatron ne permet pas de mesurer sa largeur. Son temps de vie prédit dans le cadre du Modèle Standard est de l'ordre de 0.5×10^{-24} ps. C'est un ordre de grandeur plus petit que le temps nécessaire à l'hadronisation d'un quark. Le quark top est ainsi le seul quark qui se désintègre avant de se trouver lié par interaction forte dans un hadron. Cette particularité importante permet d'étudier les propriétés du quark top à partir de ses produits de désintégration. Par exemple le spin d'un quark lié dans un hadron est modifié *via* des interactions avec le spin de l'autre quark de ce hadron,

et ces modifications se font avant la désintégration par interaction faible du quark. Les produits de désintégration d'une particule belle ne permettent ainsi pas de mesurer le spin de cette particule, contrairement au cas du quark top. La production de spin du quark top peut être mesurée par exemple à partir des distributions angulaires de ses produits de désintégration. Le spin peut aussi être étudié au travers de l'interaction forte lors de la production d'une paire $t\bar{t}$. Le top produit est non polarisé, c'est-à-dire que les hélicités positives et négatives ont la même occurence, mais les spins du t et du $\bar{t}$ sont corrélés. Au Tevatron, contrairement au LHC, il est possible de définir un axe de quantification selon lequel la corrélation entre les spins du t et du $\bar{t}$ est de 100 %, car 85 % des paires $t\bar{t}$ sont produites à partir d'un gluon de spin 1. Cependant une telle étude expérimentale nécessite une grande statistique et la mesure précise des directions des quarks top et n'est pas envisageable au Tevatron, de même pour la mesure du spin.

Quant à sa charge, la possibilité que la particule découverte et étudiée soit une particule de charge $-\frac{4}{3}e$ plutôt que $+\frac{2}{3}e$ est exclue par les données de DØ avec un niveau de confiance de 92 % [21]. Cette possibilité vient du mode de reconstruction des quarks top. Il s'agit d'une paire de deux quarks top, se désintégrant respectivement en W^+b et $W^-\bar{b}$. La charge du quark b n'étant pas mesurée, il s'ensuit une ambiguïté sur l'association correcte des deux W et des deux b et donc de la charge des deux quarks top reconstruits.

1.3.2 Production du quark top au Tevatron

1.3.2.1 Production forte du quark top

Le quark top est produit majoritairement par paire au Tevatron (deux à trois fois plus souvent que seul). Il s'agit d'une production par interaction forte dont les diagrammes dominants sont présentés sur la figure 1.5 : $q\bar{q} \rightarrow t\bar{t}$ pour 85 % des paires et $g\bar{g} \rightarrow t\bar{t}$ pour 15 % des paires (estimation NLO). C'est ce canal de production qui a permis la découverte du quark top au Tevatron en 1995 [18].

La section efficace de production d'une paire $t\bar{t}$ a été mesurée avec une précision de 20 % au *Run I* du Tevatron avec $\sqrt{s} = 1.8$ TeV ($\sigma_{t\bar{t}} = 5.9 \pm 1.7$ pb), en accord dans la limite des incertitudes avec la prédiction théorique ($\sigma_{t\bar{t}} = 6.7^{+0.71}_{-0.88}$ pb [22]). Pour l'énergie du *Run II* de $\sqrt{s} = 1.96$ TeV, la section efficace prédite (au NNLO et pour $m_{top} = 175$ GeV/c^2) est $\sigma_{t\bar{t}} = 7.91 \pm 0.71$ pb [23] et la mesure individuelle la plus précise a été obtenue par l'expérience DØ avec l'état final lepton+jets : $\sigma_{t\bar{t}} = 8.18^{+0.90}_{-0.84}$ pb [24].

1.3.2.2 Production faible du quark top

Le quark top est aussi produit au Tevatron *via* l'interaction faible, qui produit un top unique (*single top*) dans l'état final. Les processus menant à cet état final, dont les diagrammes de Feynman dominants sont montrés sur la figure 1.6, sont les suivants :

- le processus $p\bar{p} \rightarrow t\bar{b} + X$, $\bar{t}b + X$ de production dans la voie s d'un *single top*, avec une section efficace de 0.88 ± 0.14 pb [25] (au NLO et pour $m_{top} = 175$ GeV/c^2) ;

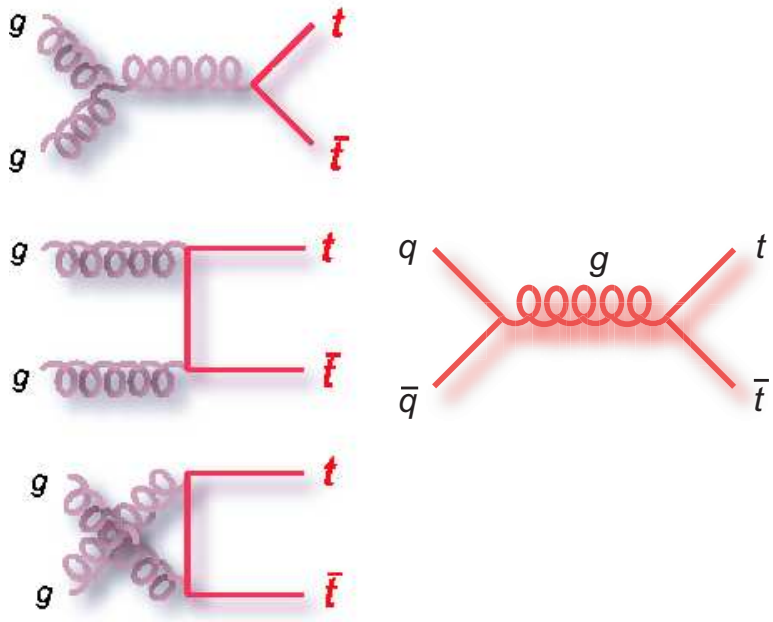

FIG. 1.5: Les quatre diagrammes de Feynman à l'ordre de l'arbre correspondant à la production de paires de quarks top par interaction forte au Tevatron dans le cadre du Modèle Standard.

- le processus $p\bar{p} \to tq\bar{b}+X$, $\bar{t}\bar{q}b+X$ de production dans la voie t d'un *single top*, avec une section efficace de 1.98 ± 0.30 pb [25] (au NLO et pour $m_{top} = 175$ GeV/c^2) ;
- enfin le processus $p\bar{p} \to tW^- + X$, $\bar{t}W^+ + X$, dont la section efficace est très faible au Tevatron : 0.08 ± 0.02 pb [26] (au LO et pour $m_{top} = 175$ GeV/c^2).

La section efficace de production d'un *single top* par interaction faible au Tevatron est donc de l'ordre de 3 pb au total, c'est-à-dire seulement 2.5 fois moins que la section efficace de production de paires de quarks top. Néanmoins ce processus a été très difficile à mettre en évidence au Tevatron. En plus de la section efficace plus faible, le bruit de fond est plus difficile à éliminer dans l'état final avec un seul top parce que la multiplicité en jets est plus faible. Ainsi une analyse basée sur des coupures séquentielles ne permet pas d'observer un excès d'événements par rapport à la quantité de bruit de fond attendu. Le signal de *single top* est estimé en combinant des outils statistiques multi-variables sophistiqués (réseaux de neurones bayesiens, arbres de décision *boostés* et fonctions de vraisemblance discriminantes utilisant les éléments de matrice des processus). Ce mode de production est ainsi mis en évidence par DØ en décembre 2006 [27]. Les sections efficaces mesurées pour les voies $s+t$ sont de 2.2 ± 0.7 pb par CDF [28], correspondant à une significance de 3.7 écarts standards et 4.7 ± 1.3 pb par DØ [29], correspondant à une significance de 3.6 écarts standards.

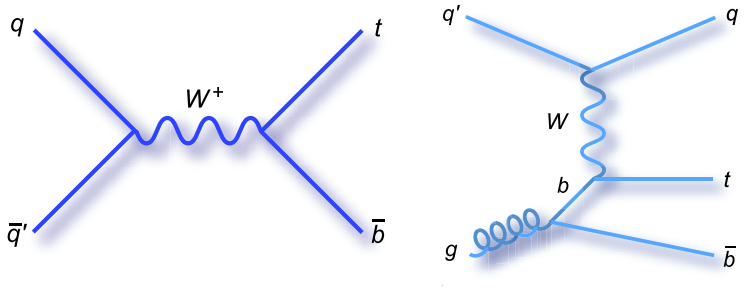

FIG. 1.6: Les diagrammes de Feynman des processus dominants de production d'un top unique par interaction faible dans le cadre du Modèle Standard : production dans la voie s (à gauche) et dans la voie t (à droite)

La mesure de la section efficace de production de *single top* permet de contraindre les théories de physique au-delà du Modèle Standard, comme cela est illustré en exemple sur la figure 1.7, notamment si le boson W est en fait un nouveau boson lourd W' dans la voie s ou un boson neutre (Z^0, γ) permettant un changement de saveur de quark dans la voie t (voir la figure 1.6).

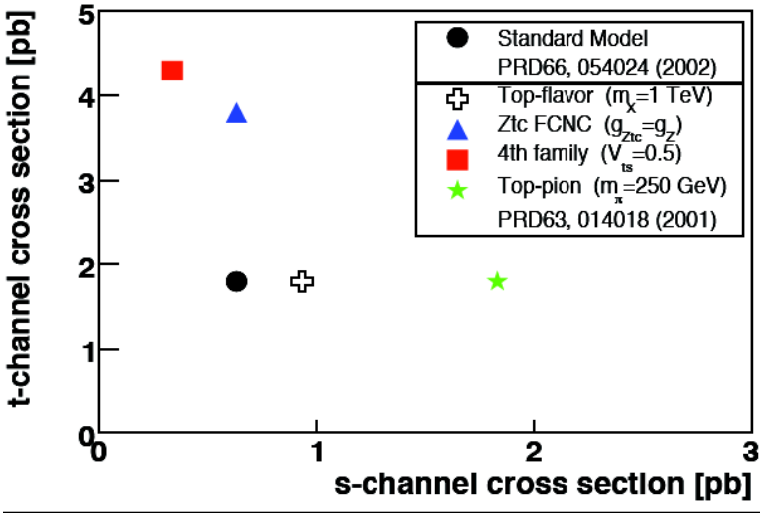

FIG. 1.7: Exemples de contraintes sur les théories au-delà du Modèle Standard obtenues à partir des mesures des sections efficaces de production dans la voie s et dans la voie t de *single top* par interaction faible.

La section efficace de production de *single top* étant directement proportionnelle au module au carré de l'élément $|V_{tb}|$ de la matrice CKM, sa mesure permet de mesurer directement cet élément de matrice. Cette section efficace peut différer de la prédiction

dans le cadre du modèle standard parce que $|V_{tb}| < 1$ mais aussi parce que le couplage tWb n'est pas un couplage standard de structure $V-A$. Des analyses sont en cours dans DØ pour contraindre les contributions anormales (de type tensorielles par exemple) au vertex tWb.

1.3.3 Désintégration du quark top et d'une paire de quarks top

Le quark top, s'il est effectivement standard, se désintègre en tant que quark libre, en un quark b et un boson W^+ réel dans quasiment 100 % des cas (voir figure 1.8).

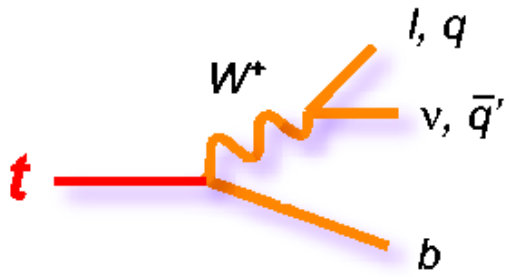

FIG. 1.8: Diagramme de Feynman décrivant la désintégration du top dans le cadre du Modèle Standard.

Comme nous l'avons vu dans la section 1.3.2, le quark top est étudié au Tevatron *via* le mode de production par interaction forte, produisant une paire $t\bar{t}$. Les états finals d'une paire de top se distinguent par le mode de désintégration de chacun des deux bosons W, comme le résume la figure 1.9 :

- **le canal dilepton :** les deux bosons W se désintègrent chacun en un lepton chargé et son neutrino associé. Ce canal comporte deux leptons isolés de grande impulsion transverse et de charges opposées, beaucoup d'énergie manquante emportée par les neutrinos et deux jets issus de l'hadronisation de quarks beaux. La sélection d'un tel état final se fait avec une bonne pureté grâce aux deux leptons. Peu d'analyses utilisent les états finals comportant des leptons τ et le canal dilepton fait généralement référence aux états finals e+jets et mu+jets seulement. Dans ce cas, cet état final de la paire de quarks top a un rapport d'embranchement très faible, de l'ordre de 5 % seulement et le poids des analyses dilepton dans les mesures combinant plusieurs canaux est faible. Pour des mesures nécessitant la reconstruction des quadri-impulsions des quarks top, comme la mesure de la masse du top, cet état final présente l'inconvénient d'être sous-contraint par la présence d'au moins deux neutrinos.
  A l'inverse, cet état final est très intéressant par sa simplicité, par exemple au démarrage d'une nouvelle machine comme le LHC. Le quark top peut jouer le rôle

de chandelle standard pour comprendre et calibrer un détecteur, et le canal dilepton sera le tout premier à être étudié dans ce but (par exemple pour la mesure de la section efficace de production de paires de quarks top).

- **le canal lepton+jets :** un boson W se désintègre leptoniquement et l'autre hadroniquement. Cet état final comporte un lepton isolé de haute impulsion transverse, de l'énergie manquante emportée par le neutrino et au moins 4 jets dont deux sont issus de l'hadronisation de quarks beaux. Si l'on ne prend pas en compte les canaux faisant intervenir des τ, dont la signature est plus complexe, cet état final intervient dans 30 % des désintégrations de la paire de quarks top.
- **le canal hadronique :** les deux bosons W se désintègrent en quarks, ce qui mène à un état final comportant au minimum 6 jets. Le rapport d'embranchement en ce canal est important, de l'ordre de 45 % (en ne prenant pas en compte le canal τ+jets), mais évidemment la sélection sur un tel état final multi-jets est difficile dans un environnement hadronique. L'identification des jets de quarks beaux est indispensable dans ce canal, mais malgré tout le bruit de fond reste important.

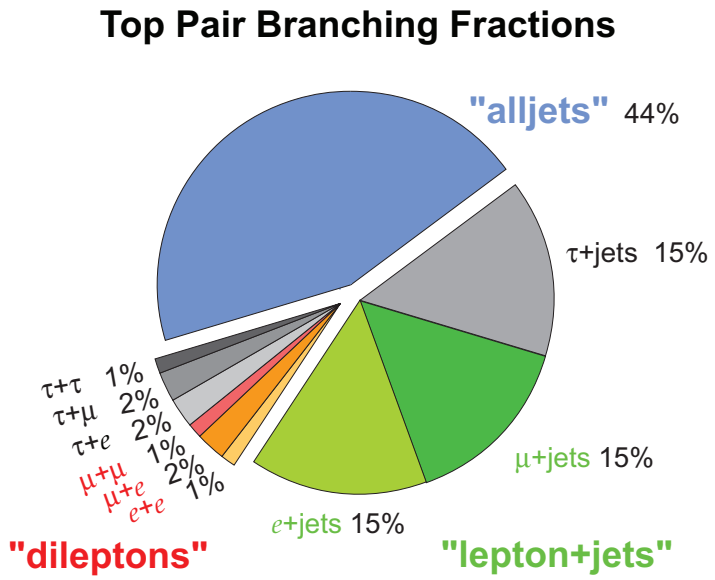

FIG. 1.9: Canaux de désintégration d'une paire de quark $t\bar{t}$ et leur proportion relative.

2

Le Tevatron et l'expérience DØ

Le Tevatron est un collisionneur proton-antiproton installé au *Fermi National Accelerator Laboratory* (F.N.A.L., ou Fermilab) [30] dans la banlieue de Chicago. Le Fermilab est un laboratoire en service depuis 1967, tout d'abord sous la simple dénomination *National Accelerator Laboratory*. Il fut renommé en 1974 en l'honneur d'Enrico Fermi. Il fut le théâtre de la découverte des deux derniers quarks, le quark b en 1977 et le quark top en 1995. Depuis 1987, le complexe d'accélération à atteint sa forme actuelle avec la mise en service du collisionneur proton-antiproton Tevatron.

Outre le Tevatron et ses deux détecteurs CDF[1] [31] et DØ [32] localisés le long de son périmètre aux points de collision, le Fermilab accueille un certain nombre d'expériences dans divers domaines, et plus particulièrement les expériences MiniBooNE et MINOS dédiées à l'étude des neutrinos.

Le Tevatron a connu une première phase de prise de données entre 1992[2] et 1996, appelée *Run I* qui permit de collecter 125 pb^{-1} de données par expérience. L'accélérateur délivrait alors des collisions avec une énergie dans le centre de masse[3] de $\sqrt{s} = 1.80$ TeV. Plusieurs années d'arrêt de la prise de données physiques ont été mises à profit pour modifier en profondeur les détecteurs DØ et CDF, et améliorer la chaîne d'accélération. Le résultat a été en 2001 la reprise des collisions avec cette fois-ci une énergie dans le

[1] *Collider Detector at Fermilab*

[2] En 1988 seul CDF était en fonctionnement. Il fallut attendre 1992 pour que les deux détecteurs commencent à enregistrer les collisions.

[3] s est la variable de Mandelstam : $s = (p_{proton} + p_{antiproton})^2$, p représentant la quadri-impulsion des particules incidentes.

centre de masse $\sqrt{s} = 1.96$ TeV et une luminosité augmentée.

Une première partie de ce chapitre s'attèlera à décrire succintement chacune des étapes nécessaires au fonctionnement du collisionneur Tevatron. Viendra ensuite une description plus complète du détecteur DØ installé auprès d'un des deux points de collision des faisceaux. Enfin, une dernière partie décrira les principes sous-jacents à la reconstruction de tous les objets à partir des signaux enregistrés par le détecteur DØ. Dans cette dernière partie nous préciserons les critères de sélection appliqués pour définir un jet, un électron, un muon et le vertex primaire tels qu'ils sont utilisés dans l'analyse présentée dans cette thèse.

2.1 Le complexe de production et d'accélération du Tevatron

Les paragraphes qui suivent présentent toutes les étapes permettant au complexe Tevatron de faire entrer en collisions les faisceaux de protons et d'antiprotons, de leur production à leur collision en passant par leur accélération. La figure 2.1 présente une vue aérienne du Fermilab et du Tevatron, ainsi qu'une schématisation détaillée des divers constituants de la chaîne de production et d'accélération des protons et des antiprotons qui entreront en collision aux points nommés DØ et CDF.

2.1.1 Production du faisceau de protons

2.1.1.1 Le Pré-accélérateur

Un pré-accélérateur de Cockroft-Walton ionise un gaz d'hydrogène au moyen d'un champ électrostatique. Les électrons sont tout d'abord arrachés des atomes d'hydrogènes, les protons libres sont ensuite attirés par une cathode de Cesium où ils récupèrent deux électrons pour donner des ions H^-. Ces derniers sont alors expulsés de la cavité et accélérés jusqu'à une énergie de 750 keV. Une fois le faisceau d'ions H^- accéléré, il traverse une ligne de transfert appelée la *"750 keV line"* avant d'entrer dans le *Linac*.

2.1.1.2 Le Linac

Le *Linac* est un accélérateur linéaire de 130 m de long constitué de cavités accélératrices radio-fréquence, au pouvoir d'extraction de 3 MeV.m^{-1}. Il produit un faisceau (un *pulse*) structuré en paquets espacés de 5 ns et contenant $6.3 \cdot 10^{12}$ ions H^- de 400 MeV, qui traversent ensuite une autre ligne de transfert appelée la *"400 MeV line"* pour rejoindre le *Booster*.

Fermilab Tevatron Accelerator With Main Injector

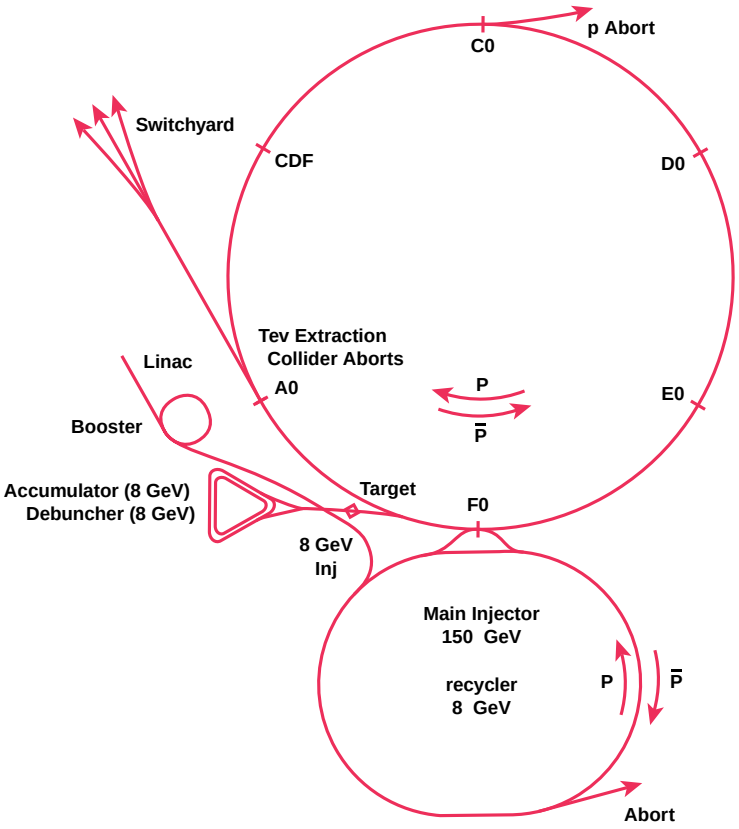

FIG. 2.1: En haut : Vue aérienne du complexe d'accélération et du collisonneur Tevatron. En bas : Schéma de la chaîne de production et d'accélération.

2.1.1.3 Le Booster

Les ions H^- passent tout d'abord au travers d'une feuille de carbone, ce qui a pour effet d'arracher leurs électrons pour donner les protons. Un champ magnétique élimine les ions H^- et les atomes d'hydrogène restants. Le *Booster* est le premier accélérateur circulaire (synchrotron) de la chaîne d'accélérateurs. Il mesure 475 m de circonférence. Il est constitué d'un ensemble d'aimants qui courbent et focalisent le faisceau, ainsi que de 18 cavités radio-frequence qui permettent d'accelerer les protons d'environ 0.5 MeV par tour pour fournir au bout de 25 ms des protons de 8 GeV. Ceux-ci sortent alors du *Booster* par paquets d'environ $5 \cdot 10^{10}$ protons, pour être envoyés dans l'Injecteur Principal.

2.1.2 L'Injecteur Principal

L'Injecteur Principal est le changement majeur de la chaîne d'accélération lors du passage du *Run I* au *Run II*. C'est un synchrotron d'environ 1 km de diamètre, composé de cavités radio-fréquence et d'un guide magnétique formé d'aimants dipolaires et quadripolaires. Il remplit plusieurs fonctions :

- accélérer des protons jusqu'à 150 GeV pour les injecter dans le Tevatron après accumulation en paquets de $2.7 \cdot 10^{11}$ protons ;
- accélérer des protons jusqu'à 120 GeV pour la production des antiprotons après accumulation en paquets de $2 \cdot 10^{12}$ protons, ainsi que pour les expériences sur cibles fixes ;
- réaccélèrer les antiprotons issus du cycle de création et de refroidissement (cf. ci-dessous) jusqu'à 150 GeV pour les injecter dans le Tevatron par paquets de $1.3 \cdot 10^{12}$;
- décélérer des antiprotons jusqu'a 8 GeV à la fin du cycle de collision pour les transférer dans le recycleur.

2.1.3 Production du faisceau d'antiprotons

Cette étape de production et de stockage des antiprotons est celle qui limite actuellement l'augmentation de la luminosité. Former un faisceau d'antiprotons de haute intensité est en effet plus complexe et difficile que pour les protons.

2.1.3.1 La Station-Cible

Une cible de Nickel, large de 10 cm et épaisse de 2 cm, est bombardée par le faisceau de protons à 120 GeV. De la sorte, de très nombreuses particules secondaires sont produites, dont en particulier des antiprotons d'énergie moyenne de 8 GeV. Ceux-ci ont néanmoins une distribution en énergie très large. Toutes les particules produites sont focalisées dans une même direction par une lentille magnétique au lithium (sa faible densité limitant son interaction avec les antiprotons), tandis qu'un dipôle de 1.5 T, utilisé comme spectromètre de masse, sélectionne les antiprotons produits. Le rendement

est d'environ 1 à 2 antiprotons pour 10^5 protons. Les antiprotons ainsi produits gardent la structure en paquet des protons incidents.

2.1.3.2 Le Debuncher

Ces paquets d'antiprotons sont ensuite envoyés dans le *Debuncher*, dont le rôle est de réduire leur dispersion en énergie. Cet accélérateur triangulaire aux sommets arrondis, de 505 m de long, est composé d'aimants dipôlaires et quadrupôlaires ainsi que d'une cavité radio-fréquence dans une des sections droites, qui se charge du refroidissement stochastique du faisceau. Ce refroidissement reduit la taille transverse du faisceau et stabilise son énergie, en corrigeant à chaque tour la dispersion en énergie au moyen d'un champ électromagnétique. Cependant, la réduction de la dispersion en énergie se fait au détriment des dispersions spatiales et temporelles. Ainsi le faisceau d'antiprotons, initialement pulsé, est maintenant continu. Il est envoyé dans l'Accumulateur, qui est situé dans le même tunnel que le *Debuncher*.

2.1.3.3 L'Accumulateur

L'Accumulateur est composé lui aussi de cavités radio-fréquence et d'un système de refroidissement stochastique. Son rôle est de conserver au faisceau sa faible dispersion énergétique, tout en augmentant sa densité et en reconstituant les paquets d'antiprotons afin de correspondre au faisceau de protons. Plusieurs heures sont nécessaires (environ 8h) pour atteindre la quantité d'antiprotons voulue, de $1.3 \cdot 10^{12}$ antiprotons.
Une fois le faisceau constitué, les antiprotons sont envoyés dans l'Injecteur Principal où ils seront accélérés jusqu'à 150 GeV, avant d'être eux aussi injectés dans le Tevatron, dans la direction opposée à celle des protons.

2.1.3.4 Le Recycleur

Considérant que seule une faible partie des particules circulant dans l'anneau participe effectivement aux collisions, un système permettant de recupérer les antiprotons restants pour les réutiliser a été mis au point. Ce Recycleur est un synchrotron à aimants permanents, situé dans le même tunnel que l'Injecteur principal. Cependant, il s'est finalement avéré que ce système n'était pas réalisable, et depuis début 2004 le Recycleur est utilisé comme simple anneau de stockage des antiprotons de 8 GeV issus de l'Accumulateur, en complément de l'Accumulateur (qui a une capacité de stockage limitée à environ $3 \cdot 10^{12}$ antiprotons). Il permet de stocker un total de $2 \cdot 10^{12}$ antiprotons. De plus, depuis mi-2005, un faisceau d'électrons de 4.8 MeV est également utilisé pour améliorer le refroidissement et ainsi augmenter l'intensité du faisceau d'antiprotons dans le Recycleur.

2.1.4 Le Tevatron

Le Tevatron est le dernier et le plus grand élément de la chaîne d'accélération. C'est un collisionneur d'un kilomètre de diamètre. Il s'agit du premier synchrotron à avoir utilisé la supraconductivité. L'anneau est divisé en six secteurs nommés de A à F.

Le Tevatron est constitué de :
- 772 dipôles supraconducteurs, qui produisent un champ magnétique de 4.33 T afin de courber les faisceaux. Ils nécessitent d'être refroidis par un système cryogénique fonctionnant à la température de l'hélium liquide (3.6 K).
- 180 quadrupôles, qui contrôlent la taille transverse des faisceaux et assurent leur guidage. Ils assurent la focalisation des faisceaux aux points de collision une fois l'énergie de 980 GeV atteinte.
- 8 cavités accélératrices radio-fréquence, qui accélérent les protons et les antiprotons de 150 GeV à 980 GeV.

Les faisceaux sont constitués en réalité de 3 super-paquets (ou *superbunch*) espacés de 2.6 μs, eux-mêmes constitués de 12 sous-paquets distants de 396 ns. Ils entrent en collision en deux points de l'anneau, B0 et D0, où sont situés les détecteurs CDF et DØ. La situation stable où les deux faisceaux à 980 GeV collisionnent est appelée un *store*. La durée de vie des faisceaux est cependant limitée à une quinzaine d'heures par les interactions avec le gaz résiduel dans le tube à vide. Une nouvelle phase d'injection de protons et d'antiprotons est alors nécessaire, elle dure environ huit heures.

Comme indiqué précédemment, la figure 2.1 présente le complexe de production, d'accélération et de collision des faisceaux de protons et d'antiprotons. Le tableau 2.1 quant à lui récapitule les fonctions des divers éléments de la chaîne et les gammes d'énergies correspondantes atteintes par les protons et les antiprotons. Enfin, le tableau 2.2 compare les caractéristiques principales du Tevatron des trois phases de prise de données. Une description technique très détaillée de chaque élément du complexe est disponible dans les *Operations Rookie Books* accessibles depuis [33].

	Nom de l'élément	Fonction	Gamme d'énergie
Production des protons	Source de protons	production des H-	$\leq$ 10 KeV
	Cockroft-Walton	accélération des H-	750 KeV
	Linac	accélération des H-	0.750 - 400 MeV
	Booster	accélération des p	0.400 - 8 GeV
	Injecteur principal	accélération des p	8 - 150 GeV
	Tevatron	accélération, collisions	150 - 980 GeV
Production des antiprotons	Station-Cible	production, collection	8 GeV
	Debuncheur, accumulateur	stockage	8 GeV
	Injecteur principal	accélération	8 - 150 GeV
	Tevatron	accélération, collisions	150 - 980 GeV

TAB. 2.1: Liste des composants intervenant dans la chaîne de production et d'accélération des protons et des antiprotons à Fermilab.

	Run I	*Run IIa*	*Run IIb*
Période d'activité	1992/1996	2001/2006	2006/2009 ?
Energie $\sqrt{s}$ (TeV)	1.8	1.96	1.96
Luminosité instantanée ($\mathrm{cm^{-2}s^{-1}}$)	1.6×10^{30}	1×10^{32}	3×10^{32}
Luminosité intégrée délivrée ($\mathrm{fb^{-1}}$)	0.15	1.6	8
Nombre de paquets $p \times \bar{p}$	6×6	36×36	36×36
Nombre d'interactions	2.5	2.3	8
Temps de croisement (ns)	3500	396	396

TAB. 2.2: Principales caractéristiques du Tevatron aux *Run I*, *Run IIa* et *Run IIb*.

2.2 Le détecteur DØ

Le détecteur DØ est un détecteur généraliste de physique des particules auprès d'un collisionneur. Il a été essentiellement conçu dans le but de découvrir et d'étudier le quark top.
La conception du détecteur DØ a débuté en 1983 et son installation a pris fin en 1992, date à partir de laquelle le *Run I* [34] et la prise de données physiques ont commencé. Il est principalement constitué de trois systèmes de détection concentriques (cf. figure 2.2), le détecteur de traces chargées et de vertex au plus proche du faisceau, puis le calorimètre et enfin le spectromètre à muons. De très nombreux changements ont été entrepris entre le *Run I* et le *Run II*, que ce soit l'ajout ou la modification de tout ou partie de certains éléments (trajectographe, détecteurs de pieds de gerbe,...), ou afin de tenir compte des nouveaux temps de croisement des faisceaux (remplacement de l'électronique de lecture, révision du système de déclenchement). Le *Run II* a débuté en mars 2001, et le détecteur a fini d'être installé au printemps 2002 [35]. S'ensuivirent au cours des années plusieurs périodes de coupure du Tevatron qui furent mises à profit pour procéder à de nouveaux changements, en particulier au cours du printemps 2006 l'ajout de la *Layer 0* du détecteur de vertex.
La sélection et la reconstruction d'un événement contenant des quarks top nécessite de mesurer l'ensemble des objets qui le constituent. Les paragraphes qui suivent s'attèlent à la description de l'ensemble du système de détection, après une définition du système de coordonnées utilisé. Une partie de mon travail de thèse a consisté en l'amélioration de la reconstruction des muons dans le spectromètre à muons. La section correspondante détaillera le travail réalisé.

2.2.1 Le système de coordonnées

Le système des coordonnées cartésiennes est défini comme suit. L'axe z correspond à l'axe du faisceau, orienté positivement selon la direction des protons. L'axe y correspond à la direction orthoradiale à l'anneau du Tevatron, orientée vers le haut, et l'axe x à la direction radiale dirigée vers l'intérieur de l'anneau. Le plan (xOy) résultant définit le plan transverse. Cependant, il est plus pertinent d'utiliser les coordonnées sphériques avec (cf. figure 2.3) :

- l'angle polaire θ par rapport à l'axe z, variant entre 0 et π.
- l'angle azimutal ϕ : $\phi = \pi/2$ correspond à l'axe (0y) et $\phi = 0$ correspond à l'axe (0x). Cet angle caractérise la projection de la trajectoire dans le plan transverse (x, y) : $\phi = \arctan y/x$.
- R : distance transverse par rapport à l'axe z : $R = \sqrt{x^2 + y^2}$.

qui permettent de décrire la physique, et parfois les coordonnées cylindriqyes $Rz\phi$.

En pratique, la pseudo-rapidité η est utilisée en lieu et place de l'angle polaire θ. La pseudo-rapidité correspond à une approximation relativiste ($(mc^2/E) \rightarrow 0$) de la

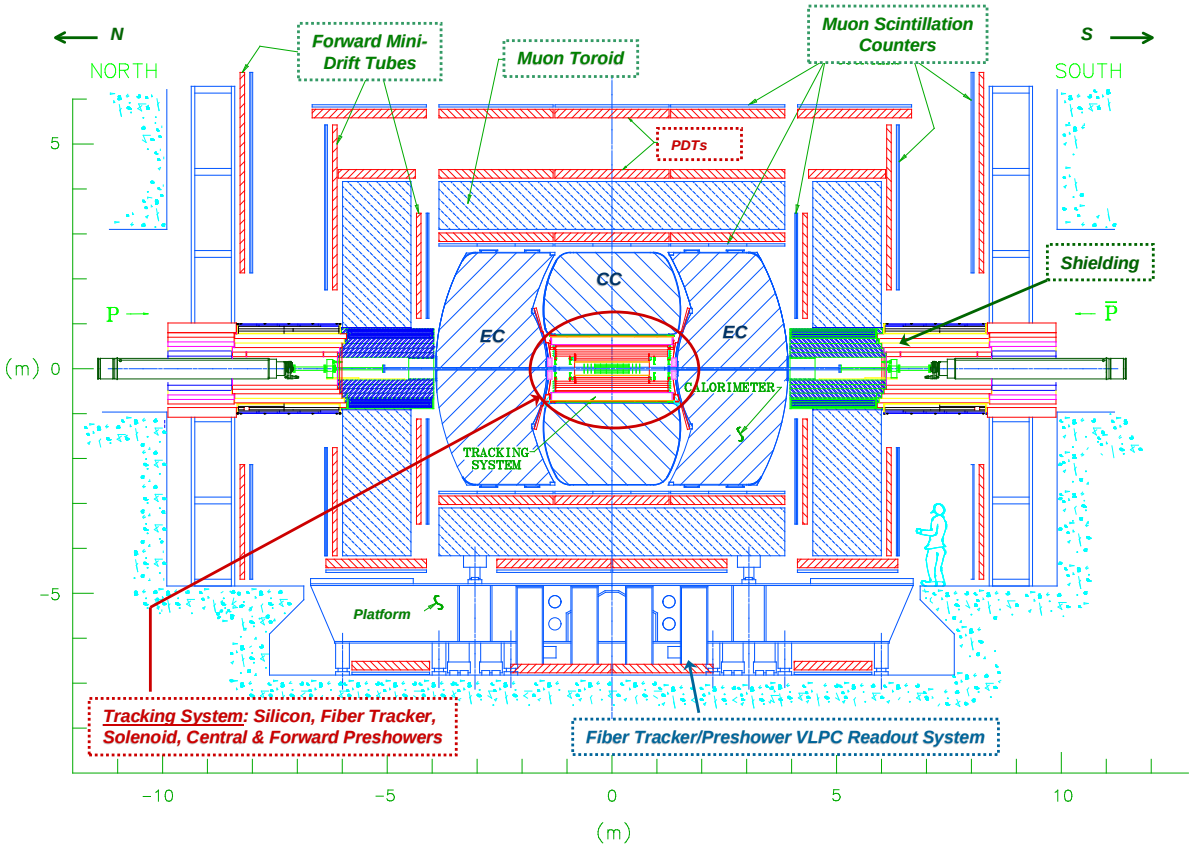

FIG. 2.2: Représentation schématique de l'ensemble du détecteur DØ (coupe longitudinale).

rapidité y pour des angles finis. Cette dernière est invariante par transformation de Lorentz selon l'axe z avec la définition suivante :

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \;\rightarrow\; \eta = -\ln\tan\frac{\theta}{2}\;(E \gg m) \tag{2.1}$$

où E désigne l'énergie, p_z la composante longitudinale de l'impulsion et m la masse. On définit ensuite la distance angulaire entre deux objets comme étant $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

Deux types de coordonnées sont définis. Les coordonnées géométriques (indicées $_{det}$), dont l'origine du repère correspondant se confond avec le centre géométrique du détecteur (tous les objets sont initialement reconstruits par rapport à ce repère), et les coordonnées physiques, pour lesquelles l'origine correspond cette fois-ci à la position du vertex primaire. En effet, de par la taille longitudinale des faisceaux, ces deux positions ne sont généralement pas confondues. La distribution selon z des vertex primaires d'interaction est une gaussienne centrée en 0 mais d'une largeur d'environ 25 cm. En l'absence de contre-indication, les coordonnées indiquées seront les coordonnées phy-

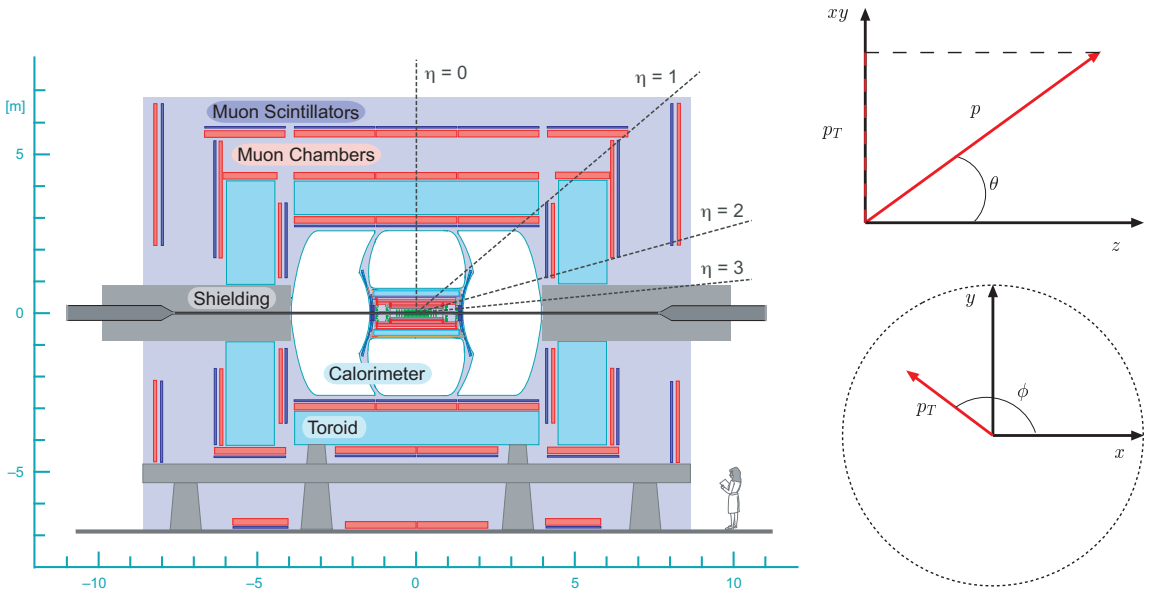

FIG. 2.3: Figure de gauche : Coupe longitudinale du détecteur DØ. Quelques intervales de pseudo-rapidité y sont représentés. Figure de droite : schéma du système de coordonnées.

siques (en particulier pour la pseudo-rapidité, cf. figure 2.4).

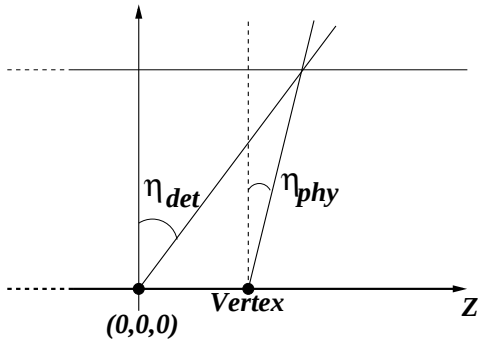

FIG. 2.4: Pseudo-rapidité *physique* et *géométrique*.

Lors d'une collision proton-antiproton aux énergies du Tevatron ce sont leurs sous-constituants, possédant une fraction inconnue de l'énergie totale, qui interagissent. Comme une partie des particules s'échappent le long du tube du faisceau sans être détectée (jets de particules à bas angles provenant essentiellement des partons dits spectateurs, qui ne participent pas à l'interaction dure), et que l'impulsion longitudinale initiale des partons interagissant n'est pas connue, l'impulsion, l'énergie et l'énergie

manquante utilisées dans les analyses sont celles mesurées dans le plan perpendiculaire à l'axe du faisceau. On peut en effet utiliser la conservation de l'énergie et de l'impulsion dans le plan transverse en considérant que l'impulsion transverse des partons initiaux est nulle. On parle alors d'impulsion transverse P_T, d'énergie transverse E_T et d'énergie transverse manquante $\not\!\!E_T$. L'énergie manquante avant toute correction résulte de la présence dans l'état final de particules interagissant très peu voire pas du tout avec le calorimètre (respectivement des muons et des neutrinos), ou d'effets de reconstruction imparfaite.

2.2.2 Le système de reconstruction des traces et des vertex

Ce trajectographe est constitué de trois sous-composants concentriques, un détecteur de vertex au silicium, un détecteur de traces à fibres scintillantes et un solénoïde. Il permet de répondre à trois objectifs :

- reconstruire les trajectoires tri-dimensionnelles des particules chargées ;
- mesurer leur impulsion ;
- reconstruire les vertex primaires et secondaires.

Les deux sous-détecteurs servent en outre aux décisions de déclenchement (cf. section 2.3).

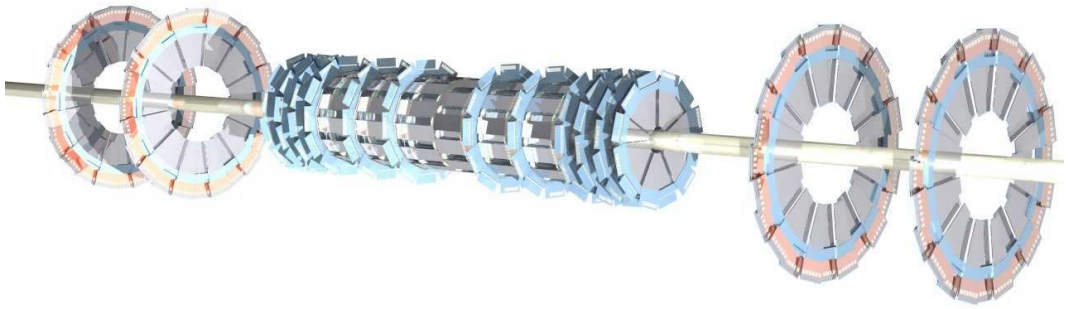

FIG. 2.5: Vue isométrique du détecteur SMT. On distingue au centre les 6 barillets encadrés des 12 disques F et aux extrémités les 4 disques H.

2.2.2.1 Le détecteur de vertex en silicium SMT (*Silicon Microstrip Tracker)*

Le SMT [36] est un détecteur de vertex à micro-pistes de silicium placé au plus près du faisceau qui a été inclus dans la prise de données physiques depuis le début du *Run II*, en mars 2001. Il permet de reconstruire en trois dimensions les traces des particules chargées et les vertex d'interaction sur presque toute la couverture en pseudo-rapidité du calorimètre et du détecteur de muons. Il a une importance cruciale pour les algorithmes

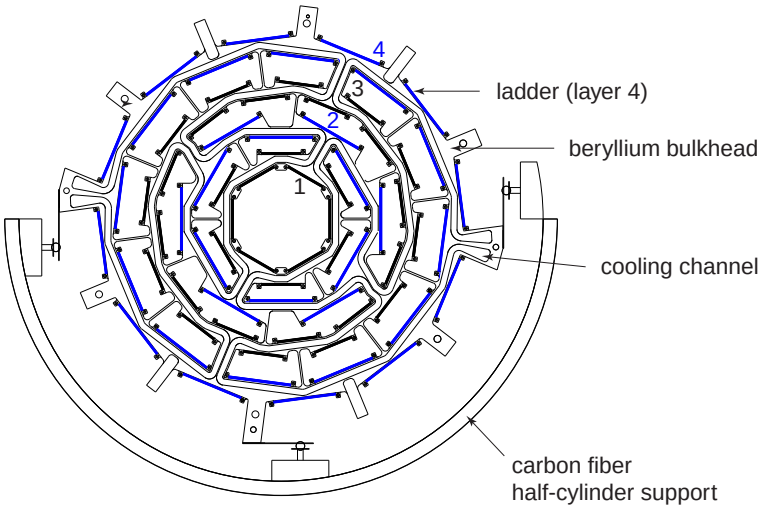

FIG. 2.6: Section transversale d'un barillet du SMT. Les couches 1 et 3 (noir) sont constituées de détecteurs simple-face, les couches 2 et 4 (bleu) de double-face.

d'identification de jets de quarks beaux. La longueur de la zone d'interaction d'environ 25 cm a déterminé sa taille.

Le SMT est constitué de barillets alignés le long du faisceau et de disques orthogonaux au faisceau, pour une longueur totale d'environ 2.4 m (cf. figure 2.5). Les six barillets de 12 cm de long entourent le faisceau et assurent une couverture angulaire de $|\eta_{det}| < 1.7$. Ils sont intégrés dans un support en béryllium. Un barillet est composé de 4 couches concentriques de capteurs à micro-pistes en silicium, de rayons respectifs de 2.7, 4.5, 6.6 et 9.5 cm. Chaque barillet comporte au total 72 détecteurs agencés en quinconce afin d'assurer une bonne herméticité. Une section transversale d'un tel barillet est illustrée par la figure 2.6. La composition des couches des barillets dépend de leur position. Les deuxièmes et quatrièmes couches des six barillets sont toutes constituées d'élements double-face à pistes axiales pour l'une et décalées d'un angle de 2°pour l'autre. Les premières et troisièmes couches diffèrent. Pour les quatre barillets centraux, elles sont composées de capteurs double-face à pistes axiales ou formant un angle stéréo de 90°, pour les barillets externes, de détecteurs simple-face à pistes axiales.
Les barillets sont séparés par des interstices de 8 mm de long dans lesquels viennent s'insérer six des douzes disques F. Les six restants se répartissent aux deux extrémités, après les barillets. Ils sont situés respectivement à $|z| = 12.5$, 25.4, 38.2, 43.1, 48.1 et 53.1 cm. Un disque F est constitué de 12 secteurs de silicium double face dont les pistes forment un angle stéréo de 30°. Les barillets et les disques F sont supportés par deux demi-cylindres en fibre de carbone.
Afin de pouvoir enregistrer des événements à plus haute rapidité, ce système est com-

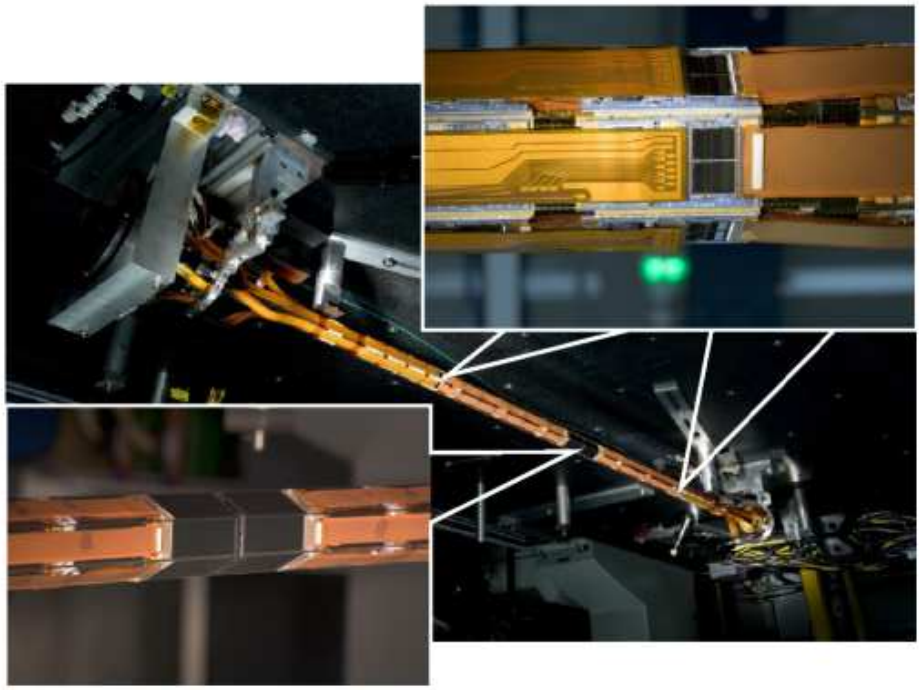

FIG. 2.7: Photographie du détecteur L0.

plété par 4 disques H positionnés aux extrémités, aux positions $|z| = 100.4$ cm et $|z| = 121.0$ cm. Le SMT atteint ainsi une couverture angulaire allant jusqu'à $|\eta_{det}| < 3.0$. Les disques H sont composés de 24 secteurs, chacun constitué de deux capteurs simple-face accolés et formant un angle stéréo de 15°.
Le SMT a ainsi une résolution selon l'axe z d'environ 10 μm et selon le plan $R-\phi$ d'environ 35 μm. Le tableau 2.3 récapitule quelques unes des caractéristiques de ses divers éléments constitutifs.

Un système de refroidissement à base d'eau et de glycol à -10°C circulant dans le support en béryllium et le support des disques maintient une température inférieure à 5 °C dans le SMT, ce qui permet de limiter les dommages dûs aux radiations et d'évacuer la chaleur engendrée par l'électronique de lecture.

Il était prévu que le seuil de radiations tolérées par le SMT soit dépassé suite au vieillissement du détecteur et à l'augmentation de la luminosité. Au-delà de ce seuil de 1 Mrad, qui correspond à une luminosité intégrée d'environ 1 fb^{-1}, il était attendu que la résolution spatiale du SMT se dégrade fortement. Pour pallier à cela, une couche supplémentaire de détecteurs au silicium, interne au SMT (d'où son nom de *Layer 0*) et située

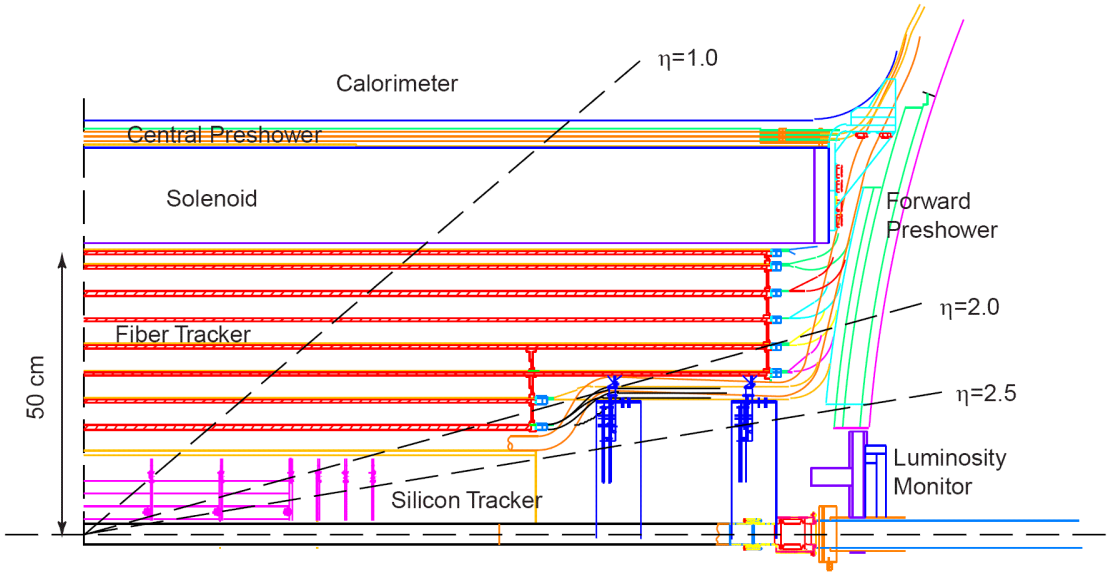

FIG. 2.8: Schéma du quart nord-est de la cavité centrale de DØ. Le SMT, le CFT, le solénoïde et les détecteurs de pied de gerbe y sont représentés.

à 1.7 cm du faisceau, a été ajoutée en 2006 pour le *Run IIb*. La figure 2.7 en présente des photographies. Cette couche très proche du faisceau améliore nettement la résolution en paramètre d'impact des traces.

Détecteur	Layer 0	Barillet	Disque F	Disque H
Couches / Secteurs	1	4	12	2
Canaux	12288	387072	258048	73728
Modules	48	432	144	96
Rayon interne (cm)	1.61	2.7	2.6	9.5
Rayon externe (cm)	1.77	9.4	10.5	26

TAB. 2.3: Quelques caractéristiques des éléments du détecteur SMT.

2.2.2.2 Le détecteur de traces à fibres scintillantes CFT (*Central Fiber Tracker*)

Le CFT est constitué d'un total de 76800 fibres scintillantes de 835 μm de diamètre montées sur 8 supports cylindriques concentriques disposés de 20 à 52 cm du faisceau, entourant ainsi le SMT. Il est représenté sur la figure 2.8. Les deux cylindres les plus in-

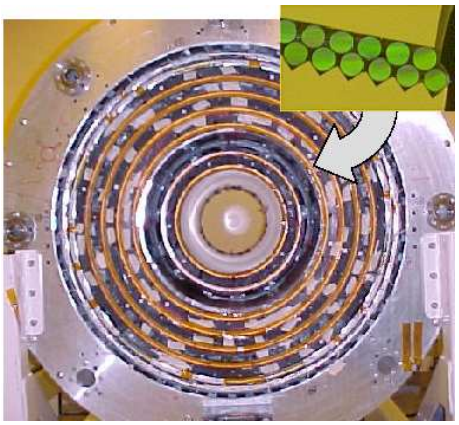

FIG. 2.9: Photographie du CFT et zoom sur une double couche de fibres scintillantes (photographie en haut à droite).

ternes sont longs de 1.66 m, les six autres de 2.52 m. Son acceptance géométrique totale correspond à $|\eta_{det}| < 1.7$ avec une couverture totale. Les six couches externes permettent d'étendre la couverture angulaire jusque $|\eta_{det}| = 2$. Chaque cylindre est composé de deux double-couches de fibres scintillantes, la première axiale, la seconde formant alternativement un angle de $+3^{\circ}$ou -3°. Au sein d'une double-couche, les fibres sont placées en quinconce afin de maximiser la surface de détection (cf. figure 2.9), pour une résolution par couche dans le plan $R - \phi$ de l'ordre de 100 μm. Toutes les fibres scintillantes sont ensuite raccordées à des guides d'onde transparents qui transmettent l'information lumineuse engendrée par le passage d'une particule à des compteurs de photons (VLPC, *Visible Light Photon Counter*). Ces VLPC fonctionnent à très basse température (9 K) et se chargent de produire un signal électrique avec une efficacité quantique d'environ 80 %.
L'information du CFT, combinée à celle du SMT, permet ainsi de reconstruire avec précision les trajectoires des particules chargées. Le CFT permet également d'augmenter le bras de levier.

2.2.2.3 L'aimant solénoïdal

Le solénoïde supraconducteur délivre un champ magnétique uniforme à 0.5 % de 1.9 Tesla, parallèle à l'axe z. Le champ, initialement de 2 T, a été diminué en 2004 d'environ 5% afin d'éviter les pertes de la supraconductivité (*quenches*). C'est un cylindre de 2.54 m de long et 61 cm de rayon qui entoure le CFT. Il est constitué de deux couches supraconductrices d'un alliage de cuivre, de niobium et de titane, stabilisées par de l'aluminium, comptant au total pour 0.9 X_0[4] de matière. Le champ magnétique courbant

[4] X_0 correspond à la distance moyenne de matière traversée pour laquelle l'énergie d'un électron est divisée par e.

la trajectoire des particules chargées, ce solénoïde permet de reconstruire et mesurer les impulsions des traces chargées.

2.2.2.4 Performances du trajectographe

La résolution σ_{P_T} sur l'impulsion transverse P_T des traces chargées de pseudo-rapidité η, obtenue en combinant les informations du SMT et du CFT correspond à :

$$\frac{\sigma_{P_T}}{P_T} = \sqrt{(S \cdot \sqrt{\cosh \eta})^2 + (C \cdot p_T)^2}$$

où S est lié aux interactions multiples et C est le terme de résolution. Les valeurs nominales (resp. mesurées) correspondantes sont $S = 0.015$ (resp. 0.015) et $C = 0.0018$ (resp. 0.0014).

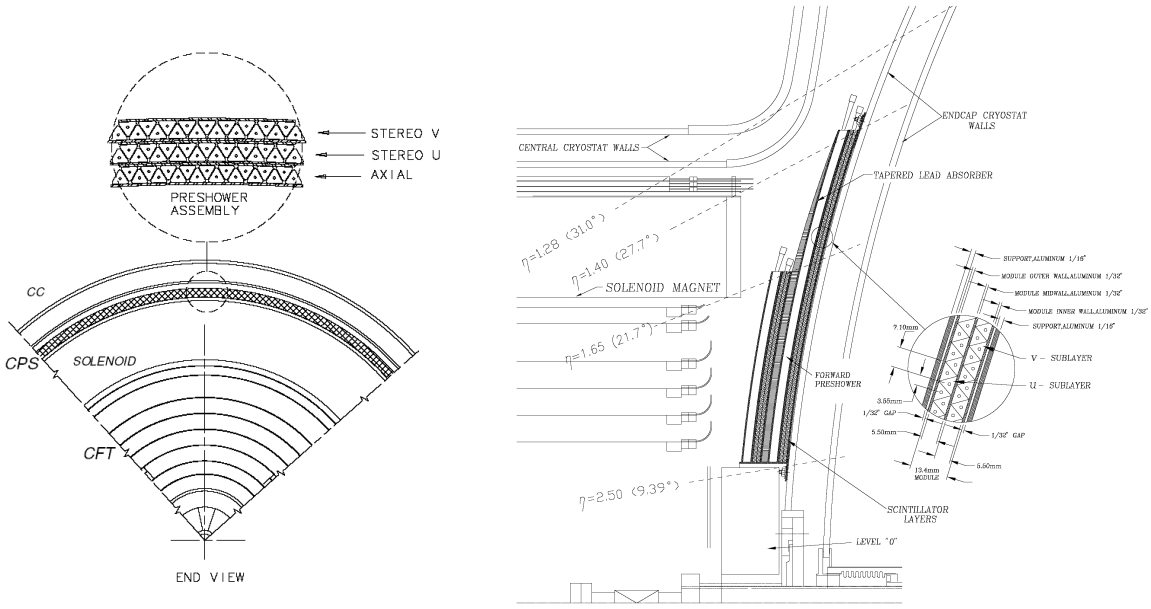

FIG. 2.10: Figure de gauche : vue en coupe radiale du CPS. Figure de droite : vue en coupe longitudinale d'un FPS.

2.2.2.5 Les détecteurs de pieds de gerbes

L'épaisseur de matière rencontrée par les particules lorsqu'elles traversent le solénoïde peut induire le développement d'une gerbe électromagnétique avant qu'elles n'atteignent le calorimètre. Pour pallier à la baisse de résolution induite sur l'énergie mesurée dans le calorimètre et mesurer le début de ces gerbes, deux types de détecteurs de pieds de

gerbes [37] ont été ajoutés. Ils fonctionnent à la fois en tant que trajectographe en mesurant la direction des particules incidentes et en tant que calorimètre en échantillonnant l'énergie perdue dans le solénoïde. Ils permettent ainsi également d'améliorer l'identification des électrons, des photons et des pions neutres.

Les deux détecteurs sont formés de bandes triangulaires de scintillateurs intercalés de façon à limiter les zones mortes, comme illustré sur la figure 2.10. Les scintillateurs ne sont pas arrangés exactement de la même façon pour les deux détecteurs. A l'intérieur de ces scintillateurs triangulaires est disposée une fibre optique à décalage de longueur d'onde de 835 μm de diamètre, reliée en sortie à un guide d'onde transparent de même diamètre qui conduit le signal lumineux à un VLPC.
Le détecteur de pieds de gerbes central, ou CPS (pour *Central Pre-Shower detector*) est situé entre le solénoïde et le calorimètre central, et couvre la région $|\eta_{det}| < 1.3$. Une couche d'absorbeur au Plomb d'environ 1 X_0 sépare de plus le CPS du solénoïde. Ce détecteur est constitué de trois couches cylindriques concentriques de 1280 scintillateurs triangulaires intercalés, la première axiale, les deux suivantes formant un angle stéréo de $\pm$ 23°avec elle.
Les deux détecteurs de pieds de gerbes avant et arrière, ou FPS (pour *Forward Pre-Shower detector*) sont montés sur la face interne des cryostats des bouchons du calorimètre, entre les détecteurs inter-cryostat et les luminomètres. Ils ont tous les deux la même structure, composée de deux double-couches de scintillateurs séparées d'une couche d'absorbeur au Plomb de 2 X_0. Au sein d'une double-couche, les scintillateurs font un angle stéréo de 22.5°. La première double-couche se limite à la région $1.65 < |\eta_{det}| < 2.5$, alors que la seconde s'étend sur $1.5 < |\eta_{det}| < 2.5$. En effet, la zone $1.5 < |\eta_{det}| < 1.65$ est dans l'ombre du solénoïde qui fournit jusqu'à 3 X_0 (la couche de Plomb ne couvre pas cette région).

La résolution spatiale de ces détecteurs est de 1.4 mm pour un électron. Les couches de Plomb permettent de s'assurer que quelle que soit la pseudo-rapidité $|\eta_{det}|$, la même quantité de matière sera traversée avant d'atteindre le calorimètre.

2.2.3 Le calorimètre

Le calorimètre de DØ est inchangé par rapport au *Run I*, seule son électronique de lecture a été remplacée afin de pouvoir faire face au nouveau temps de croisement des faisceaux. Il est constitué de trois calorimètres qui ont été élaborés dans l'objectif de fournir une mesure précise de l'énergie pour les électrons, les photons et les jets, d'aider à l'identification des électrons, des photons, des jets et des muons et enfin d'établir le bilan d'énergie transverse d'un événement. Le calorimètre de DØ se caractérise par une très grande herméticité.

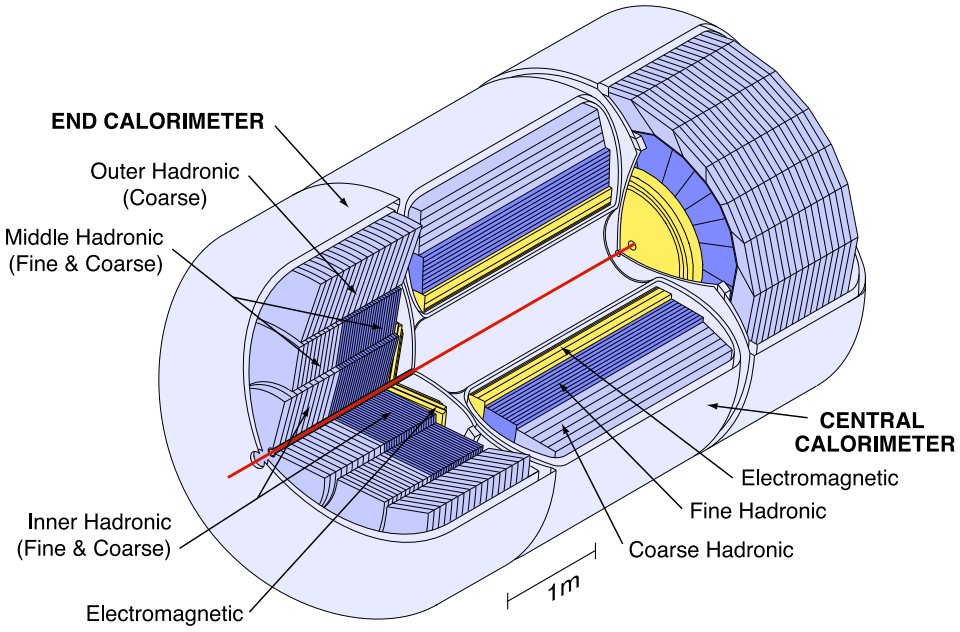

FIG. 2.11: Schéma du calorimètre central et des deux bouchons calorimétriques. Les couches électromagnétiques sont représentées en jaune, les couches hadroniques en bleu (foncé pour les couches FH, clair pour les couches CH).

Le Calorimètre Central (CC) couvre approximativement la région $|\eta_{det}| < 1$, tandis que les Calorimètres Externes Nord (ECN) et Sud (ECS) étendent cette couverture de $|\eta_{det}| \approx 0.7$ jusque $|\eta_{det}| < 4.5$ (ils sont aussi appelés Bouchons). L'interstice entre les Bouchons et le Calorimètre Central est appelé la région intercryostatique et s'étend de $|\eta_{det}| \approx 0.8$ à $|\eta_{det}| \approx 1.4$. Un système de détection additionnel y est inséré (cf. paragraphe 2.2.3.3) afin de combler l'absence d'instrumentation dans cette zone, due en partie au câblage. Les trois calorimètres sont enveloppés dans une double couche d'acier inoxydable afin de maintenir l'argon liquide à une température de 78 K.
Chaque calorimètre comporte une partie électromagnétique (la plus proche de la région d'interaction) notée EM, puis des sections hadroniques, de granularité fine puis grossière, respectivement notées FH (*Fine Hadronic*) et CH (*Coarse Hadronic*). Toutes sont composées d'un certain nombre de couches. La figure 2.11 représente de façon détaillée l'ensemble du calorimètre et de ses diverses sections et les tableaux 2.4 et 2.5) donnent ses caractéristiques.
On associe à ces deux types de calorimètres deux grandeurs physiques :

- pour les objets électromagnétiques : la longueur de radiation X_0, correspondant à la distance moyenne que parcourt une particule dans un matériau avant de perdre $1/e$ de son énergie ;

– pour les hadrons : la longueur d'interaction λ_0, qui correspond au libre parcours moyen entre deux interactions fortes dans le milieu traversé.

Les différentes couches constituant les calorimètres sont des cylindres concentriques pour le CC et des disques orthogonaux au faisceau pour les EC (sauf la derniere couche CH) regroupés en trois cylindres concentriques nommés IH (*Inner Hadronic*), MH (*Middle Hadronic*) et OH (*Outer Hadronic*). La particularité du cylindre OH est d'avoir ses plaques d'absorbeur en acier inclinées d'un angle de 60 °par rapport à l'axe du détecteur (visible sur la figure 2.12) et de ne pas avoir de partie fine. La région de granularité fine permet généralement de mesurer entièrement l'énergie et la position des hadrons incidents, tandis que la région grossière protège les chambres à muons contre des gerbes tardives, et permet un échantillonnage en bout de gerbe et la récupération d'une partie de l'énergie au minimum d'ionisation des muons, seules particules à s'échapper du calorimètre.

Le calorimètre de DØ est composé d'une succession de couches d'absorbeur qui initient les gerbes de particules, et de couches d'élément actif qui s'ionisent au passage des particules de la gerbe. L'élement de base du calorimètre est la cellule. L'enregistrement des charges qui sont créées dans cette dernière permet de mesurer l'énergie de la particule correspondante. Un tel calorimètre, dit à échantillonnage, améliore la résolution spatiale au détriment de la résolution en énergie, comparé à un calorimètre homogène.
Le type d'absorbeur diffère selon les zones, cf. tableaux 2.4 et 2.5. La partie électromagnétique est constituée d'Uranium appauvri, qui a de nombreuses qualités. Outre une faible longueur de radiation, il permet au calorimètre d'être compact, mais également, d'avoir un calorimètre compensé : le rapport entre l'énergie réelle d'une particule et l'énergie mesurée sera le même pour une particule électromagnétique et pour une particule hadronique (c'est un peu moins le cas au *Run II* que ce ne le fut au *Run I*). La compensation est un atout lors de la mesure de l'énergie de jets dont la gerbe comporte une partie électromagnétique et une partie hadronique.
L'élément actif est de l'Argon liquide, dont les avantages sont un taux d'ionisation très élevé, une calibration simple, une segmentation flexible et une bonne résistance à l'exposition continue aux radiations, ce qui assure la stabilité au cours du temps.
Une cellule est constituée d'une plaque d'absorbeur suivi d'une plaque en cuivre à 2.3 mm servant d'électrode collectant les charges, le tout baignant dans l'Argon liquide (cf. figure 2.13). Une différence de potentiel de 2 à 2.5 kV est appliquée entre la plaque d'absorbeur reliée à la masse et la plaque de cuivre. Le champ électrique induit entraîne la dérive en 450 ns environ des charges électriques créées par l'ionisation du milieu actif dû au passage d'une particule. Le temps de croisement des faisceaux étant de 396 ns, l'électronique de lecture devra gérer une superposition d'événements.
Ces cellules basiques sont regroupées en cellules de lecture, correspondant à l'association en profondeur des blocs ayant les mêmes coordonnées η et ϕ. La taille transverse des cellules de mesure correspond à la taille transverse des gerbes : 1 à 2 cm pour les gerbes électromagnétiques et environ 10 cm pour les gerbes hadroniques. L'échelle en η et ϕ correspond à la taille typique des jets de partons, i.e. $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \approx 0.5$.

Les subdivisions longitudinales sont utiles car les profils longitudinaux des gerbes permettent de discriminer entre les électrons et les hadrons. Ces modules de lecture forment ainsi des tours dites pseudo-projectives, dont les centres des cellules les constituant sont alignés avec le centre géométrique du détecteur. Le calorimètre compte un total de 5000 tours pour 55 296 cellules.

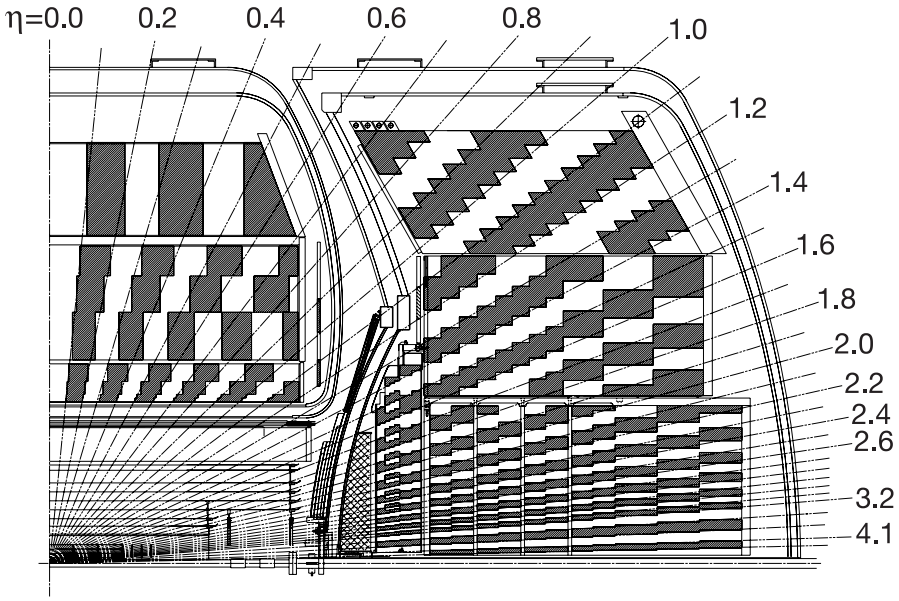

FIG. 2.12: Vue schématique d'un quart du calorimètre présentant les motifs de segmentation longitudinals et transverses.

	EM	FH	CH
Nombre de modules en ϕ	32	16	16
Nombre de couches	4	3	1
Absorbeur (épaisseur en mm)	Ur (3)	Ur-Nb à 2% (6)	Cu (46.5)
X_0	≈ 21	≈ 96	≈ 32
λ_0	0.76	3.2	3.2

TAB. 2.4: Caractéristiques du Calorimètre Central.

2.2.3.1 Granularité

Les couches des modules EM, FH et CH du CC sont constituées de cellules de lecture représentant une granularité de $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, à l'exception de la troisième couche

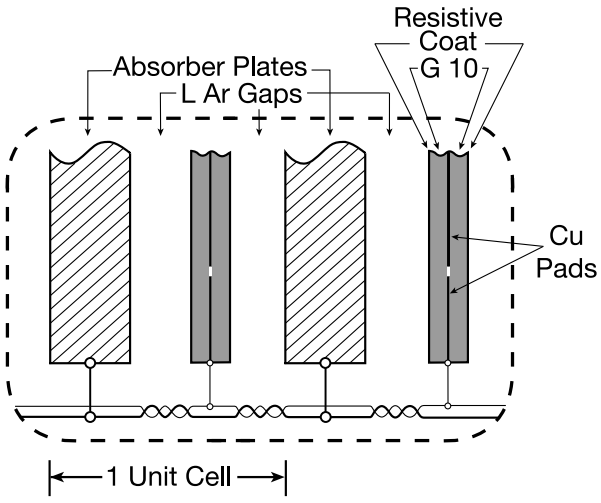

FIG. 2.13: Schéma d'une cellule du calorimètre.

	EM	IH		MH		OH
		FH	CH	FH	CH	CH
Nbre modules en ϕ	1	1	1	16	16	16
Nbre couches	4	4	1	4	1	1
Absorbeur (mm)	Ur (4)	Ur-Nb 2% (6)	Inox (46.5)	Ur-Nb 2% (6)	Inox (46.5)	Inox (46.5)
X_0	≈ 21	≈ 122	≈ 32	≈ 115	≈ 38	≈ 65
λ_0	0.95	4.9	3.6	4.0	4.1	7.0

TAB. 2.5: Caractéristiques des Calorimètres Externes.

qui présente une segmentation plus fine ($\Delta\eta \times \Delta\phi = 0.05 \times 0.05$) car elle couvrait la région de plus grand dépôt d'énergie électromagnétique des gerbes au *Run I* (ce n'est plus le cas au *Run II*, les gerbes étant initiées plus tôt car traversant plus de matière). Concernant les bouchons, la segmentation est la même, sauf à haute pseudo-rapidité ($|\eta| > 3.2$) où la granularité en $\Delta\eta \times \Delta\phi$ passe à 0.2×0.2 (pour des contraintes de construction), ou à 0.1×0.1 à $|\eta| > 2.6$ pour la troisième couche EM.
Entre deux modules se trouve une zone non-instrumentée appelée interstice en ϕ ou *crack*. Afin que les zones non instrumentées ne soient pas projectives, les modules azimutaux ont été décalés en ϕ (cf. figure 2.12).

2.2.3.2 Performances du calorimètre

La résolution $\sigma(E)$ sur l'énergie E des jets reconstruits grâce au calorimètre comporte trois contributions :

$$\frac{\sigma(E)}{E} = \sqrt{\frac{a^2}{E} + \frac{b^2}{E^2} + c^2} \tag{2.2}$$

où

- la contribution $\frac{a}{\sqrt{E}}$ est due à la réponse stochastique, provenant des fluctuations inhérentes au développement de gerbes dans le calorimètre (fragmentation, échantillonnage et fluctuations intrinsèques des jets). C'est le terme dominant de la résolution en énergie des jets sur la plupart du spectre.
- le terme $\frac{b}{E}$ est induit par le bruit électronique, dû aux effets instrumentaux (bruit, radioactivité de l'Uranium, piédestaux, interactions multiples), indépendant de l'energie. Ce facteur limite les performances du calorimètre pour les jets de faible énergie.
- le terme c est constant, dû notamment aux erreurs de calibration, aux non-uniformités et au non-linéarités de la réponse du calorimètre, aux matériaux morts précédant le calorimètre, au champ magnétique. Ce terme impose les limites des performances à très haute énergie.

Les termes a, b et c ont été mesurés sur des faisceaux-tests de pions et d'électrons, les valeurs obtenues sont répertoriées dans le tableau 2.6.

Particule	a	b	c
Electron	0.43 GeV	0.135 $\text{GeV}^{\frac{1}{2}}$	0.0115
Pion	0.975 GeV	0.45 $\text{GeV}^{\frac{1}{2}}$	0.032

TAB. 2.6: Contributions de la résolution du calorimètre, issus de faisceaux-tests.

2.2.3.3 Le Détecteur Inter-Cryostat

Etant donné que le complexe calorimétrique est contenu dans trois cryostats séparés, la couverture en pseudo-rapidité dans la région $0.8 < |\eta| < 1.4$ est incomplète. De plus, cette région comporte beaucoup de matériau non échantillonné, ce qui dégrade la résolution de l'énergie et accroît artificiellement la quantité d'énergie manquante. Afin de pallier à cela, deux dispositifs ont été ajoutés. Tout d'abord, des cellules sans absorbeur constituées uniquement d'argon liquide et appelées *massless gaps* sont disposées entre les cryostats centraux et externes. Ce sont les parois du cryostat qui tiennent lieu d'absorbeur. Ces cellules sont prises en compte dans les tours. Ensuite, couvrant la région $1.1 < |\eta| < 1.4$ et fixé sur les parois des cryostats externes, le Détecteur Inter-Cryostat (*Inter-Cryostat Detector*, ou ICD) aide à la reconstruction des gerbes électromagnétique. Il est constitué d'une série de tuiles scintillantes couvrant $\Delta\eta \times \Delta\phi = 0.3 \times 0.3$ et disposées dans la

continuité des tours. Elles sont connectées à des photomultiplicateurs par des fibres à décalage de longueur d'onde et des fibres transparentes.

2.2.4 Le spectromètre à muons

Les muons sont les seules particules chargées qui parviennent à s'échapper du calorimètre, dans lequel ils déposent très peu d'énergie. Pour les détecter, mesurer leur trajectoire, leur impulsion et permettre un déclenchement sur des événements comportant des muons, deux détecteurs sont disposés autour du calorimètre, ainsi qu'un système d'aimants toroïdaux [38, 39]. La forme globale de l'ensemble est un parallélépipède entourant l'ensemble des autres sous-détecteurs présentés précédemment. Une partie de mon travail de thèse a consisté à développer une procédure de contrôle de l'alignement du spectromètre central. Elle sera présentée dans la section suivante 2.2.5.

2.2.4.1 Les aimants toroïdaux

Ils sont représentés sur la figure 2.2. Un aimant central long de 7.51 m est positionné de 3.18 m à 4.25 m du faisceau (entre les couches de chambres à dérive dites A et B, cf. section suivante), et deux aimants externes sont placés entre 4.54 m $< |z| <$ 6.10 m. Ils sont constitués de fer magnétisé. Les champs magnétiques internes sont d'environ 1.8 T pour l'aimant central et 1.9 T pour les aimants externes, et sont orthogonaux aux faisceaux. La polarité des aimants est régulièrement inversée durant la prise de donnée pour limiter les effets systématiques. Ces aimants permettent de mesurer l'impulsion des muons détectés, de façon peu précise mais indépendamment des autres sous-détecteurs.

2.2.4.2 Le détecteur central

Il est constitué du toroïde central présenté précédemment et de plusieurs couches de chambres à dérive et de scintillateurs, et il assure une couverture en pseudo-rapidité jusque $|\eta| < 1.0$. La partie inférieure du détecteur n'est pas totalement instrumentée, afin de laisser place au système de soutien du dispositif.

• **Chambres à dérive : WAMUS**
Trois couches de Chambres Proportionnelles à Dérive (*Proportional Drift Tubes*, ou PDT) forment le détecteur WAMUS (pour *Wide Angle MUon Spectrometer*). La première couche, ou couche A, est située à l'intérieur du toroïde central, et les deux autres couches, B et C, hors du toroïde. Elles couvrent la région $|\eta| < 1$ [40]. Environ 55% de la région centrale est couverte par les trois couches simultanément, et 90% par au moins deux couches.

Chaque couche est constituée d'un certain nombre de chambres à dérive juxtaposées. La figure 2.17 illustre leur disposition et leur nombre (cf. la figure du haut). Le détecteur WAMUS est constitué d'un total de 94 chambres à dérive, réparties ainsi : 18 forment la couche A, 38 forment la couche B, 38 forment la couche C. Le nombre de chambres diffère selon la face. Le tableau 2.7 donne les nombres de chambres par face, pour chacune des trois couches.

Chaque chambre à dérive forme une structure rectangulaire de 2.8 m selon l'axe du faisceau et de 5.8 m selon la direction perpendiculaire au faisceau. Une chambre est composée de 3 ou 4 plans superposés de cellules à derives, comme représenté sur la figure 2.14. Les couches B et C et la couche A inférieure comportent trois plans de cellules et les autres couches A, quatre. Ces plans sont légérement décalés les uns par rapport aux autres afin de lever l'ambiguïté gauche-droite du temps de dérive (voir ci-après). Chaque plan est constitué d'un ensemble de jusqu'à 24 cellules à dérive. Le détecteur comporte au total 6624 cellules.

Les cellules à dérive sont des tubes rectangulaires d'aluminium parallèles au champs magnétique (qui est lui orthogonal à la direction du faisceau). Cette disposition permet que la courbure des traces corresponde à la direction de la dérive. Ces tubes font toute la longueur de la chambre (2.8 m), pour une largeur de 10.1 cm et une hauteur de 5.7 cm. Les cellules d'un même plan sont toutes juxtaposées dans la même direction.

Chaque tube comporte en son centre une anode (en tungstène recouvert d'un film d'or) de 0.6 cm de diamètre étendue sur toute la longueur du tube et portée à un potentiel de 4.7 kV, et des cathodes au cuivre, sous forme d'un motif de losange de 61 cm se répétant, sur les faces de 10.1 cm de large (les parois supérieure et inférieure d'un tube). Les cathodes sont portées à un potentiel de 2.3 kV, les parois d'aluminium constituant la masse. La figure 2.15 donne une représentation du champ électrique régnant à l'intérieur d'un PDT. Ce système fournit l'information sur la position le long du fil du coup généré par le passage d'un muon. L'électronique de lecture est placée sur la face latérale externe des chambres. Sur l'autre face latérale, les anodes sont reliées deux à deux entre tubes voisins d'un même plan, au moyen d'une ligne à retard.

Ces cellules d'aluminium sont remplies d'un gaz composé de 84% d'Argon, de 8% de CH_4 et de 8% de CF_4. Ce gaz va ainsi s'ioniser au passage d'une particule chargée. La vitesse de dérive des charges ainsi produites est d'environ 10 cm.μs^{-1}, pour un temps de dérive maximum d'environ 500 ns. Ce temps est supérieur au temps de croisement des faisceaux, ce qui empêche de déterminer l'appartenance d'un muon à une collision donnée, et par extension, de permettre un déclenchement sur celui-ci. C'est entre autres pour pallier à cela que des scintillateurs ont été ajoutés. Le gaz circule en permanence, avec un flux de 500 litres par heure. Il est remis en circulation après avoir été filtré.

Les électrons produits par l'ionisation du gaz au passage du muon chargé vont se diriger principalement vers l'anode, en suivant les lignes de champ. Une plus faible quantité rejoindra les cathodes. De plus, les électrons, accélérés par le champ électrique, vont ioniser à leur tour le gaz, déclenchant ainsi un phénomène d'avalanche. Les ions positifs quant à eux vont rejoindre beaucoup plus lentement les parois d'aluminium de la cellule.

L'impulsion du muon est calculée à partir de sa courbure induite par le toroïde. Cette courbure est déterminée par la différence de pentes entre la ligne formée à partir du vertex primaire et des coups enregistrés dans la couche A, et la ligne formée à partir des coups dans les couches B et C.

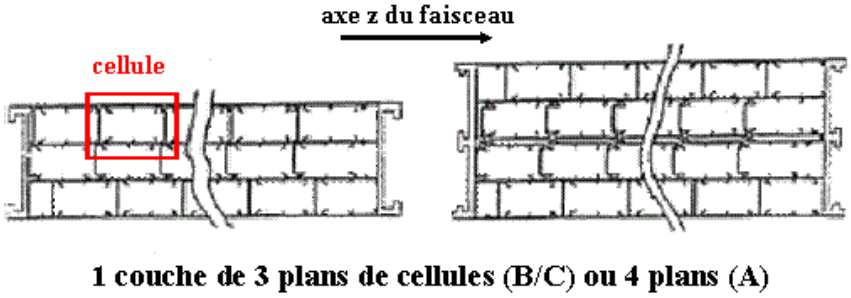

FIG. 2.14: Vue en coupe des plans de cellules à dérive. Un agrandissement d'une cellule (telle que celle encadrée en rouge) est donné sur la figure 2.15.

Pour chaque coup dans un PDT induit par le passage d'un muon, les informations suivantes sont enregistrées :

- le temps de dérive t_d des électrons produits par l'ionisation. Ce temps correspond à la différence entre le temps où le muon traverse la cellule (temps auquel se produit la collision augmenté du temps nécessaire au muon se déplaçant à la vitesse de la lumière pour atteindre la couche, et corrigé couche par couche, cf. section 2.2.5.1) et le moment où le signal est enregistré sur le fil. Ce temps de dérive peut être transformé en une distance de dérive (définie dans le plan médian de la cellule, le long de l'axe z du faisceau), comme nous le verrons dans la section 2.2.5.1. La résolution sur cette distance est de l'ordre de 500 μm. On en déduira alors la position du muon incident selon la direction z de l'axe du faisceau.
- la différence δt entre le temps d'arrivée du signal à l'électronique de lecture de la cellule touchée et le temps d'arrivée du signal à l'électronique de la cellule voisine qui lui est reliée. Cette information permet de déterminer la position du coup le long du fil avec une précision de 10 à 50 cm (selon que le coup soit éloigné ou proche de l'électronique).

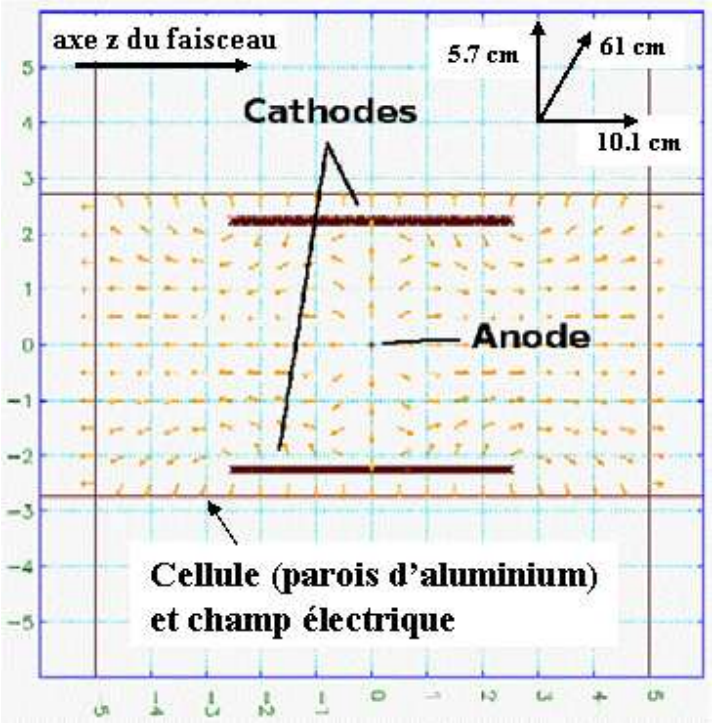

FIG. 2.15: Représentation du champ électrique régnant au sein d'une cellule à dérive. La cellule est visualisée selon sa plus petite section rectangulaire. Les dimensions sont données en cm.

– le dépôt de charge dans les cathodes. Les cathodes fournissent une mesure plus précise de la position le long du fil d'anodee, grâce à leur motif de losanges répétés. Combiné à l'information δt, le dépôt de charge amène à une précision de l'ordre de 5 mm sur la position du muon le long du fil (perpendiculaire au faisceau). Cependant, toutes les cathodes ne sont pas instrumentées. La résolution moyenne est donc moindre.

• **Les compteurs scintillants** $A\phi$

Les compteurs scintillants dits $A\phi$ sont installés devant la couche A des PDT (cf. figure 2.17). Ils permettent un déclenchement rapide sur les muons grâce à une résolution temporelle de l'ordre de 4 ns, en assurant l'association d'un muon traversant les PDT avec le croisement de faisceau correspondant, ainsi que la réjection des muons cosmiques ou provenant des interactions avec le tube à vide. Ils participent également à la reconstruction des trajectoires des muons.

Ces scintillateurs ont une segmentation en ϕ d'environ 4.5°, correspondant à celle du CFT (ce qui permet de combiner leurs informations pour le déclenchement), et une segmentation longitudinale correspondant à celle des PDT. Ils se chevauchent d'environ 3 % afin de détecter tout muon les traversant. Ces compteurs sont constitués de fibres

scintillantes à décalage de longueur d'onde reliées à des tubes photomultiplicateurs. Les scintillateurs donnent aussi le temps initial pour les PDTs.

• **Les compteurs scintillants *Cosmic Cap* et *Cosmic Bottom***
Ce sont un ensemble de compteurs scintillants (cf. figure 2.17) permettant également d'associer un muon dans un PDT avec le croisement de faisceau correspondant. L'information obtenue est utilisée par le système de déclenchement, ainsi que pour réduire le fond des rayons cosmiques.
Ces scintillateurs sont de deux types, dont la répartition par couche est indiquée dans le tableau 2.7 :

- Les *Cosmic Cap* : ils sont disposés sur la face externe des couches C supérieure et latérales des PDT (et un faible nombre sur les deux faces latérales de la couche B. Constitués de scintillateurs à fibres à décalage de longueur d'onde reliées à des phototubes, ils assurent une couverture de 4.5°en ϕ.
- Les *Cosmic Bottom* : ils sont placés sur la face externe des couches de PDT C et B inférieures. Leur structure est proche de celle des *Cosmic Cap*, mais à l'inverse de ces derniers, leur largeur est selon ϕ, leur longueur selon η (cf. figure 2.17).

Couche	A				B				C			
WAMUS PDT	18				38				38			
	$-x$	$+x$	$-y$	$+y$	$-x$	$+x$	$-y$	$+y$	$-x$	$+x$	$-y$	$+y$
	3	3	6	6	10	10	8	10	10	10	8	10
Cosmic Cap	-				16				240			
Cosmic Bottom	-				96				36			
$A-\phi$	630				-				-			

TAB. 2.7: Tableau récapitulatif du nombre de chambres à dérive et de compteurs scintillants par couche, pour le détecteur central. Concernant les PDT, leur nombre est également décomposé selon leur position. Pour la définition des axes x et y, se référer à la figure 2.17.

2.2.4.3 Les détecteurs avant et arrière : FAMUS

Le détecteur FAMUS (pour *Forward Angle MUon Spectrometer*) est constitué de plans de mini tubes à dérive (*Mini Drift Tubes*, ou MDT) accompagnés de scintillateurs à pixels, orientés dans le plan transverse à l'axe du faisceau. Il couvre la région $1.0 < |\eta| < 2.0$ [41].

Les couches de MDT sont au nombre de trois. Elles sont représentées sur la figure 2.17. La première couche, dite A elle aussi, est positionnée avant le toroïde et constituée de quatre plans de tubes, tandis que les deux autres couches, B et C, sont situées après l'aimant et composées de trois plans. Chaque couche est divisée en octants orientés le

long des lignes du champ magnétique (perpendiculairement au faisceau), selon x ou y (ne permettant donc pas de mesure selon ϕ). Un tube contient huit cellules d'une section de 0.94×0.94 cm^2 contenant chacune une anode de W-Au de 50 μm de diamètre, mise à la masse, baignant dans un mélange de $CF_4(90\ \%) - CH_4(10\ \%)$. Les cathodes sont, elles, portées à un potentiel de -3.2 kV. Le temps de dérive des électrons dans les tubes (environ 60 ns) est inférieur au temps de croisement des faisceaux, et ce système offre de plus une bonne résolution spatiale, permettant la reconstruction des segments de traces grâce à l'information récoltée sur la position de la trace et sa direction.

Les compteurs scintillants servent au déclenchement et à une détermination grossière de la trajectoire du muon [42]. Ils sont disposés en réseau de pixels, localisés sur la face interne des MDT A et C, et sur la face externe des MDT B. Chacune des couches est divisée en octants contenant environ 96 scintillateurs trapézoïdaux reliés à des phototubes positionnés sur 12 rangées concentriques, avec une segmentation en ϕ de 4.5°correspondant à celle du CFT, et une segmentation en η de 0.12 sur les neuf premières rangées, et 0.07 sur les trois dernières. Ces compteurs complètent l'information des MDT pour la détermination de la position en ϕ.

2.2.4.4 Le blindage

Le blindage, constitué de couches successives de fer (absorbe les gerbes électromagnétiques et hadroniques), de polyéthylène (absorbe de neutrons) et de plomb (absorbe les rayons gamma provenant de la capture des neutrons), est installé autour du tube à vide et au niveau des entrées du détecteur. Il permet de réduire les fonds dus aux résidus des protons et des antiprotons, et aux interactions des halos des faisceaux dans le tunnel, et assure également une protection face aux radiations.

2.2.5 La reconstruction des muons dans le spectromètre central

La reconstruction des muons traversant le spectromètre s'effectue en trois étapes successives, après traduction des informations provenant des cartes de lecture en quantités physiques. Tout d'abord, les coups dans chaque cellule sont reconstruits par l'intermédiaire de la distance de dérive. Ensuite les coups voisins ainsi déterminés sont reliés par des segments de droite pour chaque couche de PDT. La dernière étape consiste à apparier les divers segments obtenus (associés également aux coups dans les scintillateurs), afin de reconstruire la trace du muon et de calculer son impulsion à partir de la différence de pente entre le segment construit dans la couche A et ceux construits dans les couches B et C [43].
Par la suite, les informations issues du spectromètre pourront être combinées à leur tour avec celles provenant du calorimètre et du détecteur de traces interne, en particulier pour corriger l'impulsion du muon de sa perte d'énergie dans le calorimètre, et remesurer cette impulsion avec une bien meilleure résolution.

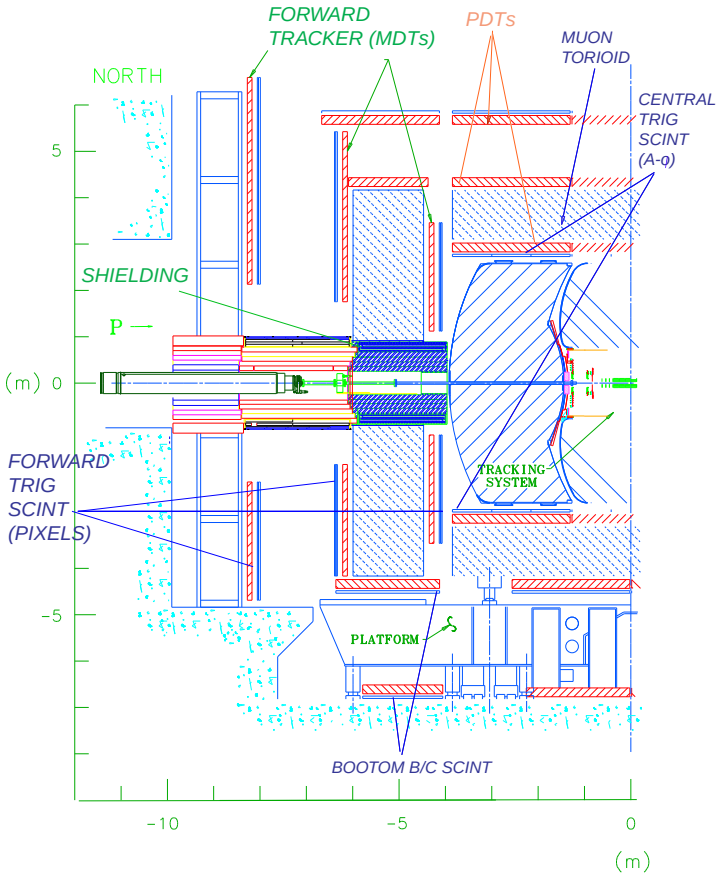

FIG. 2.16: Vue en coupe de la moitié gauche du détecteur à muons.

La procédure de reconstruction des coups et des segments dans les PDT va être détaillée dans les sections suivantes.

J'ai développé des procédures de contrôle et de correction de la résolution spatiale des chambres de PDT, de l'alignement des chambres et de la reconstruction des muons dans le spectromètre central, qui ont été mises à disposition de la collaboration dans le *package* *muon_PDT_monitoring*. De plus, des contrôles ont été effectués sur diverses périodes de prises de données, et les modifications nécessaires implémentées dans les programmes de reconstruction *online* et *offline* et de déclenchement. Cette section illustre ce travail, elle n'est toutefois ni exhaustive ni quantitative.

2.2.5.1 La détermination de la distance de dérive dans les PDT

Chaque fil d'anode des cellules a un système de lecture qui lui est propre, et qui inclut une horloge. On appelle T_0 le temps après lequel l'horloge est déclenchée, lorsque une

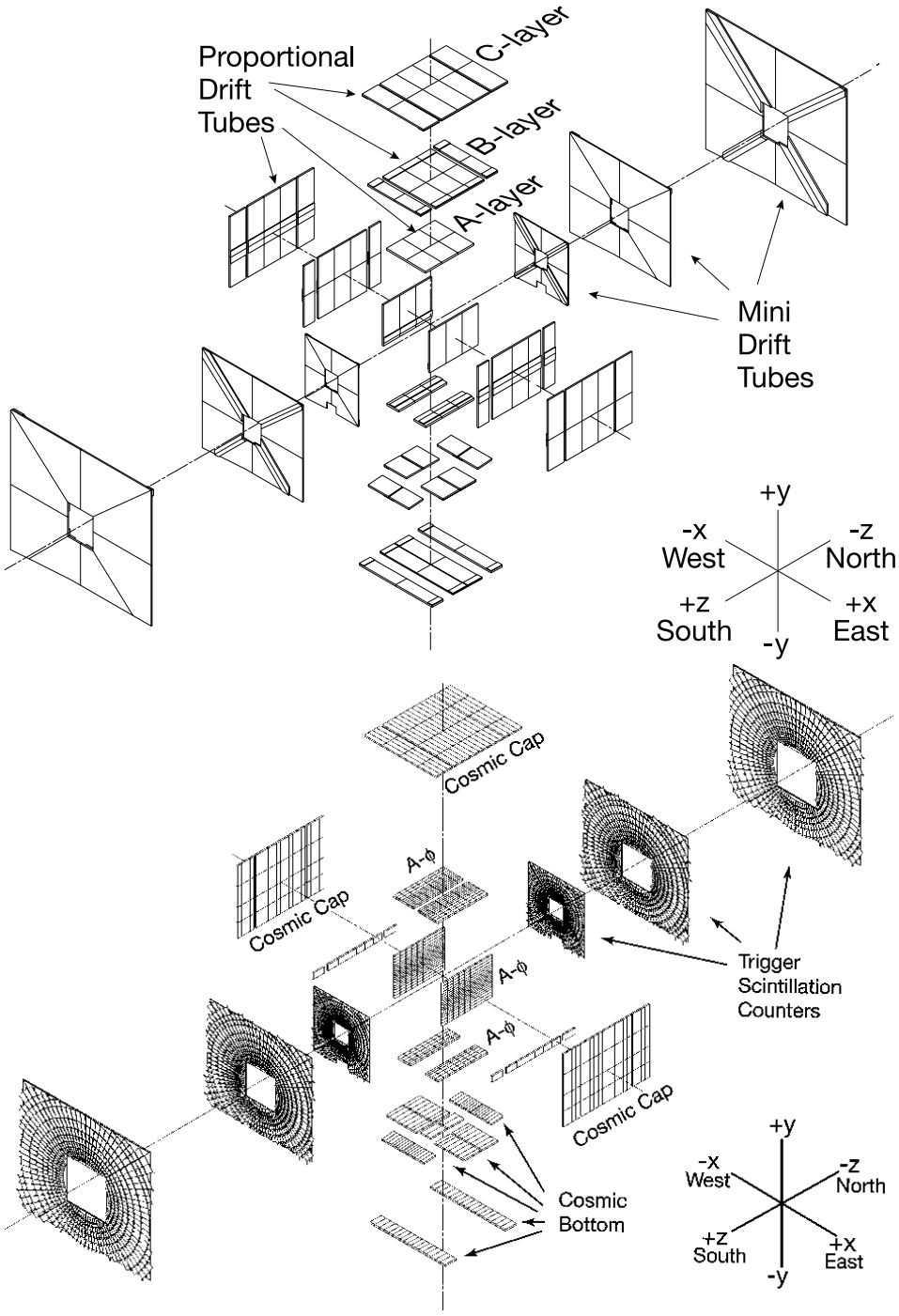

FIG. 2.17: Le spectromètre à muons (vue éclatée) : en haut, les tubes a dérive, et en bas, les scintillateurs.

collision a lieu. Le temps de dérive t_d, c'est-à-dire le temps que met le signal produit dans une cellule par le passage d'un muon pour atteindre l'anode, est ainsi enregistré. Les T_0, qui diffèrent pour chaque chambre à dérive, doivent donc être correctement décrits pour pouvoir mesurer précisement le temps de dérive. La relation entre le temps de dérive t_d et la distance de dérive d correspondante, indispensable à la reconstruction de la trajectoire du muon, ne suit pas une relation linéaire. Sa détermination n'est ainsi pas triviale. Elle dépend, entre autres, de l'angle d'incidence de la trace du muon, qui n'est pas accessible initialement. La figure 2.18 présente une coupe transversale d'une cellule à dérive, traversée par un muon ayant un angle d'incidence θ, et définit la distance de dérive d (notée X sur la figure).

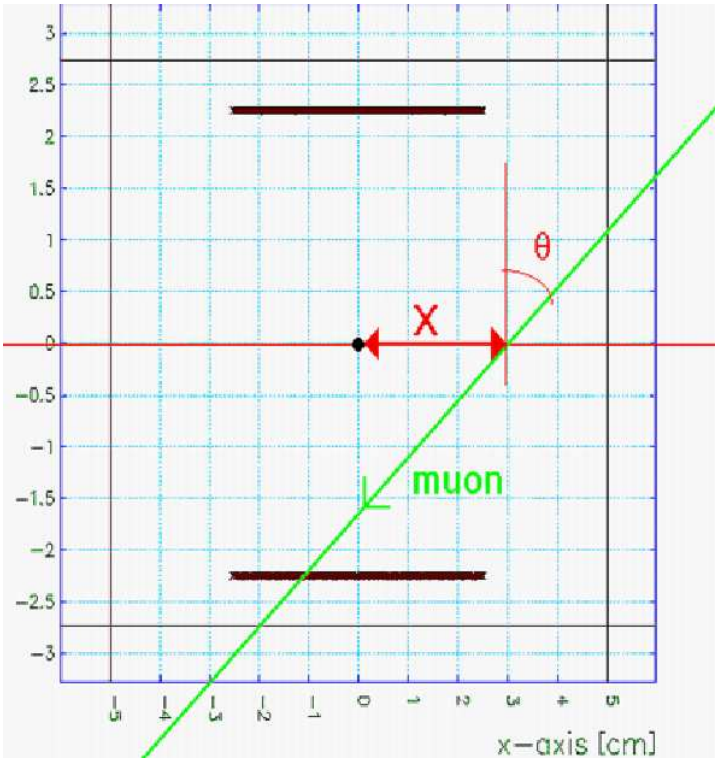

FIG. 2.18: Définition de l'angle d'incidence θ d'une trace d'un muon traversant une cellule à dérive, et de la distance au fil X (nommée d dans le texte) dans le plan médian de la cellule.

En se basant sur l'expérience du *Run I*, la paramétrisation suivante de la relation temps-distance $d(t_d, \theta)$ a été déterminée [44] :

$$\begin{aligned} d(t_d,\theta) &= (\alpha_{10} + \alpha_{11}\theta + \alpha_{12}\theta^2 + \alpha_{13}\theta^3)t_d \\ &+ (\alpha_{20} + \alpha_{21}\theta + \alpha_{22}\theta^2 + \alpha_{23}\theta^3)t_d^2 \\ &+ (\alpha_{30} + \alpha_{31}\theta + \alpha_{32}\theta^2 + \alpha_{33}\theta^3)t_d^3 \\ &+ (\alpha_{40} + \alpha_{41}\theta + \alpha_{42}\theta^2 + \alpha_{43}\theta^3)t_d^4 \end{aligned}$$

où t_d correspond au temps de dérive mesuré et θ à l'angle d'incidence de la trace du muon. Les paramètres $\alpha_{i,j}$ ont été déterminés expérimentalement à l'aide d'un dispositif spécifique de chambres à dérive reproduisant les conditions des couches A, B et C. Lors de l'étude que nous avons menée, les paramètres étaient initialement distincts entre la couche A (et quatre couches B et C particulières : celles correspondant à la partie inférieure du détecteur) et les couches B et C. La différence provient du fait que le système de circulation du gaz est différent entre les couches A et les couches B et C. Ainsi, le contenu en eau, qui réduit la vitesse de dérive, n'est pas le même. Les tables 2.8 et 2.9 listent les paramètres $\alpha_{i,j}$ utilisés lors de cette étude pour les deux types de gaz [45]. Depuis l'arrêt du détecteur de 2006, le même gaz circule dans toutes les couches, et un seul jeu de paramètres est déterminé. La présente étude est antérieure à ce changement.

Pour une valeur fixée de θ, l'équation précédente devient un polynôme d'ordre 4 selon le temps de dérive :

$$d(t_d) = \beta_1 t_d + \beta_2 t_d^2 + \beta_3 t_d^3 + \beta_4 t_d^4$$

Il est de plus possible d'appliquer un facteur correctif SF_{couche} directement sur cette distance, ce qui a pour effet de modéliser une modification de la vitesse de dérive :

$$d(t_d) = (\beta_1 t_d + \beta_2 t_d^2 + \beta_3 t_d^3 + \beta_4 t_d^4) \cdot SF_{A,B,C}$$

En pratique, deux facteurs correctifs seront appliqués, l'un pour la couche A et les chambres B et C spécifiques, l'autre pour les couches B et C. Ces facteurs correctifs sont eux aussi fournis au programme *d0reco*.

$\alpha_{i,j}$	θ^0	θ^1	θ^2	θ^3
$\alpha_{1,j}$	$9.85253 \cdot 10^{-3}$	$1.01407 \cdot 10^{-4}$	$-7.72285 \cdot 10^{-6}$	$1.97738 \cdot 10^{-7}$
$\alpha_{2,j}$	$1.33748 \cdot 10^{-6}$	$-5.88543 \cdot 10^{-7}$	$4.89143 \cdot 10^{-8}$	$-1.15444 \cdot 10^{-9}$
$\alpha_{3,j}$	$-2.47611 \cdot 10^{-9}$	$1.21509 \cdot 10^{-9}$	$-9.32089 \cdot 10^{-11}$	$1.94994 \cdot 10^{-12}$
$\alpha_{4,j}$	$-1.54404 \cdot 10^{-13}$	$-7.37766 \cdot 10^{-13}$	$4.62811 \cdot 10^{-14}$	$-6.91967 \cdot 10^{-16}$

TAB. 2.8: Paramétrisation de la relation temps-distance pour la couche A et les quatres couches B et C particulières.

$\alpha_{i,j}$	θ^0	θ^1	θ^2	θ^3
$\alpha_{1,j}$	$9.13928 \cdot 10^{-3}$	$4.29118 \cdot 10^{-5}$	$-2.96607 \cdot 10^{-6}$	$6.64569 \cdot 10^{-8}$
$\alpha_{2,j}$	$-2.82045 \cdot 10^{-8}$	$-1.05518 \cdot 10^{-7}$	$8.48049 \cdot 10^{-9}$	$-2.10862 \cdot 10^{-10}$
$\alpha_{3,j}$	$1.66737 \cdot 10^{-9}$	$1.75045 \cdot 10^{-10}$	$-9.74414 \cdot 10^{-12}$	$1.91771 \cdot 10^{-13}$
$\alpha_{4,j}$	$-4.73803 \cdot 10^{-12}$	$-1.79514 \cdot 10^{-13}$	$3.81363 \cdot 10^{-15}$	$3.83472 \cdot 10^{-17}$

TAB. 2.9: Paramétrisation de la relation temps-distance pour les couches B et C.

La connaissance des T_0 de chacune des 94 chambres est indispensable. Ceux-ci sont implémentés dans les programmes *online* et *offline* de reconstruction des coups et de calcul des temps de dérive dans le spectromètre central (*d0reco*). Ces temps T_0 incluent le temps physique que met le muon pour atteindre, à la vitesse de la lumière, la couche correspondante, ainsi que le temps instrumental correspondant à la propagation du signal le long des câbles et son traitement par l'électronique de lecture. Ceci explique que les T_0 diffèrent d'une chambre à l'autre, puisqu'ils dépendront de leur position, de leur cablâge et de leur électronique de lecture. Il est indispensable de vérifier régulièrement que les T_0 chargés sont corrects, et de procéder à leur correction le cas échéant. En effet le gaz surtout, mais également le cablâge et l'électronique de lecture peuvent évoluer au cours des années, en particulier après les périodes d'arrêt du détecteur.

2.2.5.2 Détermination des T_0 et procédures de contrôle de l'alignement

Les données utilisées pour déterminer les T_0 sont des événements ayant subi une présélection assez lâche et contenant un muon associé à une trace d'impulsion transverse $P_T > 5$ GeV/c dans le détecteur de trace central. De plus, des coups doivent avoir été enregistrés dans tous les plans des chambres traversées.

Une fois les coups reconstruits grâce à la détermination des distances de dérive, ils sont associés, au sein d'une chambre donnée, pour former des segments (la courbure due au champ magnétique est totalement négligeable au sein d'une chambre). Cependant, bien que la distance de dérive ait été déterminée, il reste une inconnue. Le système ne permet pas en effet de savoir directement si le muon est passé à gauche ou à droite de l'anode. Ceci est appelé l'ambiguïté gauche-droite. Pour lever cette ambiguïté, plusieurs segments sont reconstruits à partir des coups enregistrés en imposant un choix sur leur position. Ainsi, pour une chambre comportant 3 plans de PDT nous pouvons reconstruire 8 segments en combinant les trois coups enregistrés. Une couche comportant 4 plans de cellules permettra de reconstruire 16 segments. Le segment sélectionné parmi toutes les combinaisons est celui dont le χ^2 de l'ajustement sur les coups est le plus faible.
Maintenant que le segment est connu, on peut déterminer l'angle d'incidence de la trace du muon dans cette chambre traversée. Une procédure itérative est alors menée afin de recalculer les distances de dérive (et donc la position des coups) dans chaque PDT tra-

versé et de recalculer le segment en imposant cette nouvelle valeur de l'angle d'incidence (l'angle était pris nul lors de la première itération). Finalement, le segment est retenu si l'angle d'incidence converge vers une valeur et le χ^2 de l'ajustement sur le segment est inférieur à un certain seuil. Cette procédure itérative est utilisée pour la couche A, mais pas pour les couches B et C. Ces deux couches se trouvant du même côté par rapport au toroïde, les deux segments déterminés dans les deux couches sont théoriquement colinéaires. L'ambiguïté gauche-droite est levée séparement pour les couches B et C, puis le segment BC est recherché. Pour les couches B et C, l'angle d'incidence correspond à l'angle que fait la droite reliant le barycentre du segment B au barycentre du segment C.

Lors de cette reconstruction des segments, nous pouvons jouer sur deux types de paramètres : les T_0 de chaque chambre et le facteur correctif SF_{couche} sur la vitesse de dérive appliqué à chaque couche. Dans un premier temps, nous avons déterminé les corrections à apporter aux T_0 enregistrés dans la base de données de la reconstruction (pour des données enregistrées en juillet 2004 et janvier 2005). Ces ΔT_0 sont calculés ainsi :

- l'angle d'incidence du muon est calculé par la méthode itérative présentée précédemment, sans correction des T_0, pour chaque événement du lot de données étudié ;
- en utilisant ces angles et en appliquant une correction donnée ΔT_0 sur le T_0 d'une chambre, la somme des χ^2 de l'ajustement du segment correspondant pour tous les muons des événements utilisés est calculée ;
- pour chaque chambre, la correction ΔT_0 retenue et implementée dans les algorithmes de reconstruction sera celle minimisant la somme des χ^2.

La distribution des résidus Δd_i de chaque plan de cellules permet d'étudier la qualité de la correction. Ces résidus sont définis comme suit : $\Delta d_i = d_i(t) - d_i^{seg}(t)$ où $d_i(t)$ est la distance de dérive enregistrée dans une cellule d'un plan i et $d_i^{seg}(t)$ la distance de dérive correspondant au segment ajustant les coups déterminé précédemment. Ils permettent de s'assurer que les corrections appliquées améliorent réellement la résolution obtenue. La figure 2.19 correspondant à des données enregistrées en juillet 2004 illustre l'influence de la correction des T_0 pour une chambre de PDT de la couche A. A gauche est représentée la distribution des résidus calculés sans correction des T_0, à droite est représentée la même distribution mais déterminée avec la correction des temps de dérive. Un ajustement gaussien de la distribution des résidus est également reporté sur ces figures. De la largeur σ_i de cette gaussienne ajustant la distribution des Δd_i, nous pouvons déduire la résolution du plan de cellules i d'une chambre donnée, au moyen des formules suivantes, pour une chambre comportant quatre plans de cellules :

- $\sigma = \sqrt{3} \times \sigma_{1,4}$
- $\sigma = \sqrt{3/2} \times \sigma_{2,3}$

et pour une chambre comportant trois plans de cellules :

- $\sigma = \sqrt{6} \times \sigma_{1,3}$

– $\sigma = \sqrt{3/2} \times \sigma_2$

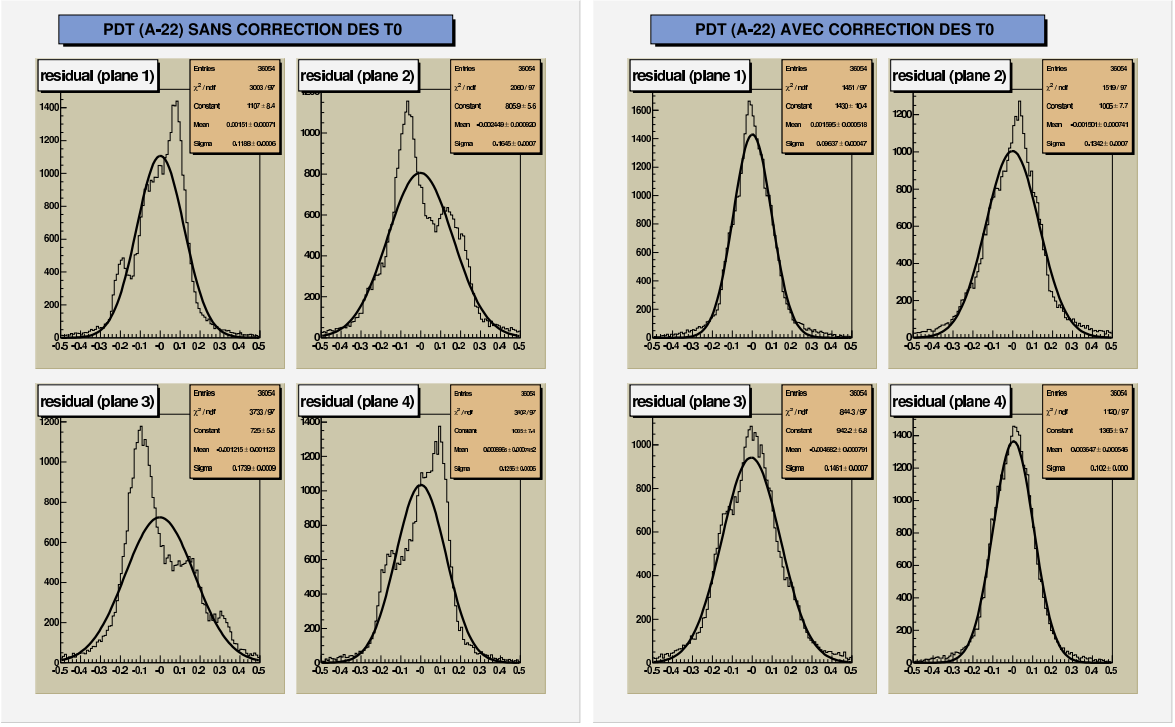

FIG. 2.19: Distributions des résidus déterminés pour chacun des quatre plans de cellules de la chambre de PDT 22 de la couche A. Les figures de gauche ont été obtenues sans correction des T_0, celles de droite avec correction. Un ajustement par une gaussienne est superposé aux distributions. L'échelle des abscisses est le centimètre.

Alors que dans un premier temps nous comptions uniquement corriger les T_0 des chambres, il s'est avéré que les résultats obtenus n'étaient pas très bons, ni parfaitement compris. Nous nous sommes aperçus alors qu'une augmentation de la vitesse de dérive (par l'intermédiaire des facteurs correctifs SF_{couche}), couplée à une mise à jour des T_0, était nécessaire et qu'elle améliorait grandement la résolution. De plus, l'utilisation d'un seul jeu de paramètres $\alpha_{i,j}$ pour toutes les couches, correspondant à celui utilisé par la couche A (tableau 2.8), améliore plus encore les résultats obtenus. Par la suite, ce seul jeu de paramètres sera utilisé, et les T_0 et vitesses de dérive seront régulièrement corrigés.

En sus des distributions des résidus, de nombreuses autres variables permettant de contrôler la qualité de la reconstruction des traces et de l'alignement des segments dans les couches de PDT ont été implémentées dans cette procédure de contrôle. Les figures 2.20 et 2.21 présentent, à titre illustratif, les distributions de quelques unes de ces variables. Il est possible de comparer des couches superposées (telles que A-021,

B-121 et *C*-221, présentées sur les figures 2.20 et 2.21) mais également des couches adjacentes, puisque la trace d'un muon peut dans certains cas être compatible avec une telle configuration.

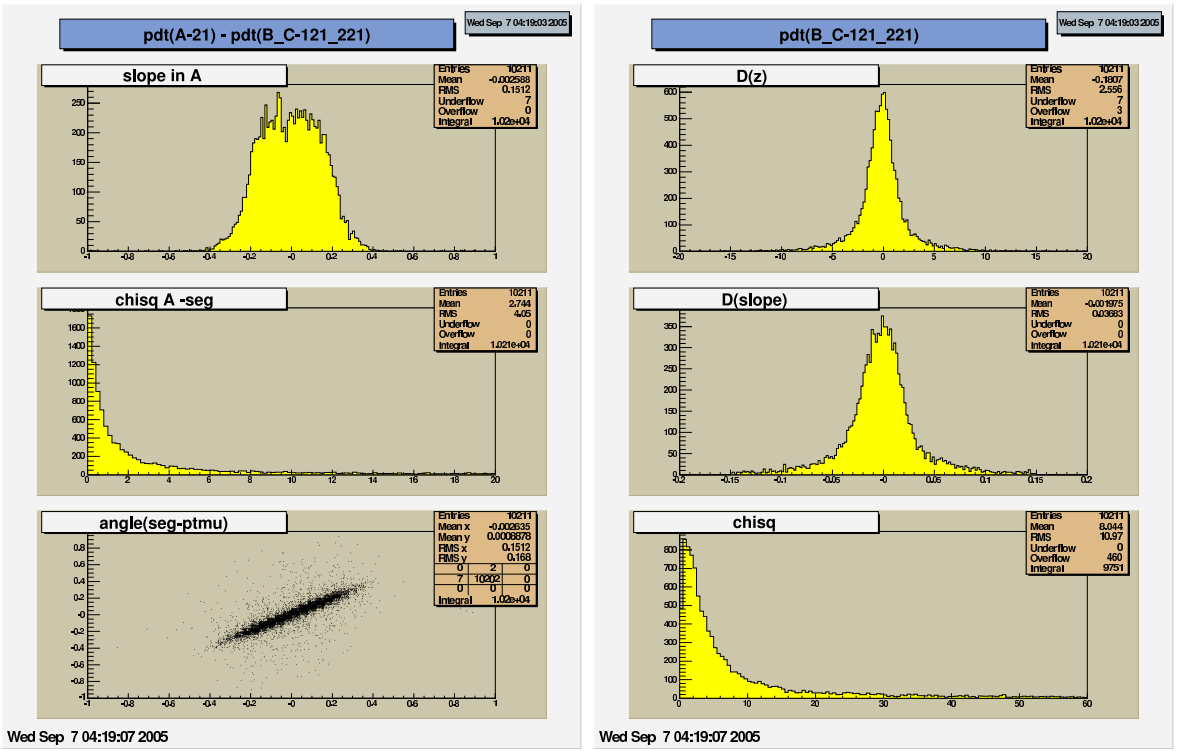

FIG. 2.20: Distributions de quelques variables de contrôle de la reconstruction des muons dans le spectromètre central, et de contrôle de la qualité de l'alignement des PDT.
La figure de gauche correspond au PDT 21 de la couche A. De haut en bas :
– distribution de la pente des segments reconstruits (l'échelle des abscisses est le centimètre) ;
– distribution du χ^2 de l'ajustement des segments ;
– comparaison de la pente déterminée à partir du segment à celle issue de l'impulsion transverse du muon.
La figure de droite présente des variables concernant la combinaison des segments reconstruits dans les PDT 121 de la couche B et 221 de la couche C. De haut en bas :
– distribution (en cm) de la distance, selon l'axe du faisceau, entre l'intersection du segment C au plan médian des couches B et C et l'intersection du segment B au même plan ;
– distribution (en cm) de la différence entre la pente du segment de la chambre C et la pente du segment de la chambre B ;
– χ^2 de l'ajustement du segment combiné BC.

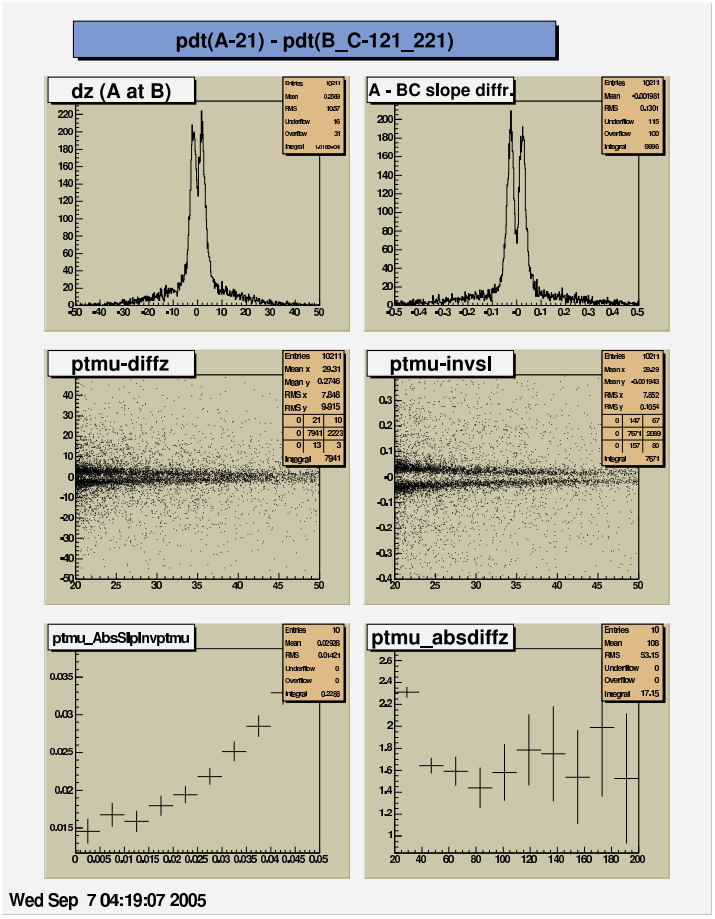

FIG. 2.21: Contrôle de l'alignement des segments des couches A, B et C.
- Figure haut gauche : distribution de la distance Δz (en cm), selon l'axe du faisceau, entre l'intersection du segment BC au plan médian des couches A et B et l'intersection du segment A au même plan ;
- Figure haut droite : distribution $\Delta(pente)$ (en cm) de la différence entre la pente du segment de la chambre A et la pente du segment combiné BC ;
- Figure milieu gauche : distribution bidimensionnelle de Δz en fonction de l'impulsion transverse du muon $ptmu$ (en GeV/c) ;
- Figure milieu droite : distribution bidimensionnelle de $\Delta(pente)$ en fonction de $ptmu$;
- Figure bas gauche : distribution normalisée de $|\Delta(pente)|$ en fonction de $1/ptmu$;
- Figure bas droite : distribution normalisée de $|\Delta z|$ en fonction de $ptmu$.

2.2.6 Le détecteur de protons avant

Le détecteurs de protons avant (FPD, pour *Forward Proton Detector*) mesure les protons et les antiprotons dispersés aux faibles angles (de l'ordre du mrad). Il est indispensable pour accéder à la totalité de la cinématique de la particule dispersée. Le FPD est constitué d'un ensemble d'aimants combinés avec des détecteurs de position à fibres scintillantes placés le long du tube à vide (cf. figure 2.22). Les détecteurs de positions sont placés à quelques millimètres du faisceau, afin de maximiser l'acceptance du FPD. Mais cette proximité impose que ces détecteurs soient éloignés du faisceau lors de la phase d'injection des protons, ainsi que dans les conditions de haute luminosité. Ils sont au nombre de 18, contenus dans des pots Romains et regroupés en six "châteaux" (cf. figure 2.22) dans lesquels le vide est fait. Six couches de fibres scintillantes orientées selon trois directions différentes offrent une résolution spatiale de 130 μm. L'information du FPD est en particulier utilisée au niveau 1 du système de déclenchement.

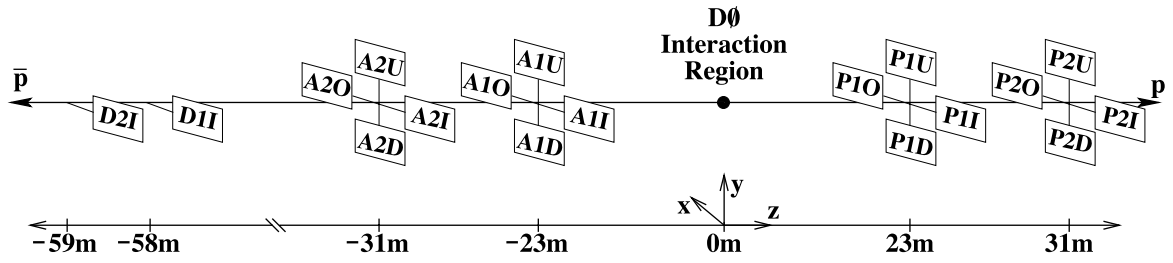

FIG. 2.22: Schéma du FPD et localisation des détecteurs de position (pots romains). Un "château" correspond au regroupement des pots romains appartenant au même plan transverse à l'axe du faisceau. Ils sont au nombre de six, comportant 1 ou 4 détecteurs de position.

2.2.7 La luminosité

La luminosité délivrée par un collisionneur conditionne la physique accessible. Une bonne connaissance de la luminosité enregistrée par l'expérience est indispensable pour de nombreuses analyses menées par la collaboration DØ, par exemple pour normaliser une contribution de bruit de fond ou pour extraire une section efficace à partir d'un taux de comptage.

On distingue la luminosité instantanée $\mathcal{L}$, qui correspond à un nombre de collisions se produisant par unité de temps (elle s'exprime généralement en $\mathrm{cm}^{-2} \cdot \mathrm{s}^{-1}$) et la luminosité intégrée L sur une certaine durée T, durant laquelle le collisionneur produit

des interactions (elle s'exprime en cm^{-2} ou plus communément[5] en pb^{-1} ou fb^{-1}) :

$$\mathcal{L} = \frac{1}{\sigma_{eff}} \frac{dN}{dt} \tag{2.3}$$

où dN/dt correspond au nombre d'interactions inélastiques par unité de temps et σ_{eff} à la section efficace des collisions inélastiques. La luminosité instantanée $\mathcal{L}$ peut s'exprimer en fonction des caractéristiques du collisionneur :

$$\mathcal{L} = \frac{fBN_pN_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)} F(\frac{\sigma_1}{\beta^*}) \tag{2.4}$$

avec :
- f : fréquence de rotation des paquets ;
- B : nombre de paquets par faisceau ;
- N_p ($N_{\bar{p}}$) : nombre de protons (d'antiprotons) par paquet ;
- σ_p ($\sigma_{\bar{p}}$) : taille transverse des faisceaux au point d'interaction ;
- $F(\frac{\sigma_1}{\beta^*})$: facteur de forme.

Les caractéristiques des faisceaux indiquées ci-dessus n'étant pas aisément accessibles avec précision, des luminomètres (figure 2.23) ont été installés entre les bouchons du calorimètre et le trajectographe du détecteur DØ afin de procéder à une mesure précise de la luminosité par comptage de coïncidences [46]. Ces luminomètres détectent les collisions inélastiques $p\bar{p}$ et permettent également de mesurer les taux d'interaction des halos des faisceaux, afin de réaliser une mesure rapide de la coordonnée z du vertex primaire et d'identifier les croisements à interactions multiples. Ces luminomètres sont constitués de 24 sections formant un disque de scintillateurs plastiques et de photomultiplicateurs, ils sont placés à $z = \pm 140$ cm (où règne un champ magnétique plus faible, d'environ 1 T) et couvrent la région $2.7 < |\eta_{det}| < 4.4$. Lorsque les faisceaux se croisent, les partons spectateurs des protons et des antiprotons concernés par les interactions produisent des jets de particules qui iront traverser les compteurs de luminosité. Par l'étude des coïncidences avant-arrière, il est possible de remonter au nombre de collisions survenues pendant un laps de temps donné, et de distinguer les interactions $p\bar{p}$ du fond induit par les halos des faisceaux.

Si f est la fréquence de croisement des faisceaux et μ le nombre moyen d'interactions par croisement, on a $dN/dt = \mu f$. Or μ suivant une loi de Poisson, on peut en déduire la probabilité d'avoir n interactions par croisement : $P(n) = \mu^n e^{-\mu}/n!$. On en déduit la probabilité de mesurer un signal dans les luminomètres, qui correspond à la détection d'au moins une interaction : $P(n > 0) = 1 - e^{-\mu}$. La luminosité instantanée s'exprime finalement ainsi :

$$\mathcal{L} = -\frac{f}{\sigma_{eff}} ln(1 - P(n > 0)) \tag{2.5}$$

[5] 1 barn = 1 b = 10^{-24} cm^{-2}

La section efficace σ_{eff} est composée de trois types de processus inélastiques : diffractifs simples, doublement diffractifs, et non diffractifs. Elle vaut $\sigma_{eff} = 43.26 \pm 2.07$ mb. Ce résultat tient compte de l'acceptance, de l'efficacité et des temps morts des luminomètres. L'incertitude totale sur la détermination de la luminosité finale est estimée dans DØ à 6.1 %, et provient essentiellement de l'erreur sur la mesure de l'efficacité du détecteur et de l'erreur sur σ_{eff}.

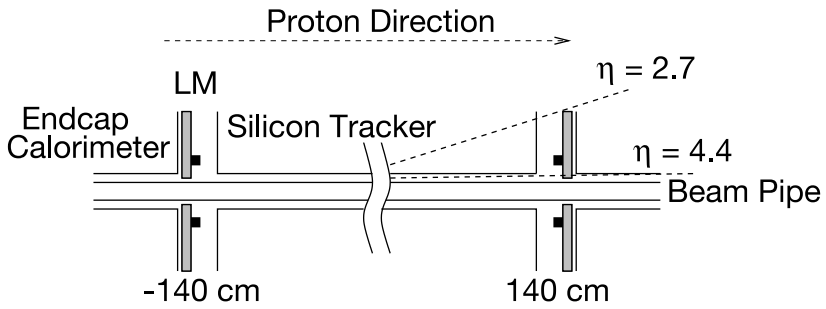

FIG. 2.23: Les luminomètres (en gris) sont disposés de part et d'autre du détecteur DØ, entre les bouchons du calorimètre et le trajectographe. En noir sont représentés les photomultiplicateurs.

Chaque événement enregistré par DØ est complètement identifié par un numéro d'événement, un numéro de *Run* (séquence de prise de données de une à quatre heures regroupant généralement plusieurs milliers d'événements consécutifs) et un numéro de bloc de luminosité (LBN). Un bloc de luminosité correspond à l'intervalle de temps durant lequel les quantités relatives à la luminosité sont enregistrées. Cet intervalle de temps est nominalement de 60 s. La figure 2.24 montre l'évolution des luminosités délivrée par le Tevatron et enregistrée par l'expérience DØ au cours du *Run II*, d'avril 2002 à avril 2008. Au mois d'avril 2008, les luminosités délivrée et enregistrée s'élèvent respectivement à 3.99 fb^{-1} et 3.44 fb^{-1}. On notera la présence de paliers qui correspondent aux périodes durant lesquelles le Tevatron était en arrêt pour maintenance, introduction de nouveaux détecteurs et amélioration de ses performances. La seconde figure illustre l'efficacité de la prise de donnée au cours de ces 6 ans. Elle est assez élevée, de l'ordre de 90 %.

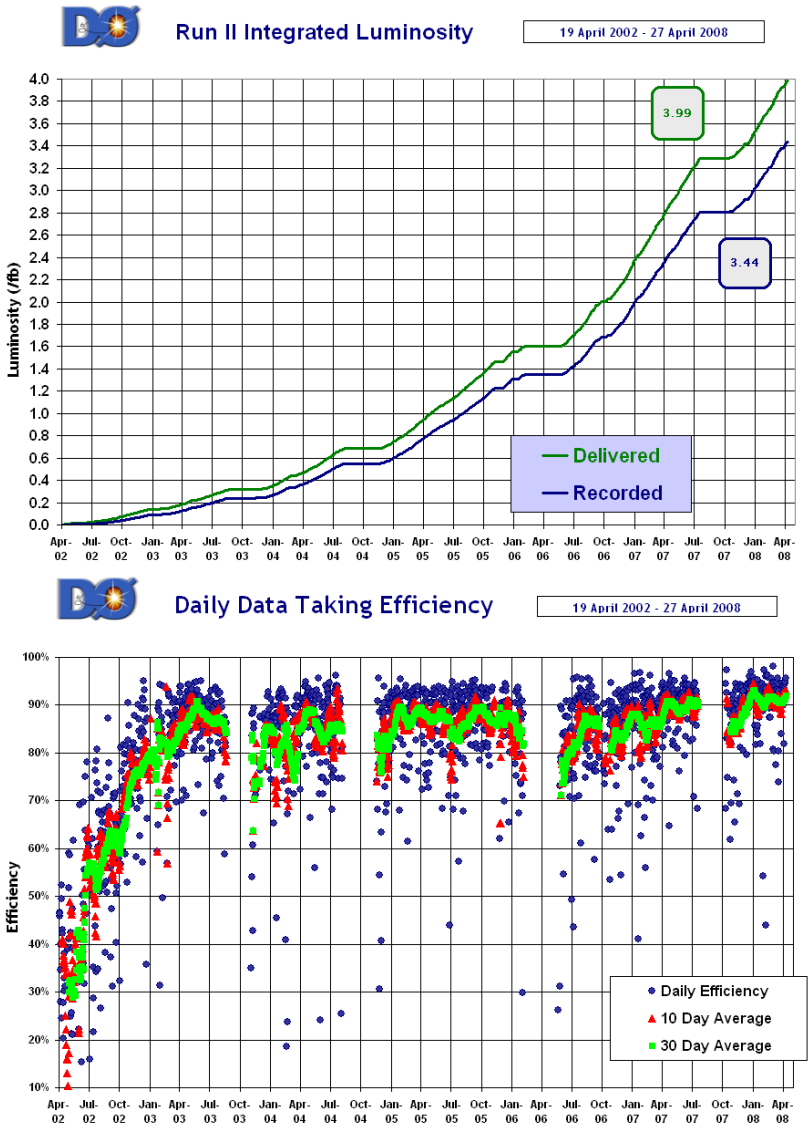

FIG. 2.24: Figure du haut : Luminosité intégrée délivrée par le Tevatron (en vert) et enregistrée par l'expérience DØ (en bleu) du début du *Run II* à nos jours (avril 2008). Une luminosité intégrée totale de 3.99 fb^{-1} a été délivrée en six ans, dont 3.44 fb^{-1} ont été enregistrés. Le *Run IIa* s'étend d'avril 2002 à février 2006, et le *Run IIb* a débuté en juin 2006. Les plateaux correspondent aux périodes de *shutdown* du collisionneur.
Figure du bas : Efficacité de la prise de données journalière (points bleus), ou intégrée sur 10 jours (triangles rouges) et 30 jours (carrés verts) d'avril 2002 à avril 2008.

2.3 Le système de déclenchement

Le système de déclenchement est un élément primordial des expériences placées auprès des collisionneurs hadroniques. En effet, des limitations pratiques évidentes empêchent que tous les événements (signaux mesurés dans les détecteurs lorsqu'une collision entre protons et antiprotons a eu lieu) produits soient enregistrés pour être analysés plus tard, du fait du très faible temps de croisement des faisceaux (396 ns, soit 2.5 MHz).

Les différents critères de déclenchement (ou *triggers*) seront élaborés dans l'optique de ne garder que les processus durs Seuls les événements passant certains critères seront enregistrés. Une autre utilité du système de déclenchement est de classer par catégories les événements sélectionnés. Cela permet ainsi, par le choix de *triggers* bien spécifiques pour une analyse donnée, de procéder à une première préselection des données enregistrées.

Le système de déclenchement de DØ est organisé en 3 niveaux successifs utilisant des critères de plus en plus évolués, désignés par L1, L2 et L3. Il est représenté schématiquement sur la figure 2.25, sur laquelle on remarquera le *Trigger Framework* qui est chargé de la gestion globale du déclenchement. Le système de déclenchement doit être suffisamment rapide pour traiter la totalité des événements produits lors des collisions. Il a été entièrement repensé entre le *Run I* et le *Run II* du fait de l'augmentation de la luminosité. Une nouveauté par rapport au *Run I* est l'introduction de mémoires (*buffers*) entre chaque niveau de déclenchement pour limiter le temps mort.
Un niveau "0" de déclenchement existe par ailleurs. En effet, lors d'une collision inélastique, les fragments du proton et de l'antiproton n'ayant pas participé à l'interaction sont émis dans des directions proches de l'axe des faisceaux. Si les détecteurs de luminosité mesurent un signal en coïncidence temporelle, les événements correspondants (dit de biais minimum, ou *minimum bias*) sont alors et alors seulement traités par le niveau 1.
Enfin, certains types d'événements ont une fréquence d'apparition trop élevée pour pouvoir être tous traités. Un *prescale* est alors appliqué, c'est-à-dire que seule une fraction de ces événements est conservée.

2.3.1 Le niveau 1

Ce niveau purement électronique n'utilise aucun algorithme de reconstruction. Il base ses décisions sur les informations brutes de tous les sous-détecteurs (calorimètre, CFT, détecteurs de pieds de gerbe, spectromètre à muons - cf. figure 2.26), hormis le SMT qui ne possède pas une électronique de lecture suffisamment rapide. En combinant ces informations, 128 conditions logiques "AND/OR" sont obtenues et utilisées afin de décider si l'événement devra être traité par le niveau 2. La décision est prise par la combinaison des quatres sous-systèmes de déclenchement suivants :

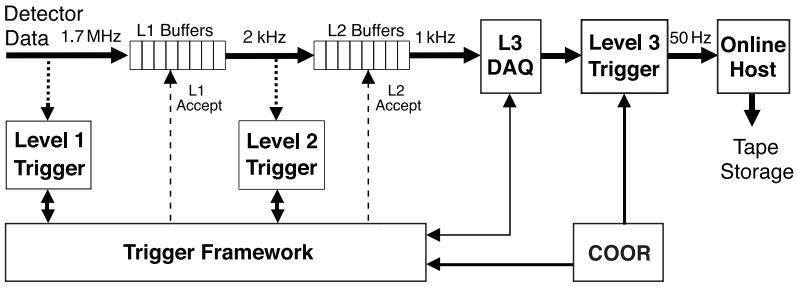

FIG. 2.25: Représentation schématique du système de déclenchement.

- **L1CAL :** déclenchement sur les électrons, les jets, et les énergies transverses totale et manquante dans le calorimètre. A l'aide des dépots d'énergie dans les tours de déclenchement du calorimètre de largeur $\Delta\eta \times \Delta\Phi = 0.2 \times 0.2$ (regroupant quatre tours ordinaires) et des détecteurs de pieds de gerbe pour les électrons.
- **L1CTT :** déclenchement sur les traces, par l'étude des coups dans le CFT, le CPS, et le FPS.
- **L1MUO :** déclenchement sur les muons. Leur impulsion est reconstruite grâce au CFT et aux scintillateurs du système d'identification des muons. Les détecteurs de muons permettent de disposer d'un déclenchement très stable et efficace à haute luminosité. Des fenêtres temporelles sont utilisées pour rejeter le bruit de fond et les rayons cosmiques.
- **L1FPD :** déclenchement sur les protons et les antiprotons détectés par le FPD.

L'acceptance du calorimètre au niveau 1 s'étend jusqu'à $|\eta_{det}| = 3.2$, celle du CFT jusqu'à $|\eta_{det}| = 1.6$, celle du détecteur de pied de gerbes jusqu'à $|\eta_{det}| = 2.5$, et celle du système à muons jusqu'à $|\eta_{det}| = 2$.

Le taux d'acceptation au niveau 1 est de 2.5 MHz et est réduit à environ 2 kHz en sortie. Son temps de décision est de 4.2 μs.

2.3.2 Le niveau 2

Ce niveau de déclenchement mi-matériel mi-logiciel est constitué de deux étapes successives. Tout d'abord, les informations du niveau 1 sont traitées par des pré-processeurs spécifiques à chaque sous-détecteur, afin de reconstruire des traces ou des amas de cellules électromagnétiques. Le SMT est également pris en compte (cf. figure 2.26), ses informations sont combinées à partir du L2 à celles du CFT dans un système de déclenchement sur les vertex déplacés, appelé L2STT, qui sera utilisé pour sélectionner des événements comportant dans leur état final des particules à longue durée de vie. Les pré-processeurs traitent les informations en parallèle en 50 μs.

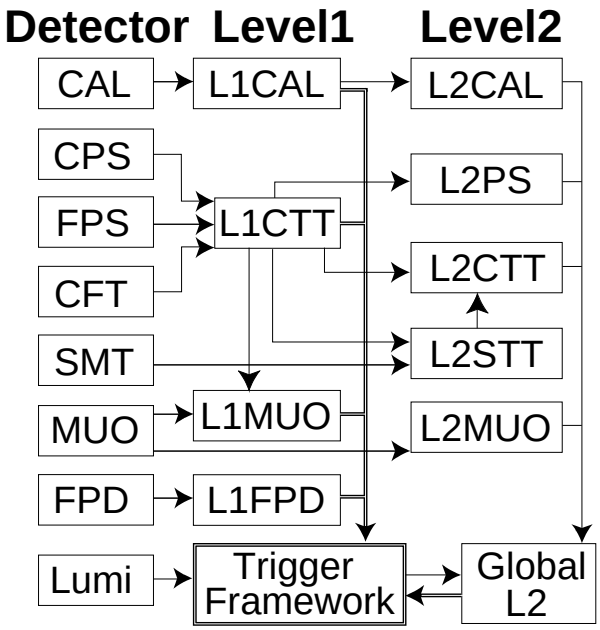

FIG. 2.26: Schéma des deux premiers niveaux du système de déclenchement.

Puis un processeur global prend en charge toutes ces informations pour les combiner et reconstruire des objets physiques bien plus élaborés, tels qu'électrons, photons, jets, vertex, muons. A partir de ces objets physiques, il calcule d'autres quantités telles que des masses ou des séparations angulaires, et teste des correlations dans les signatures physiques entre les divers sous-détecteurs. Il dispose également de 50 μs. Les algorithmes utilisés ont la possibilité d'être modifiés afin de s'adapter à la luminosité.

Le niveau 2 réduit le taux d'acceptation à environ 800 Hz, et a un temps de décision total de 100 μs.

2.3.3 Le niveau 3

Ce dernier niveau de déclenchement est purement logiciel et ne prend pas en compte l'information des niveaux précédents. Il utilise une ferme de processeurs pour reconstruire totalement les événements, en mettant en jeu des algorithmes de reconstruction complexes et proches de ceux utilisés lors de la reconstruction hors-ligne. Divers filtres sont alors appliqués sur ces événements reconstruits. Les critères de sélection sont en évolution constante, et de plus en plus contraignants au fur et a mesure que la luminosité augmente. Les versions de listes de *triggers* utilisées changent donc au cours du temps. Les critères L3 choisis pour la mesure de l'hélicité du boson W seront indiqués dans le chapitre 4.

Le niveau 3 réduit le taux d'acceptation à moins de 50 Hz et a un temps de décision total de 100 μs. En effet, au-delà de cette fréquence, tous les événements ne pourraient pas être traités lors de la reconstruction hors-ligne. Les événements passant ce dernier niveau sont ensuite enregistrés sur bandes magnétiques afin d'être reconstruits entièrement hors-ligne, puis analysés.

2.4 Reconstruction des données

Le programme de reconstruction de DØ est basé sur une bibliothèque de classes et de *templates* en langage C++ appelée *Event Data Model* (EMD). L'élément central est l'événement, i.e. le résultat d'un croisement de faisceau, qui contient toutes les informations basiques de chacun des détecteurs concernant cette interaction, ainsi que l'information liée au déclenchement. Lorsque l'événement a passé les critères de déclenchement, ces informations sont enregistrées sur bande magnétique.
La reconstruction hors-ligne est réalisée par le programme *d0reco*, qui permet également de traiter les événements simulés. Les fichiers résultants sont enregistrés dans le système de gestion et de manipulation des données SAM[6], dans le format dit *ThumbNail* [47] (d'autres formats existent, conservant une quantité d'information plus importante, parfois redondante).
Ensuite le *Common Samples Group* (CSG) se charge du retraitement éventuel des données au moyen du programme *d0correct*, qui applique toutes les corrections nécessaires aux objets, et de fournir des sous-lots de données (*skim*) ayant passé certaines présélections. Les lots (enregistrés dans SAM également) peuvent être au format *ThumbNail* mais également dans un nouveau format développé pour les analyses p17 (cf. ci-dessous) et ultérieures, nommé *Common Analysis Format* (CAF) et regroupant des ROOT *trees*. Le programme *tmb_analyze* permet de réaliser cela. Les fichiers correspondants, nommés par analogie CAFTREES, sont les fichiers de base utilisés par toutes les analyses de physiques de la collaboration menées depuis la version p17 des algorithmes de reconstruction. L'architecture d'analyse permettant de traiter de façon commune et collaborative dans DØ ces fichers est l'environnement CAFE (*Common Analysis Format Environment*).

Des versions différentes des algorithmes se succèdent au cours du temps, en fonction de l'évolution de la prise de données et de la compréhension du détecteur et des données. A chaque itération, toutes les données doivent être de nouveau reconstruites et les simulations reproduites en conséquence. Ainsi, les analyses dites p14 exploitent jusqu'à environ 370 pb^{-1} de données, enregistrées d'avril 2002 à août 2004 et reconstruites avec la version p14 des algorithmes de reconstruction. Les analyses p17 couvrent tout le *Run IIa* soit plus d'1 fb^{-1} de données, et les analyses p20 utilisent toutes les données du *Run II* reconstruites jusqu'à ce jour (les données enregistrées avant 2006 ont été reprocessées en p20).

2.5 Simulation des événements

Les simulations sont indispensables à l'analyse des données réelles enregistrées par un détecteur. Reposant sur divers modèles théoriques et paramètres, elles permettent par

[6] *Sequential data Acces via Meta-data.*

comparaison aux données et à condition de bien les décrire, de contraindre ou rejeter lesdits modèles et paramètres. La première étape est la génération d'un processus physique donné à partir du calcul des diagrammes de Feynman et de fonctions de distribution des partons. Il faut ensuite modéliser le comportement et la réponse de chaque sous-détecteur de DØ au passage des particules générées, et enfin reconstruire et traiter ces événements comme s'il s'était agit de données réelles.

2.5.1 Générateurs

L'interaction dure entre les partons du proton et de l'anti-proton est traitée par QCD perturbative. Ce ne sont pas les protons et les antiprotons qui interagissent, mais les partons les constituant. L'impulsion des partons interagissant, fraction de l'impulsion du proton et de l'antiproton, est dérivée par le programme CTEQ6L1 à l'aide des distributions de densité de partons (PDF). Ensuite, pour le calcul des éléments de matrice de l'interaction, des radiations dans l'état initial et l'état final et de l'hadronisation, les générateurs PYTHIA 6.323 [48] et ALPGEN 2.05 [49] sont utilisés.

PYTHIA se charge de l'hadronisation des quarks et des gluons colorés en jets de hadrons. Celle-ci se décompose en deux phases, la fragmentation et la désintégration des particules instables ainsi formées. PYTHIA gère également le développement de la gerbe de partons qui permet de modéliser les radiations dans l'état initial et l'état final.

ALPGEN lui, calcule les interactions dures et les quadri-vecteurs résultants à partir de l'élément de matrice au premier ordre du processus physique simulé.

2.5.2 Simulation du détecteur et des interactions au Tevatron

Une fois les processus physiques générés, il est nécessaire de simuler la réponse de l'ensemble du détecteur à l'événement. L'interaction des diverses particules avec chaque sous-détecteur et la réponse de ces derniers sont déterminées grâce au programme *d0gstar* [50]. Celui-ci repose sur une simulation GEANT4 [51] de la géométrie du détecteur. L'étape suivante est la numérisation de ces simulations par le programme *d0sim*. Celui-ci comporte deux aspects supplémentaires : il simule diverses sources de bruit instrumental, et modélise des interactions multiples pouvant apparaître dans les données réelles lors d'un croisement de faisceau (celles-ci représentent l'*underlying event*, ou événement sous-jacent). Ces interactions supplémentaires ont deux sources. L'une, indépendante de la luminosité, correspond aux interactions des résidus des protons et antiprotons initiaux de l'interaction. L'autre correspond à l'empilement d'interactions dures provenant des croisements de faisceaux prédédents (*pile-up*). Pour prendre en compte l'*underlying event*, comme la simulation de ces interactions multiples se révèle en mauvais accord avec les données, un type spécifique d'événements (appelés événements de type *zero-bias*),

issus des données réelles et vérifiant des conditions de déclenchement particulières, est utilisé. Il est traité par le programme *d0raw2sim* puis superposé aux événements simulés.
Ces événements simulés sont alors transmis au programme de correction *d0reco*, puis suivent le même traitement que les données réelles.

2.6 Reconstruction et identification hors ligne des objets

Un ensemble d'algorithmes codés en langage C++ et dénommé *d0reco* traite les données brutes enregistrées par les divers sous-détecteurs afin de reconstruire les traces puis les vertex primaires et secondaires d'interaction, et afin d'identifier notamment des candidats électrons, muons et jets. Il permet aussi de mesurer l'énergie, l'impulsion et la charge des objets. Le programme *d0reco* applique également des corrections déduites de la calibration des détecteurs.

2.6.1 Les traces

Le trajectographe, constitués des détecteurs SMT et CFT, est utilisé pour reconstruire les traces, qui sont par définition un ensemble de coups enregistrés par ces détecteurs et formant la trajectoire d'une particule chargée. Les traces obtenues seront ensuite extrapolées en dehors du trajectographe au moyen d'un modèle du détecteur, en prenant en considération les dispersions multiples, la perte d'énergie et les effets dûs au champ magnétique. Lorsqu'elles atteignent les autres sous-détecteurs, ces traces peuvent être associées à des coups dans ces systèmes [52].

Les techniques classiques de reconstruction de traces ne sont pas adaptées aux événements enregistrés dans DØ, qui comptent entre 10^4 et 10^6 coups enregistrés dans le trajectographe, avec une trace pouvant compter jusqu'à 21 coups[7]. Les algorithmes de reconstruction se doivent d'être efficaces et rapides. Deux algorithmes, dont les performances diffèrent selon les types d'objets associés aux traces, sont utilisés. Le premier s'appelle le *Global Track Finder* [52], il est associé au *Histogramming Track Finding* [53], qui après un ajustement de Kalman [54] sur l'information des coups dans le trajectographe, divise le détecteur en régions en $(qB/P_T,\phi)$ et utilise une transformation de Hough afin de réduire le nombre initial de combinaisons. Le second s'appelle l'*Alternative Algorithm* [55], il est basé sur les coups SMT à 1 dimension. Les lots de candidats traces produits par chaque algorithme sont finalement combinés et recalculées par un algo-

[7] Pour le CFT : 16 coups pour la région centrale $|\eta_{det}| < 1.7$, de 8 à 15 coups pour $1.7 < |\eta_{det}| < 2.0$, 0 coups pour $|\eta_{det}| > 2.0$. Pour le SMT : 4 coups plus 1 coup dans la *Layer 0* pour les données les plus récentes

rithme de Kalman, qui fournira des paramètres de qualité concernant les traces (en particulier le χ^2 de l'ajustement). Seules les traces de $p_t > 183$ MeV sont considérées.

2.6.2 Le vertex primaire

La principale tâche de la recherche du vertex primaire consiste à identifier les traces provenant de la collision primaire et à les distinguer des traces produites par les événements de biais minimum[8] qui s'y superposent. Pour une luminosité instantanée $\mathcal{L} \approx 10^{32}$ cm^{-2}s^{-1} par exemple, 2.5 interactions de biais minimum additionnelles sont attendues. Le vertex primaire étant ensuite utilisé par *d0reco* pour reconstruire les jets, les électrons et l'énergie transverse manquante, il est crucial d'optimiser l'algorithme de reconstruction et de sélection de ce vertex.

Deux algorithmes de reconstruction et sélection du vertex primaire, à la philosophie bien différente, sont utilisés pour la version p17 de la reconstruction des données : l'algorithme 2PPV et l'algorithme APV. Les deux types de vertex seront utilisés et étudiés lors de la calibration p17 de l'algorithme d'identification des jets beaux JLIP, présentée à la section 3.1.
Quel que soit l'algorithme, la reconstruction des vertex primaires se décompose en trois étapes :

- sélection des traces formant un vertex (en fonction de leur impulsion transverse P_T, du nombre de coups SMT et de la significance en paramètre d'impact[9]) ;
- ajustement des vertex à partir des traces, afin d'estimer leur position précise et d'y calculer le moment des traces associées ;
- sélection du vertex primaire.

Les deux paragraphes suivant décrivent succintement le principe de ces algorithmes de reconstruction des vertex. En pratique, depuis la version p17 de la reconstruction des données, l'algorithme utilisé est l'APV, car il permet de mieux séparer les vertex primaires et secondaires. L'efficacité de reconstruction des vertex primaires obtenue dans les données est supérieure à 97% pour des vertex de coordonnée le long du faisceau $|z| < 100$ cm. La résolution déterminée dans des événements multijets est de 10 μm dans le plan transverse et 25 μm selon z.

[8] Evenements de biais minimum ou *minimum bias* : événements qui déclenchent les luminomètres mais ne valident aucun critère de déclenchement. Il correspondent donc à une collision inélastique enregistrée par le détecteur, et sont à dissocier des événements de biais nul ou *zero bias* qui eux correspondent à un croisement de faisceau sans vertex primaire reconstruit ni déclenchement des luminomètres.

[9] Le Paramètre d'Impact (IP ou encore DCA, *Distance of Closest Approach*) correspond à la distance minimale d'approche entre la trajectoire et le vertex primaire, cf. section 3.1.3

2.6.2.1 L'algorithme *2-Pass Primary Vertex*

Cet algorithme (noté 2PPV), décrit dans la référence [56] et utilisé sous une forme un peu différente pour la reconstruction p14, est basé sur une approche en deux itérations (*2-pass*), qui permet en particulier d'augmenter la multiplicité de traces du vertex. Dans chaque événement, un premier lot de vertex primaires est obtenu en utilisant des critères de sélection très lâches sur les traces. Ce lot de vertex sert alors de point de départ pour l'algorithme qui redétermine un nouveau lot de vertex en imposant cette fois-ci des coupures plus contraignantes sur les traces. Les vertex ainsi trouvés subissent un ajustement de Kalman, puis une sélection basée sur une approche probabiliste est effectuée dans cette liste de vertex. En effet, les traces provenant des interactions de biais minimum ayant une impulsion transverse plus faible que celles provenant de l'interaction dure, il est possible d'en déduire une probabilité que le vertex provienne d'une interaction de biais minimum. Le vertex de la liste ayant la plus faible probabilité sera alors choisi comme vertex primaire.

2.6.2.2 L'algorithme *Adaptive Primary Vertex*

La différence entre cet algorithme (noté APV) [57, 58] et le précédent provient essentiellement de la méthode d'ajustement du vertex. Cet algorithme n'utilise pas un simple ajustement de Kalman, mais une technique permettant de réduire la contribution des traces distantes lors de l'ajustement du vertex, qui prend en pratique la forme d'un filtre de Kalman itératif et pondéré. Ceci est relativement important car ces traces, qui proviennent souvent de désintégrations de saveurs lourdes (donc de vertex secondaires), sont parfois associées par erreur au vertex primaire avec un simple filtre de Kalman.
Une autre modification importante par rapport au 2PPV concerne la région du SMT prise en compte : le nombre de coups exigés dans le SMT (qui améliore la résolution du vertex primaire) n'est effective que dans sa région fiducielle ($|z| < 36$ cm), ce qui permet d'avoir une bien meilleure efficacité de reconstruction des vertex à haut z.
La selection des traces diffère quelque peu de celle de l'algorithme précédent, mais la technique de la double itération est elle aussi employée. La sélection du vertex primaire final parmi une première liste de vertex repose quant à elle sur le même principe que pour l'algorithme 2PPV.

2.6.3 Les jets

Un jet n'est pas réellement un objet physique mais correspond à un flot de particules globalement collimées, issues de la fragmentation d'un parton acteur de l'interaction dure. Un tel jet est représenté schématiquement sur la figure 2.27. Les jets sont reconstruits dans l'expérience DØ en utilisant l'information enregistrée par le calorimètre.

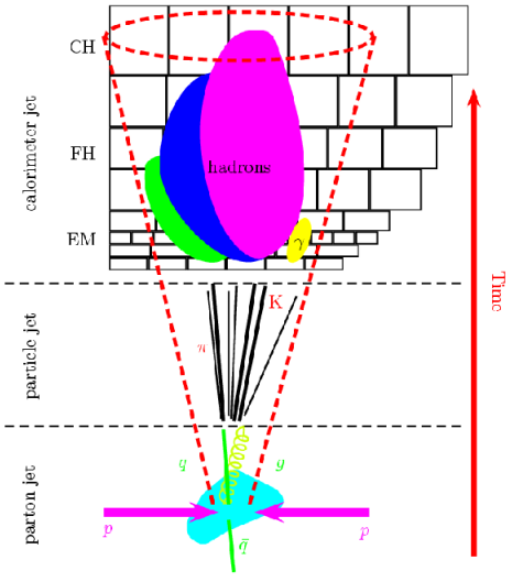

FIG. 2.27: Schéma de l'évolution de l'interaction dure partonique au jet calorimétrique. On note les trois définitions d'un jet : jet partonique, jet de particule ou jet calorimétrique.

2.6.3.1 Reconstruction

Les jets sont reconstruits depuis le *Run II* au moyen d'algorithmes dits de cône. La DØ Note [59] décrit en détail ces algorithmes et leur mise en application dans l'expérience DØ. L'idée sous-jacente est d'associer des objets, qui ont leur origine au point d'interaction, dans des cônes de rayon angulaire donné. Un jet correspondra alors à une configuration stable, où l'axe du cône qui définit les objets appartenant au jet coïncide avec la direction du jet déduite des objets qui le composent.

La distance angulaire entre deux objets est définie par $\Delta R = \sqrt{\Delta y^2 + \Delta\phi^2}$, où y correspond à la rapidité et ϕ à l'angle azimutal. Toute l'information des quadri-vecteurs des objets est utilisée pour les combiner. La reconstruction se déroule en deux étapes distinctes. L'algorithme de *Simple Cone* forme tout d'abord des pré-amas d'au moins deux tours calorimétriques ayant une énergie totale de plus de 1 GeV/c et comprises dans un rayon $\Delta R < 0.3$.
Ces pré-amas, ordonnés en énergie, servent de base à un second algorithme (l'*Improved Legacy Cone Algorithm*) pour former des proto-jets à partir des tours du calorimètre. Le cône des proto-jets doit avoir cette fois-ci un rayon $\Delta R < 0.5$ (algorithme dit JCCA) ou $\Delta R < 0.7$ (algorithme dit JCCB, utilisé par les analyses en physique du top dans DØ et en particulier pour la mesure de l'hélicité du W présentée dans cette thèse). Une seconde itération est réalisée sur l'ensemble des proto-jets obtenus et permet par ailleurs

de limiter la sensibilité de l'algorithme aux radiations. Enfin, les proto-jets ayant passé les deux premières itérations sont fusionnés ou séparés selon leur éventuel recouvrement, et rejetés si leur énergie transverse est inférieure à 6 GeV/c. Nous obtenons ainsi la liste finale des jets.

2.6.3.2 Identification

Une fois des jets reconstruits, il est nécessaire de statuer quant à leur qualité. Ainsi, trois critères sont définis pour déclarer qu'un jet est un bon jet ou non, i.e. sélectionner les jets physiques et rejeter les jets de bruit. L'étude visant à optimiser l'identification des jets p17 est présentée dans la DØ Note [60]. Les critères sont appliqués sur les variables suivantes :

- **la fraction d'énergie transverse E_T déposée dans le calorimètre hadronique grossier CH :** variable CHF pour *Coarse Hadronic Fraction*. Requérir que cette variable ne dépasse pas une certaine valeur permet de rejeter les jets correspondant à un bruit typique de ce calorimètre. La coupure sur CHF varie en fonction de la pseudo-rapidité du jet afin d'avoir une efficacité d'identification uniforme.
- **la fraction d'énergie transverse E_T déposée dans le calorimètre électromagnétique :** variable EMF pour *ElectroMagnetic Fraction*. La coupure inférieure sur EMF varie en fonction de la pseudo-rapidité du jet et de sa largeur en η, et permet de rejeter des jets dominés par du bruit hadronique. Une coupure maximale EMF $= 0.95$ est aussi utilisée indicativement pour discriminer les jets des électrons et des photons, sans appliquer de réjection.
- **la confirmation du niveau 1 du système de déclenchement :** variable L1Conf. Elle correspond au rapport de l'impulsion transverse (occultée des couches du CH) lue par l'électronique de précision sur l'impulsion transverse correspondant approximativement à celle lue par l'électronique du niveau L1. Une coupure supérieure à un certain seuil permet de rejeter les jets simulés par du bruit de lecture dans la chaîne de lecture de précision. Le choix de ce seuil dépend de l'impulsion transverse du jet et de sa pseudo-rapidité.

2.6.3.3 Correction

La correction appliquée sur l'énergie d'un jet mesurée dans le calorimètre permet de se ramener à l'énergie du jet de particule avant son interaction avec le détecteur (et non pas à l'énergie du jet partonique, cette correction ne traitant pas les radiations dures de gluons faisant de grands angles avec le jet et ayant lieu durant l'évolution de la gerbe partonique). C'est cette énergie qui est utilisée dans la majorité des analyses, en particulier pour la mesure de l'hélicité du W. Cette correction est communément nommée *Jet Energy Scale*, ou JES [61, 62].

La procédure de correction consiste en l'application séquentielle de plusieurs sous-corrections, comme le montre l'équation 2.6. Elles sont au nombre de quatre, et déterminées séparément pour les données et la simulation.

$$E_{jet}^{part} = \frac{E_{jet}^{mes} - O}{F_\eta \times R \times S} \tag{2.6}$$

où, dans l'ordre d'application :

- E_{jet}^{mes} est l'énergie calorimétrique du jet mesurée, non corrigée.
- O est la correction d'*offset*. L'énergie sans rapport avec l'interaction dure (due au bruit électronique, à l'énergie résiduelle des collisions précédentes (*pile-up*) ou aux *underlying events*[10] par exemple) est soustraite à l'énergie mesurée.
- F_η est la correction de réponse relative (ou "inter-calibration η"). La réponse du calorimètre n'est pas uniforme en pseudo-rapidité, de par sa géométrie. Cette correction a pour objectif de rendre sa réponse uniforme sur toute sa couverture en pseudo-rapidité.
- R est la correction de réponse absolue. Une fois la réponse du calorimètre uniformisée géométriquement, il est nécessaire de corriger l'énergie d'effets tels que les pertes induites par les régions non instrumentées du calorimètre, ou encore la réponse plus faible du calorimètre aux hadrons qu'aux électrons ou aux photons.
- S est la correction de gerbe (*showering*). Son objectif est de corriger les variations d'énergie dues aux effets instrumentaux, tels que notamment le développement de la gerbe dans le calorimètre et la courbure due au champs magnétique.

Notons que cette correction est appliquée de manière inclusive quelle que soit la saveur des jets. Or pour bien faire, un jet de quark beau ou de gluon par exemple nécessiterait un traitement différent qu'un jet d'une autre saveur.

2.6.3.4 Correction spécifique à la simulation : JSSR

La correction explicitée précédemment est commune aux données et à la simulation, tout en étant paramétrisée séparément pour chaque cas. Néanmoins, l'accord après correction entre les données réelle et le Monte-Carlo n'est pas parfait, la résolution et l'efficacité de reconstruction se révélant meilleures dans la simulation. Une correction supplémentaire est donc nécessaire pour les jets simulés afin de prendre en compte les différences de résolution de l'énergie des jets simulés, les différences d'échelle d'énergie, et les différences de reconstruction et d'identification des jets. Dénommée JSSR pour *Jets Smearing, Shifting and Removing*, elle se décompose en trois sous-corrections corrélées [63, 64]. La première, dite *smearing*, dégrade l'impulsion des jets simulés. La seconde, le *shifting*, décale les impulsions de ces jets. Cette méthode est délicate à appliquer en particulier pour propager le décalage à l'énergie transverse manquante, et n'est pas uti-

[10] L'*underlying event* correspond au bruit résultant des interactions multiples des partons et des résidus des faisceaux.

lisée par toutes les analyses[11]. En particulier, le *shifting* ne sera pas appliqué pour la mesure de l'hélicité du boson W. Enfin, le *removing* rejette les jets d'impulsion transverse $P_T < 13$ GeV/c (cette coupure permet d'avoir une efficacité de reconstruction et d'identification constante en fonction de l'impulsion transverse corrigée du jet).

2.6.4 Les objets électromagnétiques

2.6.4.1 Reconstruction des objets électromagnétiques

La reconstruction des objets électromagnétiques se base sur l'information provenant des tours du calorimètre électromagnétique et de la couche FH1 de la partie hadronique. L'information est traitée par l'algorithme de *Simple cône* (le même que celui utilisé lors de la reconstruction des jets, présenté dans la section 2.6.3.1) avec un rayon $\Delta R < 0.4$. Le pré-amas le plus énergétique obtenu sert alors à définir un cône de rayon 0.2 dans la partie centrale du calorimètre ou de 10 cm dans les bouchons. Toutes les tours appartenant à ce cône forment ensuite le candidat objet électromagnétique. Cette procédure est efficace pour des électrons isolés de haute impulsion transverse. Lorsque ceux-ci possèdent une faible impulsion transverse ou bien sont présents dans un jet, la *Méthode de la Route* [65] est préférée à cet algorithme. L'énergie des objets électromagnétiques est ensuite corrigée pour prendre en considération la géométrie du détecteur et la matière traversée. Ces objets électromagnétiques peuvent correspondre à un électron, à un photon ou encore à un pion neutre décroissant en deux photons.

2.6.4.2 Identification et discrimination

Un certain nombre de variables permettent de discriminer les électrons des photons et des autres jets, et de définir des critères de qualité. En voici une description succinte.

• **La probabilité sur le χ^2 de l'association de l'objet électromagnétique à une trace**, qui se base sur l'expression suivante :

$$\chi^2 = \left(\frac{\Delta z}{\sigma_z}\right)^2 + \left(\frac{\Delta \Phi}{\sigma_\Phi}\right)^2$$

où Δz et $\Delta\Phi$ correspondent aux écarts selon z et Φ entre l'objet électromagnétique et la trace. La probabilité que l'association entre le jet et une trace donnée ait un certain χ^2 en est ensuite dérivée. Les électrons sont sensés correspondre à une trace dans le trajectographe, et celle-ci doit être alignée avec l'amas calorimétrique. Cela permet de rejeter les particules neutres, celles-ci déposant de l'énergie dans le calorimètre mais pas de trace dans le trajectographe. De plus, en cas de mauvaise association d'une particule

[11] Les erreurs systématiques associées à cette correction sont alors plus importante que lors de l'application de toute la correction JSSR.

neutre à une trace, la qualité de l'association permettra de rejeter cette particule.

• **La *Matrice H* :** permet de mesurer la compatibilité entre la forme du dépôt d'énergie dans le calorimètre à la gerbe électromagnétique théorique. Elle correspond au χ^2 associé à l'inverse d'une matrice de covariance construite à partir de 7 variables[12]. Ces variables sont les fractions d'énergie dans les couches EM1 à EM4 du calorimètre, la largeur de la gerbe dans le plan $R-\phi$ dans la couche EM3, le logarithme de l'énergie et la position longitudinale du vertex primaire divisée par son erreur.

• **La fraction électromagnétique EMF :** rapport de l'énergie déposée dans les couches électromagnétiques EM1 à EM4 et dans la première couche hadronique fine FH1 sur l'énergie totale déposée par la gerbe. Un électron ou un photon aura une grande fraction électromagnétique car les jets ont une gerbe plus grande et qui débute plus tard. Cette variable permet ainsi de bien discriminer ces objets.

• **L'isolation *Iso***, définie comme suit :

$$Iso = \frac{E_{tot}(\Delta R < 0.4) - E_{EM}(\Delta R < 0.2)}{E_{EM}(\Delta R < 0.2)}$$

où E_{tot} correspond à l'énergie totale des tours comprises dans un cône de rayon 0.4 autour de l'objet électromagnétique, et E_{EM} est l'énergie électromagnétique des tours appartenant à un cône de 0.2 autour de l'objet. Une valeur faible de cette variable d'isolation est représentative d'un objet isolé.

• **La vraisemblance électromagnétique L_{EM} :** une fonction de vraisemblance est élaborée à partir de sept variables. Celles-ci sont la fraction électromagnétique, le χ^2 de la Matrice H, le χ^2 de l'association spatiale à une trace, le nombre de traces dans un cône de 0.5 autour de la trace associée, le paramètre d'impact par rapport au vertex primaire, le rapport de l'énergie transverse de l'amas calorimétrique sur l'impulsion transverse de la trace, et enfin la somme des impulsions transverses des traces appartenant à un cône de 0.4 centré sur la trace associée à l'objet électromagnétique.

2.6.4.3 Définition des électrons *tight*

Un électron est considéré de qualité *tight* s'il vérifie les propriétés suivantes :
- une trace est associée à l'objet calorimétrique (avec le critère de χ^2 de la section 2.6.4.2) ;
- $H_{Matrix}7 < 50$;
- $EMF > 0.9$;

[12] Certaines versions de l'identification des objets électromagnétiques utilisent une huitième variable correspondant à la largeur longitudinale de la gerbe dans la couche EM3 du calorimètre.

- $Iso < 0.15$;
- $L_{EM} > 0.85$;
- $P_T^{trace} > 5$ GeV/c.

les variables ayant été définies précédemment. C'est le type d'électron sélectionné dans les analyses utilisant un état final électron+jets telles que la mesure de l'hélicité du boson W.

2.6.5 Muons

2.6.5.1 Reconstruction des muons

Les muons sont reconstruits en utilisant les informations du spectromètre à muons, mais également du détecteur de traces et du calorimètre. La reconstruction des muons dans le spectromètre central a été amplement traitée dans la section **??**, s'y référer pour plus de détails. Un muon identifié par le spectromètre à muons est appelé un muon local. Les chambres à muons fournissent une mesure de son impulsion. Le trajectographe permet ensuite d'améliorer la mesure de l'impulsion. Lors qu'un muon local est associé avec une trace enregistrée dans le trajectographe, le muon est dit global.

2.6.5.2 Identification des muons

L'identification des muons repose sur des critères de qualité de l'objet, des critères de qualité de la trace qui lui est associée dans le trajectographe et des critères d'isolation. Les références [66, 67] présentent tous ces critères et leur efficacité, tandis que les paragraphes ci-dessous décrivent rapidement les spécificités de ces critères et plus particulièrement de ceux qui interviendront pour la mesure de l'hélicité du boson W.

• **Critères de qualité du muon :**

Les muons reconstruits sont classés selon leur type et leur qualité. Le type du muon est donné par le paramètre *nseg* qui peut prendre sept valeurs (0, ±1, ±2, ±3). En particulier $|nseg| = 3$ correspond à un muon local reconstruit à l'aide des coups enregistrés dans les trois couches A, B et C du spectromètre à muons et associé ($nseg = 3$) ou non ($nseg = -3$) à une trace centrale. La qualité d'un muon peut être *loose*, *medium* ou *tight*. Celle-ci dépend de la valeur de la variable *nseg* ainsi que des impacts enregistrés dans le spectromètre, et éventuellement de la convergence de l'ajustement de la trace du muon local.

Un muon sera de qualité *medium* s'il remplit les conditions suivantes :

- $|nseg| = 3$;
- au moins deux coups sont enregistrés dans les chambres à fils de la couche A du spectromètre ;
- au moins un coup dans les scintillateurs de la couche A ;
- au moins deux coups dans les chambres à fil des couches BC ;

– au moins un coup dans les couches de scintillateurs BC, hormis pour les muons globaux comportant moins de quatre coups dans les chambres à fils des couches BC.

• **Critères de qualité de la trace associée au muon :**

Ces trois critères permettent de contrôler la qualité de la trace associée au muon, ainsi que la qualité de cette association. Ce sont :

– le nombre de coups dans le SMT et le CFT ;
– le χ^2 par degré de liberté de l'ajustement de la trace (χ^2/dof) ;
– son paramètre d'impact dans le plan transverse par rapport au vertex primaire $(|dca|)$.

Une trace est de qualité *loose* si $|dca| < 0.2$ cm dans le cas où elle n'a pas de coups dans le SMT ou $|dca| < 0.02$ cm dans le cas contraire. Une trace *medium* est une trace *loose* telle que $\chi^2/dof < 4$. Enfin, une trace *tight* est une trace *medium* ayant au moins un coup dans le SMT.

• **Critères d'isolation du muon :**

Cinq variables sont utilisées pour décrire l'isolation d'un muon et permettre de définir la qualité de cette isolation :

– **TrackHalo** $= |\sum_{traces} P_T|$
dans le cône $\Delta R(trk, trk_\mu) < 0.5$;
– **ScaledTrackHalo** $= |\sum_{traces} P_T/P_T(\mu)|$
dans le cône $\Delta R(trk, trk_\mu) < 0.5$;
– **CalorimeterHalo** $= |\sum_{cellules} E_T|$
dans le cône $\Delta R(cell^{calo}, trk_\mu^{calo}) < 0.4$;
– **ScaledCalorimeterHalo** $= |\sum_{cellules} E_T/P_T(\mu)|$
dans l'anneau $0.1 < \Delta R(cell^{calo}, trk_\mu^{calo}) < 0.4$;
– $\Delta R(\mu, jet)$: distance au jet le plus proche selon $R - \phi$.

L'isolation du muon de qualité *TopP14Tight* (utilisée dans l'analyse faisant l'objet de cette thèse) contraint trois de ces variables aux valeurs suivantes :

– $ScaledTrackHalo < 0.06$;
– $ScaledCalorimeterHalo < 0.05$;
– $\Delta R(\mu, jet) > 0.5$.

• **Veto sur les muons cosmiques :**

Afin de rejeter les signaux des rayons cosmiques, l'information sur le temps des coups dans les couches de scintillateurs est utilisée. Pour chacune des couches A, B et C, si l'information sur le temps est disponible et que ce temps est pour toutes les couches inférieur à 10 ns, alors le muon correspondant est rejeté.

2.6.5.3 Définition des muons *tight*

Lors de l'analyse visant à mesurer l'hélicité du boson W qui est présentée aux chapitres 4 et 5, des muons isolés de qualité *tight* sont recherchés pour sélectionner des événements correspondant à l'état final muon+jets d'une paire de quarks top. La définition d'un tel muon est la suivante :

- le muon doit être de qualité *medium* avec $nseg = 3$;
- la qualité de la trace associée doit être *medium* ;
- l'isolation doit être de qualité *TopP14Tight* ;
- le muon ne doit pas être un muon cosmique.

De plus, ces événements ne doivent pas comporter d'autres muons passant les mêmes critères hormis l'isolation, pour laquelle une seule contrainte est maintenue : $\Delta R(\mu, jet) > 0.5$.

Dans le canal électrons+jets, aucun muon *tight* aux critères d'isolation relachés (conservant uniquement $\Delta R(\mu, jet) > 0.5$) ne doit être présent.

2.6.6 L'énergie transverse manquante

C'est une donnée dont la connaissance est primordiale quelle que soit l'analyse de physique menée. En effet, l'énergie transverse manquante $\not\!\!E_T$ est la signature des neutrinos et des autres éventuelles particules n'interagissant pas avec le détecteur [68]. Sa détermination combine des quantités mesurées dans plusieurs des sous-détecteurs : le calorimètre (dépôt d'énergie), le trajectographe (pour la détermination du vertex primaire) et le spectromètre à muons (pour la correction de l'énergie de l'événement par l'énergie des muons, car ceux-ci n'ont pas déposé leur énergie dans le calorimètre). Cependant, une mesure précise de la $\not\!\!E_T$ nécessite de lui appliquer diverses corrections dues aux autres objets reconstruits (jets, muons, etc.), ce qui lui impose d'être déterminée en fin de la chaîne de reconstruction. Elle est calculée en plusieurs étapes [69]. Dans un premier temps, la décomposition dans le plan transverse de l'énergie visible est définie :

$$E_{x,y}^{vis} = \sum_{cellules} E_{x,y}^{i} \tag{2.7}$$

où la somme porte sur les cellules du calorimètre électromagnétique et hadronique fin, du détecteur inter-cryostat, et du *massless gap* (cf. section 2.2.3.3). La prise en compte des dépôts d'énergie dans le calorimètre hadronique grossier CH, trop bruyant, induirait bien trop de fausse $\not\!\!E_T$, mais une correction due aux dépôts de cette région sera appliquée ultérieurement. On définit alors la projection dans le plan transverse de l'énergie manquante comme suit :

$$\not\!\!E_{T\,x} = -E_x^{vis} \tag{2.8}$$

$$\not\!\!E_{T\,y} = -E_y^{vis} \tag{2.9}$$

$$\not\!\!E_T = \sqrt{(\not\!\!E_{Tx})^2 + (\not\!\!E_{Ty})^2} \tag{2.10}$$

Les jets, les objects électromagnétiques et les muons présents dans l'événement ayant vu leur impulsion modifiée par le programme *d0correct*, ces changements sont ensuite propagés dans le calcul de la $\not\!\!E_T$ [70].

3

L'identification des jets de quarks beaux

Nous avons vu précédemment dans le chapitre 1 que le quark top se désintègre par interaction faible chargée en un quark beau, qui s'hadronise en un jet par la suite. Un algorithme d'identification des jets de quarks beaux permet d'opérer une sélection des candidats événements top.

L'intérêt d'une telle identification est cependant bien plus général que sa seule application aux études du top. L'identification des jets de quarks beaux est notamment nécessaire pour la recherche du boson de Higgs se désintégrant en paires $b\bar{b}$ et pour certaines recherches de particules supersymétriques.

Plusieurs propriétés caractérisent les désintégrations des particules belles par rapport à celles des particules produites à partir de quarks plus légers. Tout d'abord, le temps de vie des particules belles, qui se désintègrent par interaction faible chargée avec généralement des durées de vie de l'ordre de 1.5 ps (soit une distance de vol $c\tau \simeq 400\ \mu m$), les différencie des autres particules. La plupart en effet soit sont stables ou du moins stables à l'échelle du détecteur ($\pi^{\pm}, K^{\pm}, p, n, \mu$), soit se désintègrent immédiatement au point de production comme les résonances (Φ et D^* par exemple) ou proche du point de production (D, D_s). Ces temps de vol des particules belles conduisent à l'existence d'un vertex secondaire déplacé correspondant à leur point de désintégration. Par ailleurs, leur masse est bien plus élevée que celle des autres particules. Enfin, un hadron beau emporte en moyenne 70 % de l'énergie du quark b initial, alors que cette fraction n'est que de 50 % pour un hadron charmé produit par l'hadronisation d'un quark c au vertex primaire.

Ces trois propriétés entraînent des distributions différentes pour les variables cinématiques et topologiques caractérisant les jets beaux et les jets de saveurs plus légères : l'impulsion transverse du lepton issu de la désintégration du quark b est relativement dure, il y a présence d'un vertex déplacé, la masse et la multiplicité associées à ce vertex secondaire sont élevées, la rapidité et les paramètres d'impact des traces secondaires sont élevés.

L'équipe DØ de l'IPHC Strasbourg a developpé une méthode d'identification de la beauté d'un jet basée sur les paramètres d'impact des traces appartenant à ce jet. Cette méthode est baptisée JLIP pour *Jet LIfetime Probability*. Elle est largement utilisée par la collaboration DØ, en tant qu'algorithme d'identification autonome jusqu'à la version p17 de la reconstruction des données (février 2006), puis comme variable d'entrée d'un algorithme basé sur un réseau de neurones depuis la version p17 des analyses.
J'expliquerai la méthode, qui a été developpée avant ce travail de thèse, puis donnerai tous les résultats des calibrations p17 et p20[1] que j'ai assurées pour ces deux processings de données. Enfin, je détaillerai le réseau de neurones d'identification de la beauté, basé entre autres sur JLIP et utilisé pour la mesure de l'hélicité faisant l'objet de cette thèse.

3.1 Identification des jets de quarks beaux par l'algorithme JLIP

3.1.1 Lots de données et de simulation utilisés pour la calibration

Tous les lots de données et de simulation utilisés, que ce soit pour l'algorithme p17 ou pour sa version p20, ont été utilisés dans leur format CAF_TREE (cf. paragraphe 2.4).

3.1.1.1 Données et simulations p17

Pour les données, trois lots d'événements sont utilisés : les lots dits QCD, BID et EM1TRK :

- le lot QCD est constitué d'environ 4.9 millions d'événements dont l'acquisition a été déclenchée sur des jets à grand P_T. Ce lot de données sert à la calibration de l'algorithme ;
- le lot BID est composé de 8 millions d'événements comportant un muon de $P_T > 4$ GeV/c associé à un jet. Ainsi enrichi en jets b, ce lot sert aux mesures d'efficacité de l'identification ;

[1] p20 est le processing des données correspondant au *Run IIb* qui a débuté en août 2006 après l'ajout de la couche 0 du trajectographe.

– le lot EM1TRK comporte 5.8 millions d'événements ayant au moins un objet électromagnétique associé à une trace. Il est également utilisé pour les études de performances (taux de mauvaise identification).

Concernant la simulation Monte-Carlo, trois lots d'événements multi-jets QCD sont utilisés. Pour chacun, l'un au moins des jets d'un événement a son impulsion transverse appartenant respectivement à l'intervalle $20 < P_T < 40$ GeV/c , $40 < P_T < 80$ GeV/c et $80 < P_T < 160$ GeV/c. Chaque lot est composé d'environ 200 000 événements. Ces simulations sont utilisées pour l'ensemble de la calibration de l'algorithme. D'autres lots Monte-Carlo sont employés pour les études des performances de JLIP : $t\bar{t} \rightarrow$ lepton+jets, $t\bar{t} \rightarrow$ di-jets, $Z \rightarrow b\bar{b}$ et $Z \rightarrow b\bar{b} \rightarrow \mu$ pour un total de 370 000 événements dont 60 000 avec un muon associé à un jet.

3.1.1.2 Données et simulations p20

Les données sont constituées de 16 millions d'événements QCD dont l'acquisition a été déclenchée sur des jets à grand P_T.

De même que pour la version p17 de l'algorithme, la simulation p20 est constitué de trois lots d'événements multijets QCD dont l'un au moins des jets à son impulsion transverse appartenant respectivement à l'intervalle $20 < P_T < 40$ GeV/c , $40 < P_T < 80$ GeV/c et $80 < P_T < 160$ GeV/c. Ces lots ont été utilisés pour la calibration de l'algorithme. Afin d'augmenter la statistique disponible, trois lots complémentaires $\gamma + jets$, dont l'un au moins des jets appartient aux mêmes gammes en P_T, sont également utilisés. Chaque lot est composé d'environ 200 000 événements.

Notons que dans la simulation, quelle que soit la version du processing la saveur des jets calorimétriques reconstruits est déterminée selon les critères suivants :

– si un hadron beau généré est dans un cône de rayon 0.5 autour de la direction du jet, celui-ci est considéré comme étant un jet b ;
– si un hadron charmé (mais aucun hadron beau) est dans le cône de rayon 0.5 autour de la direction du jet, alors celui-ci est considéré comme étant un jet c ;
– si ni hadron beau, ni hadron charmé n'est trouvé, alors le jet est considéré de saveur légère (jet de quarks u, d, s ou jet de gluons).

3.1.2 Sélection des événements et des objets

3.1.2.1 Critères de déclenchement

Depuis le *Run* 194500 (juin 2004), un biais est susceptible d'être introduit dans les données par les critères de déclenchement exigés (le paragraphe 2.3 donne de plus amples informations sur le déclenchement). En effet les critères de déclenchement ne sont pas définis spécifiquement pour les études d'algorithme d'identification des jets beaux, mais pour optimiser l'acceptance aux processus physiques et éviter la saturation de la prise de données. Il apparaît que certains critères du niveau L3 requièrent des coupures sur le vertex primaire (événements tels que la coordonnée du vertex primaire le long du faisceau vérifie $|z_{PV}| < 35$ cm), sur le paramètre d'impact[2] des traces ($\geq 2-4$ traces de significance $\mathcal{S}_{IP} > 3$), ou encore sur une probabilité d'identification (≥ 1 jet possédant une probabilité inférieure à 0.05 ou 0.3, selon les cas). Ainsi, un événement est biaisé pour toute étude d'identification de jets beaux lorsqu'il valide une combinaison de ces trois critères, et aucun autre critère. Pour l'étude avec la version p20 du processing, de tels événements seront rejetés.

3.1.2.2 Vertex primaire

Les divers algorithmes de reconstruction des vertex primaires utilisés par la collaboration ont été présentés dans la section 2.6.2. L'algorithme 2PPV était le seul utilisé pour la version p14 des analyses alors que pour la version p17 un autre algorithme, APV, a aussi été testé et les performances des deux algorithmes comparées. Les paramétrisations nécessaires à JLIP ont été fournies à la collaboration pour les deux cas, mais l'algorithme de reconstruction finalement retenu a été l'APV. Pour la version p20 par contre, seul ce dernier a été étudié et utilisé. Le type d'algorithme utilisé sera précisé lorsque nous présenterons des résultats portant sur l'identification p17. Les résultats obtenus pour l'identification des jets beaux se révéleront finalement très similaires pour les deux algorithmes de reconstruction étudiés.
Les événements étudiés doivent évidemment comporter un vertex primaire reconstruit afin de pouvoir calculer des variables intéressantes pour l'identification des b. Celui-ci doit de plus avoir trois traces qui lui soient attachées et vérifier $|z_{PV}| < 60$ cm. Toutes les traces sont extrapolées au vertex primaire ainsi reconstruit et leurs paramètre d'impact et erreur recalculés.

3.1.2.3 Muons

Dans certaines études menées avec la version p17 de l'algorithme, des lots enrichis en muons ont été utilisés. La présence d'un muon associé à un jet est en effet révélatrice de la présence de saveur lourde dans l'événement. Son moment transverse est utilisé pour

[2] Les notions de paramètre d'impact d'une trace, significance en paramètre d'impact et probabilité d'identification seront définies par après.

déterminer la proportion de saveurs lourdes et par la suite l'efficacité d'identification. Le muon sélectionné doit vérifier les critères suivants :

- être de qualité *medium* (cf. paragraphe 2.6.5) et reconstruit à partir de 3 segments dans le spectromètre à muons ;
- $|z_{IP}| < 1$ cm ;
- $P_T > 4$ GeV/c ;
- $\chi^2 < 100$.

Dans le cas où plusieurs muons seraient sélectionnés, seul celui de plus petit χ^2 est retenu.

3.1.2.4 Sélection des jets et des traces et notion de *taggabilité*

La comparaison entre simulation et données réelles est rendue difficile par l'existence de différences concernant l'efficacité de reconstruction des traces, ainsi que par les jets mal reconstruits dans le calorimètre, phénomène mal pris en compte dans la simulation. Pour pallier à cela, la notion de *taggabilité* (du terme anglais *taggability*) est introduite, elle correspond au rapport du nombre de jets *taggables* (ou "étiquetables") au nombre de jets reconstruits (i.e. probabilité pour un jet reconstruit d'être étiquetable). Cette probabilité d'être *taggable* est paramétrisée en fonction de l'énergie transverse E_T et de la pseudo-rapidité η du jet :

$$P^{taggabilite}(E_T, \eta) = \frac{\#\text{jets taggables } (E_T, \eta)}{\#\text{jets reconstruits } (E_T, \eta)}$$

Un jet calorimétrique est dit *taggable* si un jet de traces peut lui être associé dans un cône de rayon $\Delta R = 0.5$. Le jet de traces doit remplir les critères suivant :

- être composé d'au moins 2 traces chargées ayant chacune au moins un coup dans le SMT, dont l'une a une impulsion transverse $P_T > 1$ GeV/c et les autres $P_T > 0.5$ GeV/c ;
- avoir un paramètre d'impact inférieur à 2 mm dans le plan transverse et inférieur à 4 mm selon l'axe du faisceau ;
- le jet doit alors pouvoir être construit à partir d'une trace initiale de $P_T > 1$ GeV/c et des traces appartenant au cône de rayon 0.5 centré sur cette trace principale.

Les traces qui seront utilisées par l'algorithme d'identification JLIP sont celles qui vérifient ces contraintes. Seuls les jets calorimétriques d'impulsion transverse $P_T > 15$ GeV/c et de pseudo-rapidité $|\eta| \leq 2.5$ sont pris en considération. De plus, les événements ne possédant pas au moins 2 jets étiquetables seront rejetés. La *taggabilité* est paramétrisée en fonction de la pseudo-rapidité η des jets et de leur énergie transverse E_T. Mais elle est également dépendante de la multiplicité en jets de l'événement, ce qui nécessite de déterminer les paramétrisations en fonction des types d'analyse. Ainsi, l'introduction de cette notion de *taggabilité* permet de décomposer la probabilité d'identifier un jet comme beau en deux parties : la probabilité qu'à le jet d'être *taggable* $P^{taggabilite}$, et la

probabilité qu'un jet *taggable* soit identifié.

$$P^{jet\ b} = \frac{\#\text{jets identifies b}}{\#\text{jets reconstruits}} = \frac{\#\text{jets identifies b}}{\#\text{jets taggables}} \cdot P^{taggabilite}$$

Les paramétrisations de *taggabilité* étant communes à tous les algorithmes d'identification de jets de quark b utilisés par la collaboration[3], son utilisation permet de comparer directement l'efficacité de ces derniers.

La figure 3.1 présente l'efficacité de *taggabilité* en fonction de divers paramètres. Elle a été déterminée dans les données et la simulation QCD pour la version p17 de l'algorithme, dans le cas où le vertex primaire a été reconstruit avec l'algorithme APV. Les résultats sont similaires pour l'algorithme 2PPV. L'efficacité de *taggabilité* est globalement plus élevée dans la simulation d'environ 10 %, ceci pouvant s'expliquer par la meilleure reconstruction des traces dans les jets observée dans la simulation. La figure 3.2 présente les mêmes efficacités pour la version p20 de l'algorithme. L'efficacité est approximativement la même dans les données qu'en p17, par contre l'efficacité dans la simulation est plus faible en p20 et présente un meilleur accord à celle des données, tout en lui restant supérieure de quelques pourcents.

On définit alors l'efficacité d'identification ou de mauvaise identification par un algorithme d'identification donné comme étant son efficacité rapportée au nombre de jets *taggables*. Ainsi l'efficacité d'identification d'un jet b est le rapport du nombre de jets b ayant passé la coupure de l'algorithme sur le nombre de jets b *taggables*. De même, l'efficacité de mauvaise identification (ou *mistag*) correspondra au nombre de jets légers (u, d, s, jets de gluons) identifiés comme jets b (c'est-à-dire ayant passé cette même coupure) sur le nombre de jets légers *taggables*.

3.1.2.5 Réjection des V^0

Les algorithmes d'identification de jets de quarks beaux sont, par construction, sensibles à toute particule de grand temps de vol. Mais une telle particule n'est pas forcément un hadron beau. D'autres, communément appelées V^0, telles que certains hadrons étranges (K_S ou encore Λ) ont de grands temps de vol, ou les conversions de photons $\gamma \rightarrow e^+e^-$ dans le matériau du détecteur ont lieu loin de la zone d'interaction. Ces V^0 simulent des hadrons beaux pour les algorithmes et conduisent à l'identification des jets les incluant comme beaux. Il est donc judicieux, en amont de l'identification, de rejeter les candidats V^0. Pour ce faire, leur sélection est menée de la façon suivante, à partir de deux traces de charges opposées potentiellement issues de leur désintégration :

[3] Ces algorithmes, SVT, CSIP et un réseau de neurones, seront décrits en détails par la suite, au paragraphe 3.2.

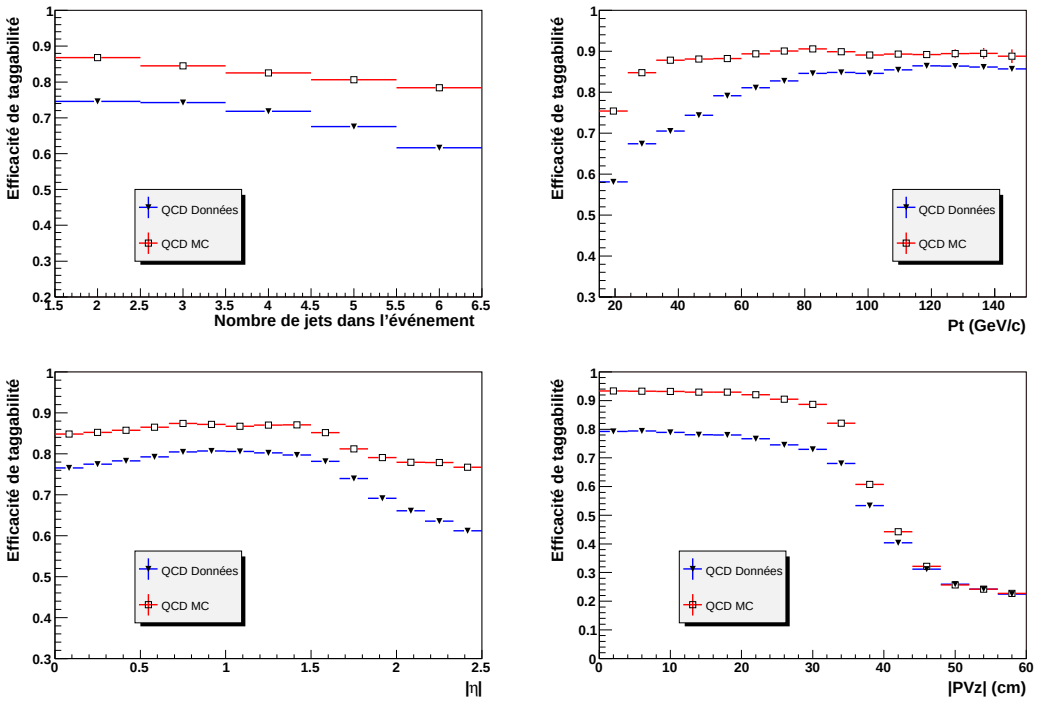

FIG. 3.1: Comparaison des efficacités de *taggabilité* entre données et simulation QCD, pour la version p17 de l'algorithme JLIP et la reconstruction du vertex primaire par l'algorithme APV. L'efficacité de *taggabilité* est représentée selon la multiplicité en jets de l'événement (figure haut gauche), en fonction de l'impulsion transverse du jet (haut droite), de sa pseudo-rapidité (bas gauche), et de sa distance au vertex primaire selon l'axe du faisceau (bas droite).

- la significance du paramètre d'impact[4] de chacune des deux traces doit être supérieure à 3 ;
- leur écart selon z par rapport au vertex primaire ne doit pas dépasser 1 cm pour rejeter les traces mal reconstruites ;
- le paramètre d'impact de la trajectoire reconstruite du candidat V^0 à partir de ces deux traces ne doit pas excéder 200 μm puisqu'il est produit au vertex primaire ;
- la masse invariante m_{V^0} reconstruite à partir de ces deux traces doit appartenir à la fenêtre de masse correspondante : $472 < m_{V^0} < 516$ MeV/c^2 pour un K^0_s, $1108 < m_{V^0} < 1122$ MeV/c^2 pour un Λ, $m_{V^0} < 25$ MeV/c^2 pour une conversion de photons ;
- pour les conversions de photons une condition supplémentaire doit être remplie par les traces : dans la zone où elles sont parallèles l'une à l'autre, leur écart ne doit pas dépasser 30 μm, simulant un angle négligeable entre l'électron et le positron.

[4] Cf. paragraphe 3.1.3

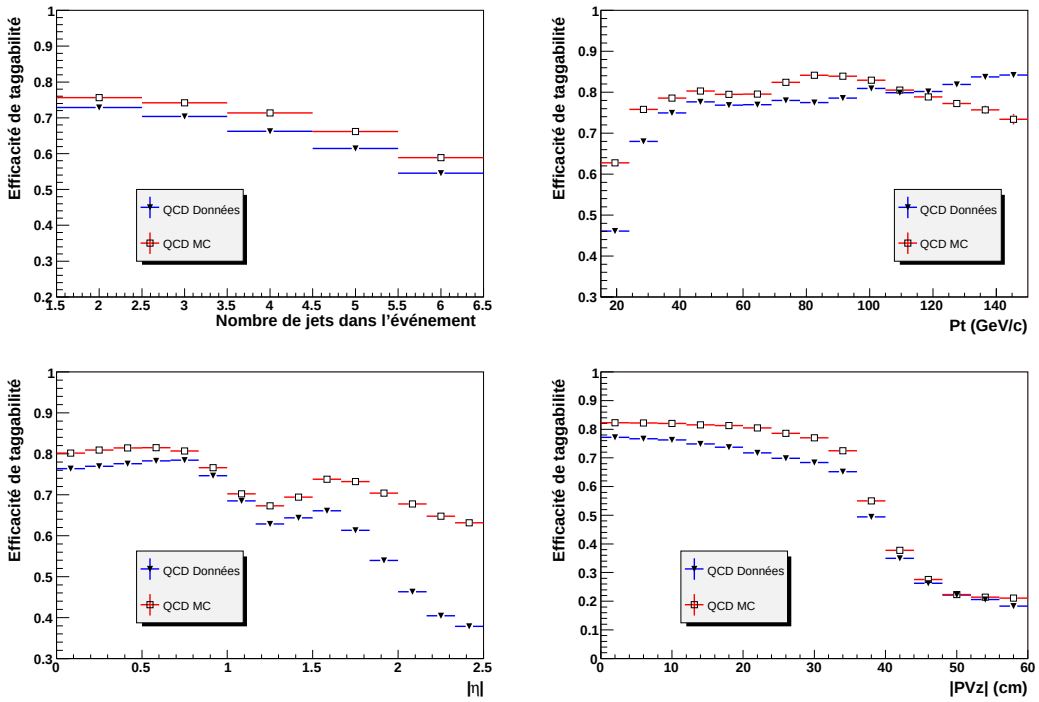

FIG. 3.2: Comparaison des efficacités de *taggabilité* entre données et simulation QCD, pour la version p20 de l'algorithme JLIP. L'efficacité de *taggabilité* est représentée selon la multiplicité en jets de l'événement (figure haut gauche), en fonction de l'impulsion transverse du jet (haut droite), de sa pseudo-rapidité (bas gauche), et de sa distance au vertex primaire selon l'axe du faisceau (bas droite).

Les paires de traces remplissant ces critères sont ainsi rejetées. Tous les algorithmes d'identification (JLIP, CSIP, SVT puis le réseau de neurones) appliquent cette réjection des V^0.

En pratique la réjection des V^0 n'est pas très efficace car elle exige la reconstruction des deux traces issues de leur désintégration. Lorsqu'une seule de ces deux traces remplit les critères de sélection des traces, elle sera utilisée par l'algorithme d'identification des jets beaux et contribuera à l'identification du jet comme effectivement beau.

3.1.3 Définition du paramètre d'impact d'une trace

L'identification des jets beaux par JLIP utilise le paramètre d'impact des traces pour identifier des jets. Dans DØ, les jets reconstruits le sont uniquement avec l'information calorimétrique. Les traces utilisées pour l'identification sont associées au jet calorimé-

trique à l'aide d'un algorithme de cône, sans que l'énergie du jet soit modifiée.

De par la présence du solénoïde, les particules chargées ont des trajectoires hélicoïdales. Le paramètre d'impact d'une trace est défini comme la distance minimale entre cette trace et le vertex primaire reconstruit. Le paramètre d'impact à deux dimensions IP_{xy} est la distance entre la projection de la trajectoire sur le plan transverse à l'axe du faisceau et le vertex primaire. La résolution suivant z est assez médiocre, et JLIP n'utilise que IP_{xy}[5]. Une particule de grand temps de vol, tel un hadron B, se désintègrera en un vertex secondaire éloigné, et les traces issues de ce vertex auront alors de grands paramètres d'impact. Ce paramètre d'impact (noté IP) est signé par $\cos(\overrightarrow{IP}, \overrightarrow{jet})$. Cela signifie que le paramètre d'impact d'une trace sera positif si cette trace se situe en avant du vertex primaire (i.e. si la trace intercepte la direction du jet), et négatif dans le cas contraire (la trace n'intercepte pas la direction du jet). La définition du paramètre d'impact et de son signe est illustrée par la figure 3.3, où θ correspond à l'angle $(\overrightarrow{IP}, \overrightarrow{jet})$. La largeur de la distribution en paramètre d'impact IP pour des traces issues du vertex primaire correspond à la résolution sur la reconstruction des trajectoires convoluée avec celle sur le vertex primaire.

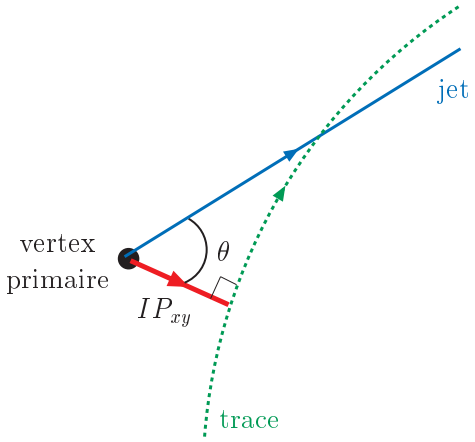

FIG. 3.3: Définition du paramètre d'impact transverse IP_{xy}. Il est signé par l'angle $\cos\theta$.

On définit ensuite la notion de significance d'une trace comme la valeur algébrique de son paramètre d'impact divisée par son erreur :

$$\mathcal{S}_{IP} = \frac{IP}{\sigma_{IP}}$$

Cependant il est apparu que l'erreur σ_{IP} n'est pas correctement estimée dans DØ. Afin de remédier à cela, une correction lui est appliquée. Elle est paramétrisée selon la qualité

[5]Par la suite, ce paramètre d'impact à deux dimensions IP_{xy} sera tout simplement noté IP. Le paramètre d'impact tri-dimensionnel n'interviendra plus.

de reconstruction de la trace, la valeur mesurée de son impulsion transverse P_T (afin de prendre en compte les diffusions multiples dues au matériau traversé), et la multiplicité de traces au vertex primaire. Cette correction sera explicitée au paragraphe 3.1.4.

Les traces provenant de la fragmentation de partons légers (u, d, s et gluons) présenteront une distribution en paramètre d'impact ainsi qu'en significance symétrique autour de 0, alors que les traces provenant de la décroissance de quarks b ainsi que de quarks c, et les traces de K_s^0 et de Λ qui sont des particules à grands temps de vie, contribueront aux valeurs positives plus élevées. La partie négative de la distribution de la significance $\mathcal{S}_{IP}$ est donc en première approximation peuplée uniquement par des particules issues du vertex primaire. La figure 3.4 compare les distributions de la significance du paramètre d'impact algébrique pour des traces appartenant à des jets issus de l'hadronisation de quarks légers, de quarks c ou de quarks beaux, distributions tirées de simulations QCD p20.

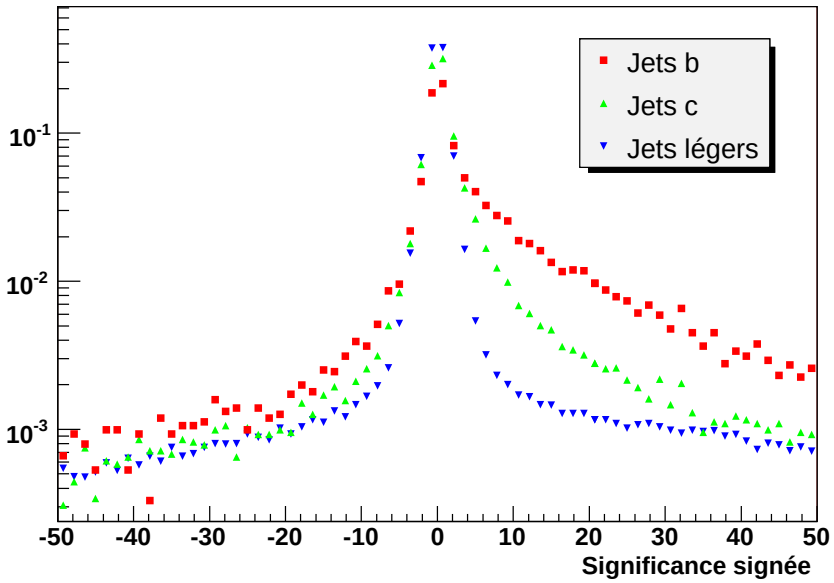

FIG. 3.4: Distributions normalisées de la significance en paramètre d'impact signée pour des jets de quarks b (carrés rouges), de quarks c (triangles verts) et de quarks légers (triangles inversés bleus) déterminés dans la simulation QCD p20. Un net excès dans la région des significances positives est visible pour les jets de saveurs lourdes.

3.1.4 Calibration de l'erreur sur le paramètre d'impact

Le paramètre d'impact vaut théoriquement zéro pour les traces provenant du vertex primaire. En pratique, à cause de la résolution imparfaite de la reconstruction des trajectoires, la distribution du paramètre d'impact est une gaussienne centrée en zéro. La largeur de cette distribution correspond alors à la résolution de la reconstruction. Si notre signal est constitué de plusieurs populations de traces ayant des résolutions différentes, alors la distribution du paramètre d'impact sera une superposition de gaussiennes centrées en zéro et de largeurs différentes.
Si l'erreur sur le paramètre d'impact est bien estimée, le *pull*[6], qui est dans ce cas IP/σ_{IP} c'est-à-dire la significance signée du paramètre d'impact, doit avoir une distribution gaussienne de largeur normalisée. Or, nous observons que ce n'est pas le cas car l'erreur est mal estimée. Il faut donc corriger l'erreur, par un facteur d'échelle correspondant à la largeur de la gaussienne ajustant les distributions de la significance signée du paramètre d'impact.

Afin de calibrer l'erreur σ_{IP}^{raw} sur le paramètre d'impact des traces, on introduit la variable suivante : $p_{scat} = p(\sin\theta)^{\frac{3}{2}}$, p étant l'impulsion de la particule et θ son angle polaire par rapport à l'axe du faisceau. En effet, dans le plan transverse, la dispersion due aux diffusions multiples est inversement proportionnelle à $P_T = p\sin\theta$ et proportionnelle à la racine carrée de la distance parcourue par la particule [71, 72]. En considérant que le détecteur a une distribution cylindrique centrée sur le faisceau, cette distance parcourue est également inversement proportionnelle à $\sin\theta$. Les distributions en paramètre d'impact ont été calculées dans 16 intervalles consécutifs en p_{scat}. De plus, les traces ont été classées en plusieurs super-catégories pour prendre en compte la qualité de la reconstruction de la trace. La correction à appliquer sera ainsi dépendante de la super-catégorie d'appartenance de la trace étudiée. Le tableau 3.1 décrit les 5 super-catégories définies pour la version p17 de l'algorithme, tandis que le tableau 3.2 décrit les catégories utilisées pour la version p20. L'ajout de la couche 0 du SMT a nécessité de définir 2 super-catégories supplémentaires pour cette version de l'algorithme, portant leur nombre à 7. Dans les deux cas, la première super-catégorie correspond aux traces avant (hors de l'acceptance du CFT), les suivantes correspondant aux traces centrales.

Les distributions de significance signée sont déterminées en fonction des super-catégories et des intervalles en p_{scat}, et ajustées dans leur partie centrale par une gaussienne. La figure 3.5 présente la distribution de significance dans les données QCD p20 pour les 7 super-catégories, dans l'intervalle $2.5 < p_{scat} < 3$ GeV/c, ainsi que l'ajustement par une gaussienne de la partie centrale. Afin de corriger l'erreur, l'écart-type de chaque gaussienne est calculé. Il est représenté en fonction de p_{scat} et pour chaque super-catégorie sur la figure 3.6. Celle-ci compare les corrections déterminées dans les données et dans

[6] Le *pull* d'une variable correspond à la différence entre la valeur théorique de cette variable et sa valeur mesurée, divisée par son erreur.

Super-catégorie	Coups CFT	Coups SMT
1	≤ 6	≥ 0
2	7	1
3	7	2
4	7	3
5	7	4

TAB. 3.1: Définition des 5 super-catégories de traces pour la version p17 de l'algorithme JLIP.

Super-catégorie	Coups CFT	Coups SMT (sans coups L0)	Coups L0 SMT
1	≤ 6	≥ 0	≥ 0
2	7	1-2	0
		0	1
3	7	1	1
4	7	3-4	0
5	7	2	1
6	7	3	1
7	7	4	1

TAB. 3.2: Définition des 7 super-catégories de traces pour la version p20 de l'algorithme JLIP.

la simulation Monte-Carlo QCD pour la version p20 de l'algorithme JLIP. Dans le cas idéal, la distribution de la largeur des gaussiennes d'ajustement devrait être une droite horizontale de valeur 1. Dans le cas de la version p20 de l'algorithme, l'erreur sur le paramètre d'impact est globalement surestimée, tant dans les données que dans la simulation, et plus fortement dans la simulation.

Par ailleurs, la résolution sur le paramètre d'impact est également sensible à la qualité du vertex primaire, dont un estimateur est le nombre de traces qui lui sont attachées. Ainsi, une fois une première correction $\alpha_{trace}(s_i, p_{scat})$ déterminée en fonction du p_{scat} des traces et de la super-catégorie s_i à laquelle elles appartiennent, une nouvelle correction $\alpha_{PV}(N_{PV})$ sur la significance du paramètre d'impact est calculée, après application de la correction précédente à chaque trace. Celle nouvelle correction ne dépend plus que du nombre de traces attachées au vertex primaire de l'événement. La distribution de la correction d'erreur en fonction de la multiplicité de traces au vertex primaire est représentée sur la figure 3.7, ainsi que la paramétrisation correspondante.

En conclusion, la correction suivante de l'erreur sur le paramètre d'impact est appliquée à chaque trace, dépendant de sa super-catégorie d'appartenance s_i, de la valeur de

son p_{scat}, et du nombre de traces associées au vertex primaire de l'événement correspondant :

$$\sigma_{IP} = \sigma_{IP}^{raw} \cdot \alpha_{PV}(N_{PV}) \cdot \alpha_{trace}(s_i, p_{scat})$$

La résolution sur le paramètre d'impact ainsi corrigée pour l'algorithme p20 est illustrée sur la figure 3.8 pour chaque super-catégorie. Cette résolution se révèle meilleure dans la simulation que dans les données réelles, ce qui indique peut-être une prise en compte imparfaite de la matière traversée par les particules dans la simulation.

Une paramétrisation de la résolution corrigée sur le paramètre d'impact peut alors être déterminée sous la forme suivante :

$$\sigma_{IP} = \frac{a}{p_{scat}} + b$$

où le paramètre a décrit les effets de la diffusion multiple et le paramètre b la résolution asymptotique. Ces paramétrisations empiriques sont reportées sur les figures correspondantes (figures 3.8, 3.12 et 3.13).

J'ai déterminé et fourni à la collaboration les paramétrisations de ces corrections de l'erreur sur le paramètre d'impact pour les versions p17 et p20 de l'algorithme JLIP. Les corrections et résolutions évaluées pour l'algorithme p17 sont quasi identiques quel que soit l'algorithme de reconstruction du vertex primaire utilisé. Les figures 3.9 et 3.10 présentent respectivement pour l'APV et le 2PPV les corrections d'erreurs en fonction des super-catégories et du p_{scat} des traces. La figure 3.11 montre la paramétrisation de la correction due à la multiplicité en traces du vertex primaire (après application de la correction précédente), et les figures 3.12 et 3.13 (respectivement toujours pour l'APV et le 2PPV) montrent les résolutions en paramètre d'impact résultantes.

3.1.5 Description de l'algorithme d'identification JLIP

La distribution de la significance en paramètre d'impact des particules légères, qui représente la distribution des traces provenant du vertex primaire, peut être modélisée par une gaussienne dont la largeur correspond à la résolution sur le paramètre d'impact. Comme nous pouvons le voir sur la figure 3.4, les particules à grand temps de vol, telles que des hadrons beaux ou charmés, présentent un excès pour la significance en paramètre d'impact positive, tandis que la significance négative est indissociable de celle des particules légères.

Dans un premier temps, on définit la fonction de résolution du paramètre d'impact des traces $\mathcal{R}(\mathcal{S}_{IP})$ comme étant l'ajustement par des gaussiennes de la partie négative de la distribution de la significance (dont l'erreur a été corrigée) du paramètre d'impact. Cette courbe de référence $\mathcal{R}(\mathcal{S}_{IP})$ (qui est donc une distribution de probabilité), obtenue à partir de traces provenant a priori uniquement du vertex primaire, permet alors de

calculer une probabilité P_{trk} pour une trace de significance S_{IP} de provenir du vertex primaire :

$$P_{trk}(S_{IP}) = \frac{\int_{-\infty}^{-|S_{IP}|} \mathcal{R}(s)ds}{\int_{-\infty}^{0} \mathcal{R}(s)ds}$$

En pratique, la borne inférieure sur la significance est fixée à -50 dans l'intégrale, la queue de distribution en-deçà étant négligée.
Pour les traces issues du vertex primaire ayant un paramètre d'impact négatif, la distribution de probabilité $P_{trk}(S_{IP} < 0)$ est plate par construction. Pour les paramètres d'impact positifs, un pic significatif à basse probabilité $P_{trk}(S_{IP} > 0)$ apparaît, dû aux traces issues des désintégrations des particules belles et charmées, mais également des désintégrations des K_s^0, des Λ et aux conversions de photons qui n'ont pas pu être rejetées. La figure 3.14 illustre les distributions de probabilité par trace pour les différentes saveurs, pour des traces de significance en paramètre d'impact positive et négative.

Enfin, toutes les N_{trk}^+ (resp. N_{trk}^-) traces d'un jet ayant une significance positive (resp. négative), peuvent être utilisées afin de calculer la probabilité P_{jet}^+ (resp. P_{jet}^-) que toutes les traces du jet proviennent du vertex primaire, c'est-à-dire que ce jet est de saveur légère :

$$P_{jet}^{\pm} = \Pi^{\pm} \times \sum_{j=0}^{N_{trk}^{\pm}-1} \frac{(-\log \Pi^{\pm})^j}{j!} \quad \text{avec} \quad \Pi^{\pm} = \prod_{i=1}^{N_{trk}^{\pm}} P_{trk}(S_{IP<0}^{IP>0})$$

La variable d'identification JLIP est par définition la probabilité P_{jet}^+. La figure 3.15 présente les distribution de probabilité P_{jet}^+ et P_{jet}^- pour les différentes saveurs. Comme on peut le voir sur ces figures, il est possible d'enrichir un lot en jets issus de décroissances de b en appliquant une coupure sur cette probabilité.

3.1.6 Détermination des fonctions de résolution

Afin de déterminer les fonctions de résolution du paramètre d'impact des traces, les super-catégories, exposées auparavant et ayant servi à calibrer l'erreur sur le paramètre d'impact, sont étendues de manière à décrire aussi bien que possible les effets géométriques du détecteur et les effets dûs à la reconstruction des traces, tout en conservant une statistique suffisante dans chacune de ces nouvelles catégories. Les paramètres pris en compte sont le nombre de coups de la trace dans les détecteurs SMT et CFT, sa pseudo-rapidité $|\eta|$, son impulsion transverse P_T et le χ^2 de son ajustement. 29 catégories ont ainsi été définies avant l'ajout de la couche 0 du SMT, pour les algorithmes p14 et p17. Elles sont détaillées dans le tableau 3.3. Une description détaillée des détecteurs SMT et CFT, en particulier en ce qui concerne le nombre de couches, leur disposition et la couverture angulaire, a été présentée au paragraphe 2.2.2 et permet de comprendre le choix des critères des catégories.

Coups SMT	Coups CFT	P_T (GeV/c)	$\|\eta\|$	χ^2
2 inner	≤ 6	≥ 1	[1.6,2[ou ≥ 2	≥ 0
1	≥ 7	≥ 1	< 1.2	[0,2[ou ≥ 2
1	≥ 7	≥ 1	≥ 1.2	≥ 0
2-4	≥ 7	[1,2[, [2,4[ou ≥ 4	< 1.2	[0,2[
2-4	≥ 7	≥ 1	< 1.2	[2,4[ou ≥ 4
2-4	≥ 7	≥ 1	[1.2,1.6[	[0,2[ou ≥ 2
2-4	≥ 7	≥ 1	≥ 1.6	≥ 0

TAB. 3.3: Définition des 29 catégories de traces utilisées pour la calibration de l'algorithme JLIP dans sa version p17.

Pour prendre en compte au mieux l'information nouvellement disponible de la couche 0, 65 catégories, distinguant les traces ayant ou non un coup dans cette couche 0, ont cette fois été définies pour l'algorithme p20. Ces nouvelles catégories sont détaillées dans le tableau 3.4. Des catégories ont été ajoutées pour prendre en compte les traces de bas P_T ($0.5 < P_t < 1$ GeV/c) qui peuvent maintenant être utilisées grâce à une couche plus proche du point de collision, et certaines des anciennes catégories ont été regroupées.

Une fonction de résolution est déterminée séparement pour chacune de ces catégories, et indépendamment dans les données et la simulation. Nous avons paramétrisé ces fonctions de résolution par un ajustement par plusieurs gaussiennes de la partie négative de la significance en paramètre d'impact (de 3 à 5 gaussiennes, le nombre retenu correspondant à la combinaison fournissant le meilleur ajustement). Durant mon travail de thèse, j'ai déterminé ces fonctions pour les données et les simulations reconstruites en p17 et p20. Elles sont ensuite fournies à la collaboration pour être utilisées dans les analyses. La figure 3.16 présente les fonctions de résolution de 4 des 65 catégories déterminées dans les données pour l'algorithme p20, ainsi que leur paramétrisation.

3.1.7 Estimation des performances de l'identification des jets beaux par l'algorithme JLIP

Une courbe de performance de cet algorithme d'identification peut être établie (par intervalles en impulsion transverse p_T et en pseudo-rapidité $|\eta|$ du jet), donnant la relation entre l'efficacité d'identification des jets beaux, la probabilité de sélectionner des jets de quarks léger (u, d, s, gluon) et la coupure appliquée sur la variable JLIP. Afin d'obtenir une estimation de ces performances il est donc nécessaire de déterminer à la fois l'efficacité d'identification des jets beaux et celle de sélection des jets de saveur légère (*mistag*, ou mauvaise identification).

Coups SMT	Couche 0 du SMT	Coups CFT	P_T (GeV/c)	$\|\eta\|$	χ^2
1 ou 2 inner	0-1	≤ 6	≥ 1	[1.6,2[ou ≥ 2	≥ 0
1	0	≥ 7	≥ 1	$\forall$	≥ 0
0	1	≥ 7	≥ 0.5	$\forall$	≥ 0
2	0	≥ 7	≥ 1	< 1.2	[0,2[ou ≥ 2
2	0	≥ 7	≥ 1	≥ 1.2	≥ 0
1	1	≥ 7	[0.5,2[ou ≥ 2	< 1.2	[0,2[ou ≥ 2
1	1	≥ 7	[0.5,2[ou ≥ 2	≥ 1.2	≥ 0
3-4	0	≥ 7	[1,2[ou [2,4[ou ≥ 4	< 1.2	[0,2[
3-4	0	≥ 7	≥ 1	< 1.2	[2,4[ou ≥ 4
3-4	0	≥ 7	≥ 1	[1.2,1.6[ou ≥ 1.6	[0,2[ou ≥ 2
2	1	≥ 7	[0.5,1[, [1,2[, [2,4[ou ≥ 4	< 1.2	[0,2[
2	1	≥ 7	[0.5,2[ou ≥ 2	< 1.2	[2,4[ou ≥ 4
2	1	≥ 7	[0.5,2[ou ≥ 2	[1.2,1.6[	[0,2[
2	1	≥ 7	≥ 0.5	[1.2,1.6[	≥ 2
2	1	≥ 7	≥ 0.5	≥ 1.6	[0,2[ou ≥ 2
3	1	≥ 7	[0.5,1[, [1,2[, [2,4[ou ≥ 4	< 1.2	[0,2[
3	1	≥ 7	[0.5,2[ou ≥ 2	< 1.2	[2,4[ou ≥ 4
3	1	≥ 7	[0.5,2[ou ≥ 2	[1.2,1.6[	[0,2[ou ≥ 2
3	1	≥ 7	≥ 0.5	≥ 1.6	[0,2[ou ≥ 2
4	1	≥ 7	[0.5,1[, [1,2[, [2,4[ou ≥ 4	< 1.2	[0,2[
4	1	≥ 7	[0.5,2[ou ≥ 2	< 1.2	[2,4[ou ≥ 4
4	1	≥ 7	[0.5,2[ou ≥ 2	[1.2,1.6[	[0,2[ou ≥ 2
4	1	≥ 7	≥ 0.5	≥ 1.6	[0,2[ou ≥ 2

TAB. 3.4: Définition des 65 catégories de traces en fonction des paramètres des traces utilisées pour la calibration de l'algorithme JLIP dans sa version p20.

3.1.7.1 Efficacité d'identification dans la simulation Monte-Carlo

La simulation Monte-Carlo permet de déterminer les efficacités d'identification des jets b et de mauvaise identification. Cependant, les résultats obtenus se révèlent être difficilement transposables aux données, nous observons une meilleure efficacité de l'algorithme dans la simulation notamment car la multiplicité en traces des jets étiquetables dans les données est globalement plus faible que dans la simulation. Nous avons essayé de comprendre de manière approfondie cet effet en étudiant cette multiplicité en fonction des caractéristiques des jets et de la qualité des traces, dans le but d'implémenter une procédure d'élimination aléatoire de traces dans la simulation. Il n'a cependant pas été possible de déterminer une correction universelle de la multiplicité en traces des jets applicable à tous les processus simulés, les désaccords entre les données et la simulation ayant des causes bien trop complexes. Plus de détails sont disponibles dans la référence [73].

En conséquence, nous n'identifions pas directement les jets issus de quarks b dans la simulation mais nous pondérons chaque événement suivant une courbe de performance de l'algorithme d'identification telle que déterminée dans les données. Ces pondérations sont déterminées séparément selon la saveur des jets, et dépendent de l'impulsion transverse et de la pseudo-rapidité du jet. Une autre solution aurait été de déterminer un facteur d'échelle qui permettrait de faire une identification événement par événement plutôt que statistique comme nous le faisons en pratique. Cependant, appliquer ces facteurs d'échelle induirait des erreurs statistiques trop grandes dans la simulation, en particulier pour les jets légers. C'est pourquoi nous n'avons pas retenu cette méthode.

3.1.7.2 Efficacité d'identification dans les données

Dans les données, l'efficacité d'identification des quarks beaux doit être mesurée dans un échantillon enrichi en jets beaux. Le groupe IPHC a mis au point trois méthodes différentes pour déterminer l'efficacité de l'algorithme JLIP dans les données. Pour cela nous avons utilisé des données pour lesquelles un muon a été associé à un jet, car il est attendu qu'une large fraction de ces événements provienne de la désintégration semi-leptonique des hadrons charmés et beaux.

La première méthode consiste à ajuster la distribution en p_{Trel} (impulsion transverse relative à l'axe du jet) du muon par l'intermédiaire de trois distributions théoriques simulées paramétrisées selon E_T et η, une pour chaque saveur (quarks b, c, ou légers). Les paramètres libres de l'ajustement sont les nombres de jets b et de jets c présents dans l'échantillon. En réalisant cet ajustement avant et après l'application de la coupure sur JLIP, puis en comparant les deux nombres de jets b obtenus, une estimation de l'efficacité d'étiquetage des jets b est obtenue. Cependant cet ajustement n'est pas toujours stable

notamment à cause de la statistique faible dans les distributions théoriques simulées et du fait de la similitude entre les distributions des quarks c et des quarks légers.
La seconde méthode consiste à comparer le nombre de jets b après coupure sur JLIP entre les événements avec un seul jet étiqueté par JLIP et les événements avec deux jets identifiés et opposés l'un à l'autre. Dans cette méthode on néglige les corrélations entre hémisphères lorsque l'on identifie deux jets (le vertex primaire est le même pour les deux jets identifiés).
Enfin une troisième méthode, *SystemD*, a été développée par l'équipe DØ de l'IPHC. Cette méthode, dans sa version utilisée par la collaboration, consiste à résoudre un système de huit équations à huit inconnues, dont justement les efficacités recherchées. *SystemD* ne requiert que très peu d'information déterminée grâce à la simulation contrairement aux autres méthodes (uniquement pour déterminer des coefficients de corrélation). Plus de précisions sur cette méthode peuvent être trouvées dans les références [74, 75].
Ces trois méthodes se révèlent être en bon accord. Les efficacités présentées dans les paragraphes suivants ont été déterminées au moyen de la méthode *SystemD*. Celle-ci est la principale méthode de détermination de l'efficacité d'identification des jets b, car elle permet une meilleure maîtrise des incertitudes systématiques. Elle est utilisée pour la détermination de cette efficacité dans le cas du réseau de neurones.

Cependant ces méthodes permettent d'évaluer l'efficacité d'identification des quarks b dans les données uniquement pour des décroissances de b en modes semi-muoniques : $\varepsilon_{b\to\mu}^{donnees}(E_T,\eta)$. Pour tenir compte d'un éventuel biais, nous déterminons un facteur correctif $C = \varepsilon_b^{donnees}(E_T,\eta)/\varepsilon_{b\to\mu}^{donnees}(E_T,\eta)$. La simulation est nécessaire pour déterminer C, or nous avons vu que l'efficacité de l'identification se révèle plus élevée dans la simulation que dans les données. Afin de corriger cet effet, un facteur d'échelle appelé SF_b, paramétrisé selon η et E_T du jet, est déterminé et appliqué au Monte-Carlo. Dans un premier temps, on calibre l'efficacité semi-muonique déterminée dans le Monte-Carlo à celle mesurée dans les données par l'intermédiaire du facteur d'échelle suivant :

$$SF_{b\to\mu}(E_T,\eta) = \frac{\varepsilon_{b\to\mu}^{donnees}(E_T,\eta)}{\varepsilon_{b\to\mu}^{MC}(E_T,\eta)}$$

où $\varepsilon_{b\to\mu}^{MC}(E_T,\eta)$ est l'efficacité d'ídentification des jets b se désintégrant en muon déduite du Monte-Carlo. Puis on fait l'hypothèse $SF_b = SF_{b\to\mu}$, ce qui nous permet d'en déduire l'efficacité d'identification inclusive des quarks beaux $\varepsilon_b^{donnees}(E_T,\eta)$:

$$\varepsilon_b^{donnees}(E_T,\eta) = \varepsilon_b^{MC}(E_T,\eta)SF_b(E_T,\eta)$$

L'efficacité d'identification des quarks c, $\varepsilon_c^{donnees}(E_T,\eta)$, est déterminée de façon identique, en faisant en plus l'hypothèse $SF_c = SF_b$.

3.1.7.3 Estimation du taux de mauvaise identification dans les données

Pour estimer le taux de mauvaise identification ε_{light}, qui correspond à l'efficacité d'identification de jets légers dû essentiellement aux effets de résolution, on fait l'hypothèse que les traces à significance négative proviennent du vertex primaire. On détermine donc ε_{-}^{data} en appliquant une coupure sur la probabilité JLIP P_{jet}^{-} basée sur les traces de paramètre d'impact négatif :

$$\varepsilon_{-}^{data}(E_T, \eta) = \frac{\#\text{jets identifiés négatifs } (E_T, \eta)}{\#\text{jets taggables } (E_T, \eta)}$$

Finalement le taux de mauvaise identification ε_{light} est obtenu en corrigeant ε_{-}^{data} par deux effets déterminés dans le Monte-Carlo QCD :

- $SF_{hf}(E_T, \eta) = \frac{\varepsilon_{-}^{light}}{\varepsilon_{-}}$: contamination de la partie négative de la significance par des saveurs lourdes ;
- $SF_{ll}(E_T, \eta) = \frac{\varepsilon_{+}^{light}}{\varepsilon_{-}^{light}}$: contribution manquante des hadrons légers à grand temps de vie.

Le taux de mauvaise identification est alors :

$$\varepsilon_{light} = \varepsilon_{data}^{-} \cdot SF_{hf} \cdot SF_{ll}$$

La figure 3.17 présente les résultats de l'étude des efficacités de mauvaise identification menée pour la version p17 de l'algorithme JLIP (ces graphes correspondent à une coupure sur la probabilité JLIP de 0.5%).

Les figures 3.18 et 3.19 montrent respectivement l'efficacité d'identification de jets beaux et l'efficacité de mauvaise identification de saveurs légères distribuées selon l'énergie transverse du jet et selon sa pseudo-rapidité, pour l'algorithme p17 et une reconstruction du vertex primaire par l'APV. Les données utilisées pour la figure 3.18 sont celles contenant un muon associé à un jet. Le point de fonctionnement choisi pour l'algorithme pour ces figures est $\mathcal{P}_{JLIP} < 0.5\%$. Les résultats sont identiques pour le 2PPV. La simulation Monte-Carlo est toujours plus optimiste que les données en ce qui concerne l'identification des b (le rapport étant d'environ 0.8).
Pour le point de fonctionnement dit *Tight* de l'algorithme JLIP, on constate une efficacité d'identifications des jets b dans les données d'un peu plus de 30 %, pour une efficacité de mauvaise identification de 0.05 %.

Jusqu'à la version p14 de l'algorithme, JLIP était utilisé dans un grand nombre d'analyse (ainsi que l'algorithme SVT), et nous devions déterminer et fournir à la collaboration ces paramétrisations de l'efficacité d'identification (*Tagging Rate Function*, ou TRF), dépendant de la saveur des jets (jets b, jets c, ou jets légers), du point de fonctionnement de l'algorithme, et paramétrisées selon la pseudo-rapidité η et l'énergie transverse E_T du jet. Elles permettent d'appliquer des pondérations aux événements pour simuler la

probabilité qu'ils comportent un ou plusieurs jets identifiés comme beaux, pour une efficacité d'identification donnée.
Depuis la version p17 des algorithmes de reconstruction, la collaboration DØ a developpé une méthode plus sophistiquée d'identification des jets de quarks beaux, utilisant un réseau de neurones permettant de combiner plusieurs variables provenant de différents algorithmes comme JLIP. La probabilité JLIP est donc devenue une des variables d'entrée du réseau de neurones. Cette nouvelle méthode permet d'accroître grandement les performances de l'identification. La comparaison des courbe de performance des versions p17 de l'algorithme JLIP et du réseau de neurones est visible sur la figure 3.21, elle correspond à l'efficacité d'identification dans les données des jets de quarks beau en fonction de l'efficacité de mauvaise identification. Une nette améliorations des performances est mise en évidence, allant jusqu'à une augmentation de l'ordre de 50 % de l'efficacité d'identification. Cela est dû au fait que les variables des autres algorithmes d'identification ne sont pas totalement corrélées avec la probabilité JLIP. La détermination des TRF propres à JLIP n'est alors plus nécessaire. Néanmoins, la problématique de l'identification directe des jets beaux du Monte-Carlo est la même dans le cas de l'identification par le réseau de neurones. Des TRF sont ainsi déterminées et utilisées globalement pour cet algorithme.

Les références [76] à [75] proposent de nombreux compléments concernant l'algorithme JLIP, ainsi que sa calibration et sa certification pour sa version p14.

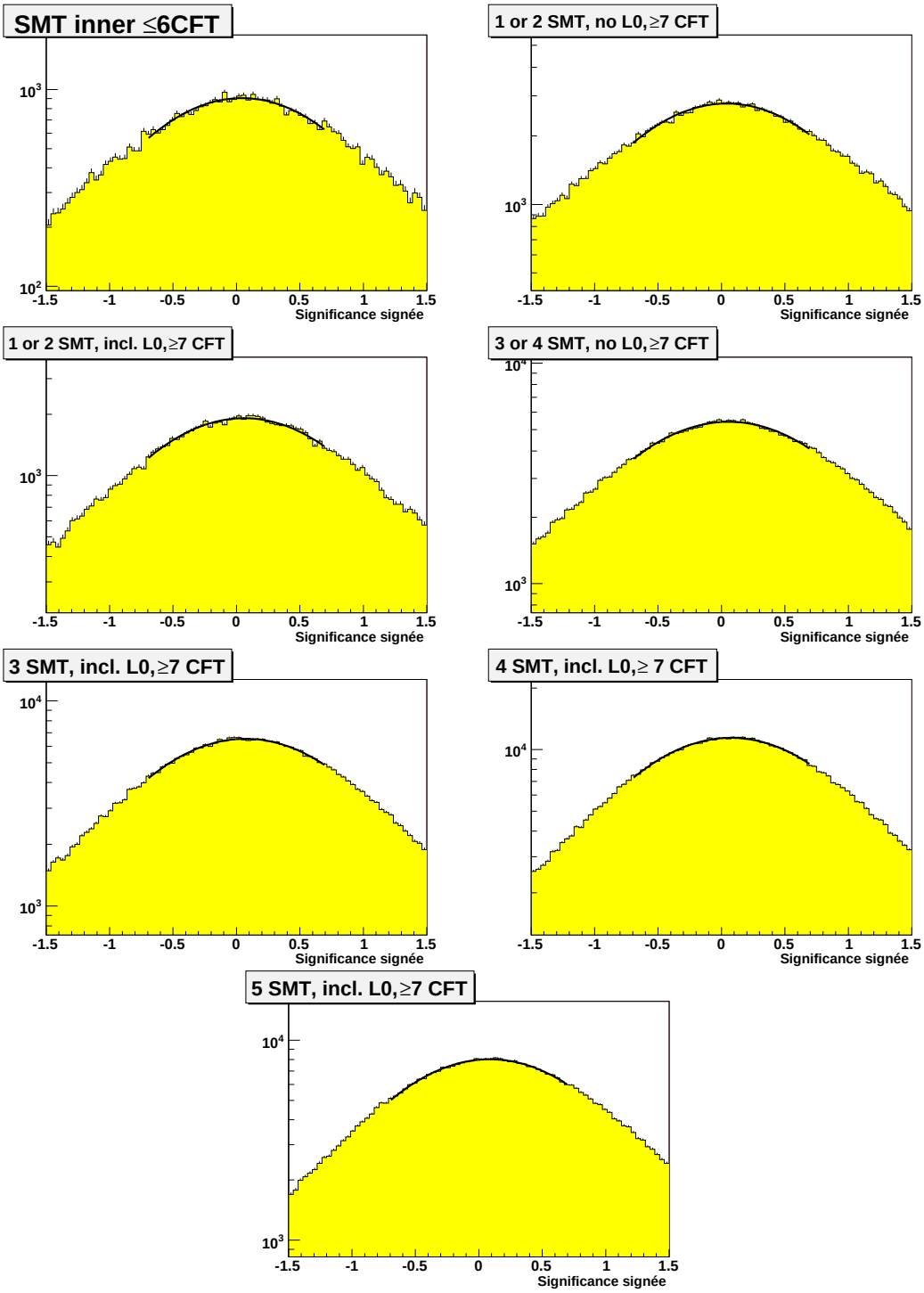

FIG. 3.5: Distribution de la significance signée du paramètre d'impact $\mathcal{S}_{IP}$ pour les 7 super-catégories et pour l'intervalle $2.5 < p_{scat} = p(sin\theta)^{3/2} < 3$ GeV/c dans les données QCD p20, et leur ajustement par une gaussienne dont l'écart-type σ est utilisé pour corriger la résolution du paramètre d'impact.

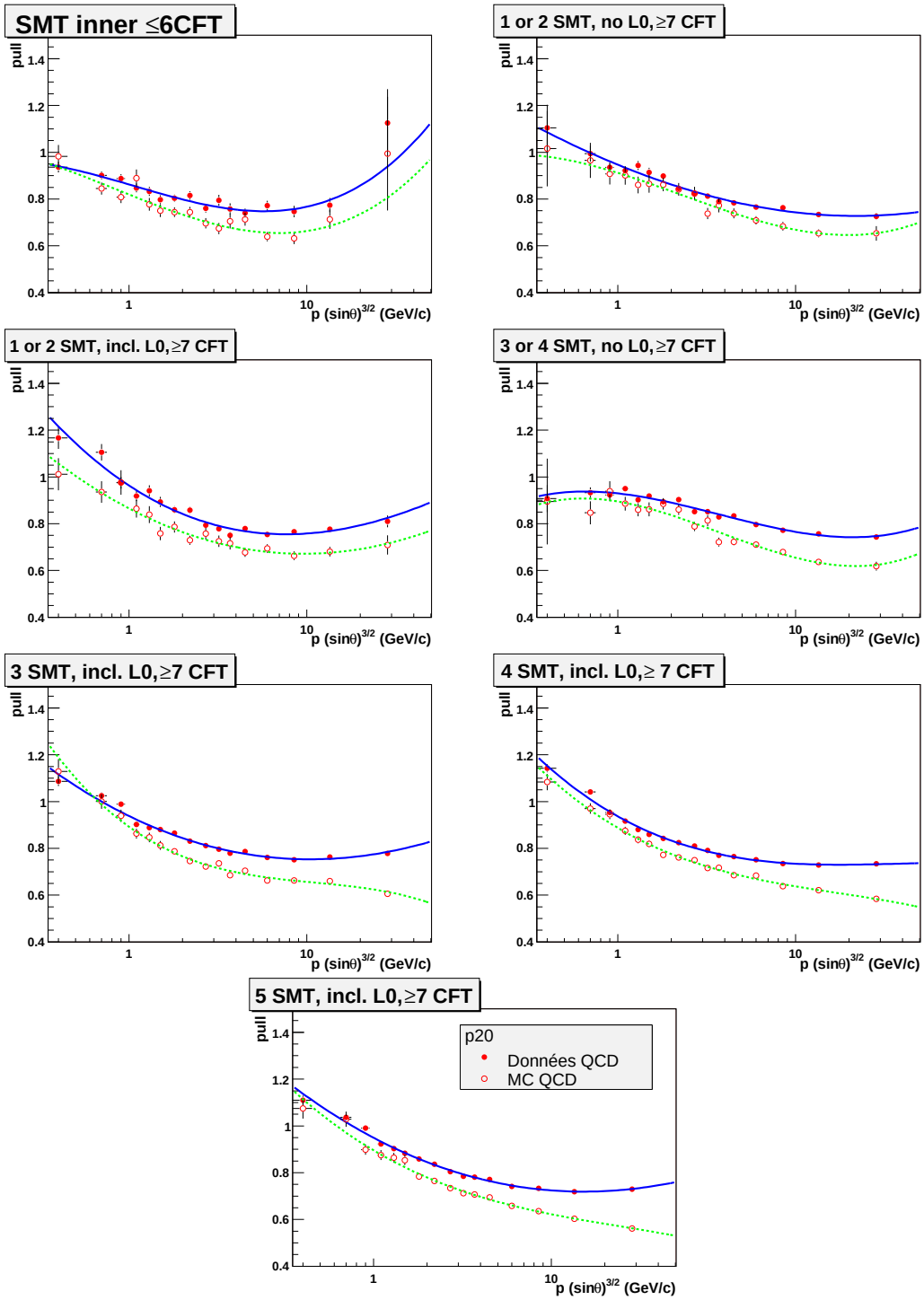

FIG. 3.6: Comparaison entre données (cercles rouges pleins) et simulation QCD p20 (cercles rouges vides) des écart-types des gaussiennes ajustant les distributions en $\mathcal{S}_{IP}^{raw}$ précédemment déterminées. Chaque graphe correspond à la distribution de corrections d'erreurs en fonction de p_{scat} pour une super-catégorie de traces donnée. La paramétrisation correspondante est représentée en trait plein bleu pour les données, en trait pointillé vert pour la simulation.

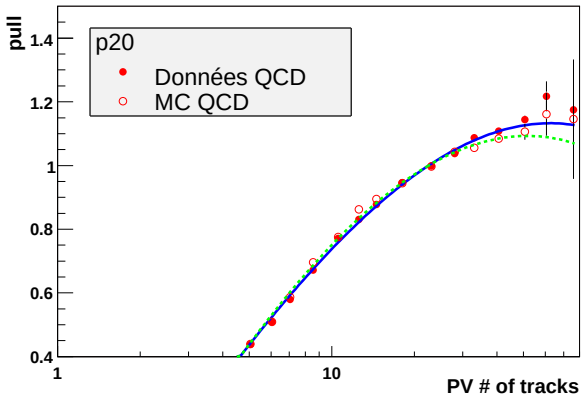

FIG. 3.7: Comparaison entre données (cercles rouges pleins) et simulation QCD p20 (cercles rouges vides) de la correction d'erreur due à la multiplicité en traces du vertex primaire. La paramétrisation correspondante est représentée en trait plein bleu pour les données, en trait pointillé vert pour la simulation.

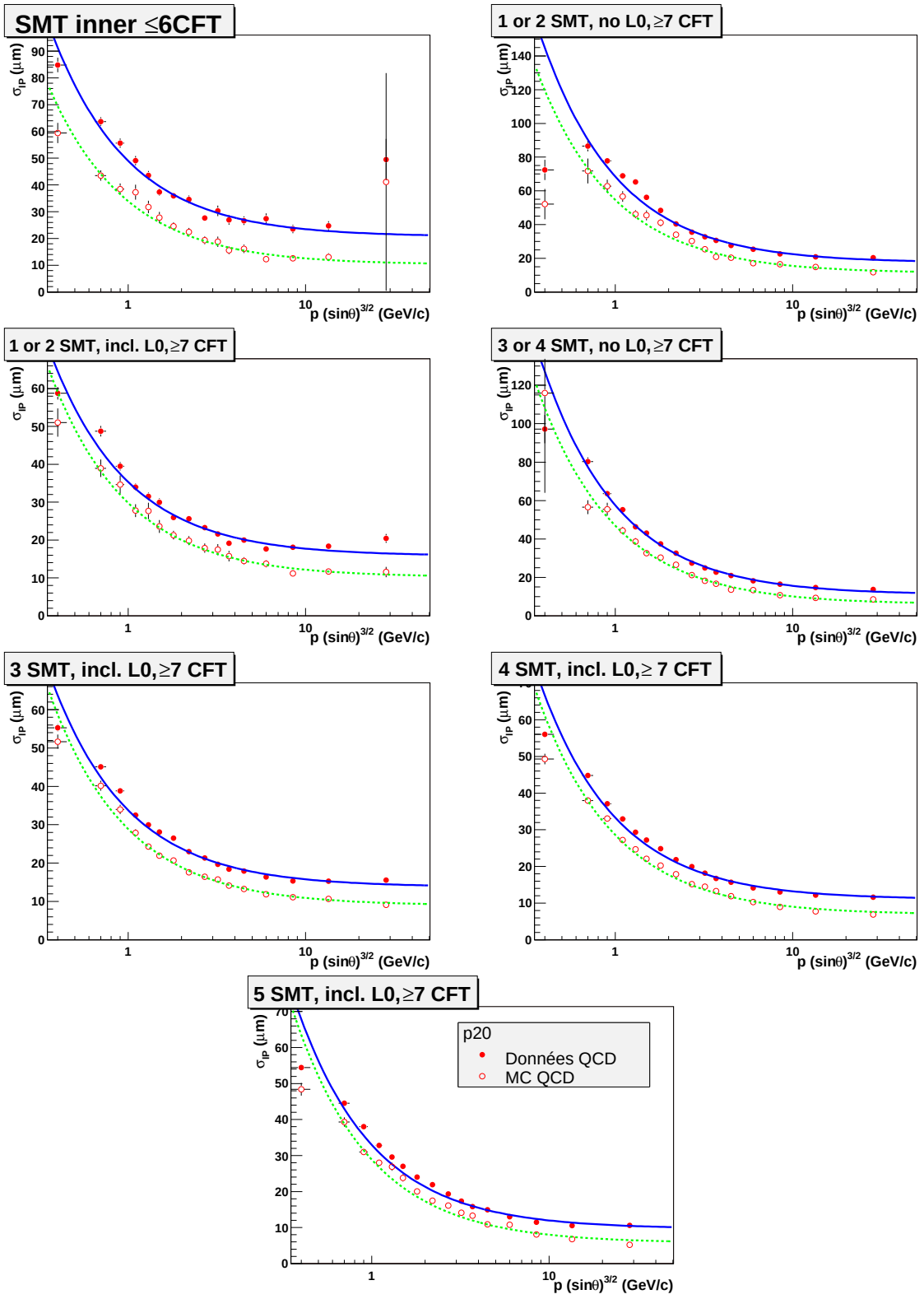

FIG. 3.8: Résolution sur le paramètre d'impact σ_{IP} corrigée pour chacune des sept super-catégories de traces, en fonction de p_{scat}. Les paramétrisations empiriques de la correction sont superposées aux distributions. Les données p20 sont en cercles rouges pleins, la simulation en cercles rouges vides et les paramétrisations correspondantes respectivement en trait plein bleu et en trait pointillé vert.

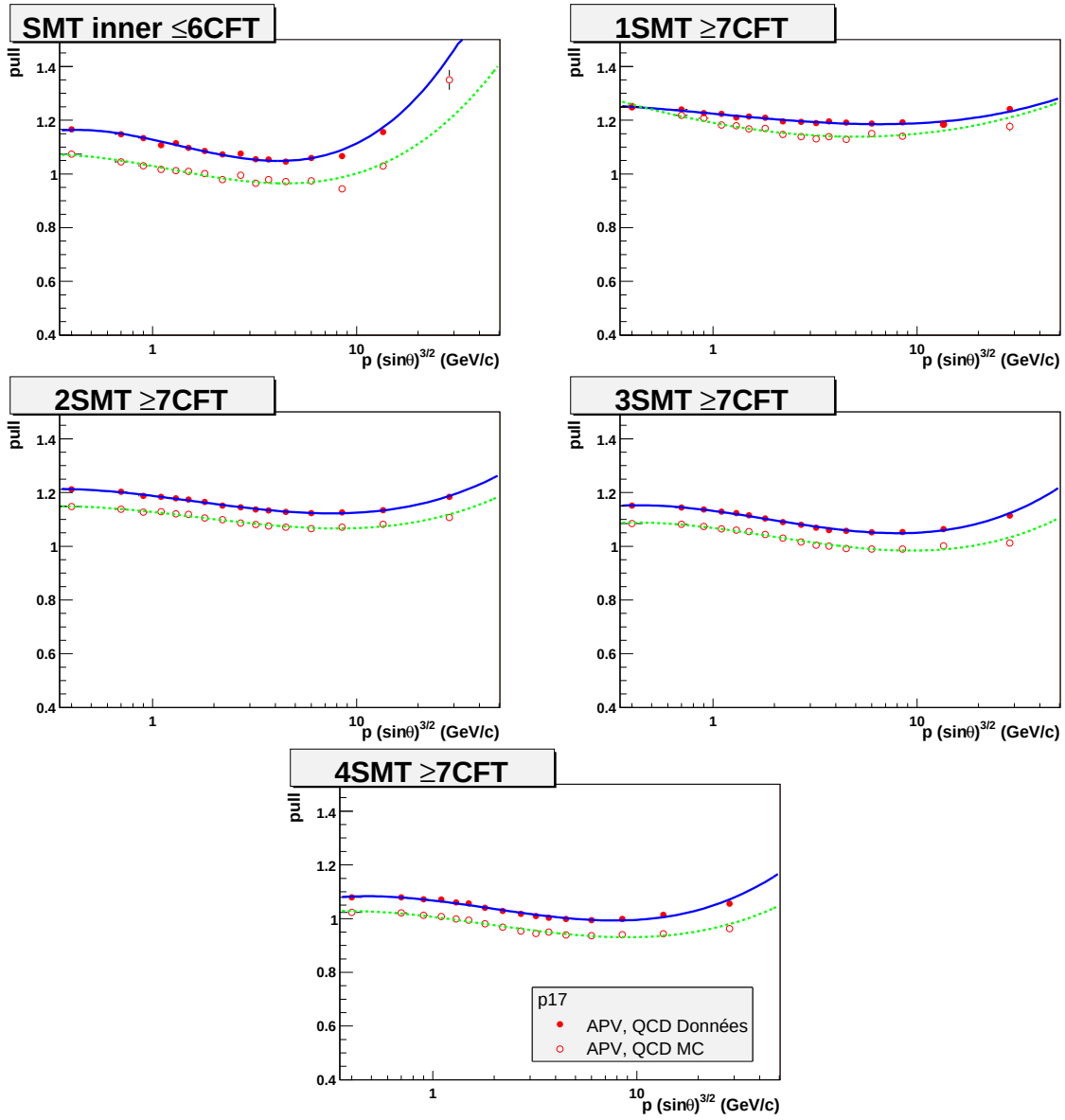

FIG. 3.9: Comparaison entre données (cercles rouges pleins) et simulation QCD p17 (cercles rouges vides) des écart-types des gaussiennes ajustant les distributions en $\mathcal{S}_{IP}$ précédemment déterminées. L'algorithme de reconstruction du vertex primaire utilisé ici est l'Adaptive Vertex (APV). Chaque graphe correspond à la distribution de corrections d'erreurs en fonction de p_{scat} pour une super-catégorie de traces donnée. La paramétrisation correspondante est représentée en trait plein bleu pour les données, en trait pointillé vert pour la simulation.

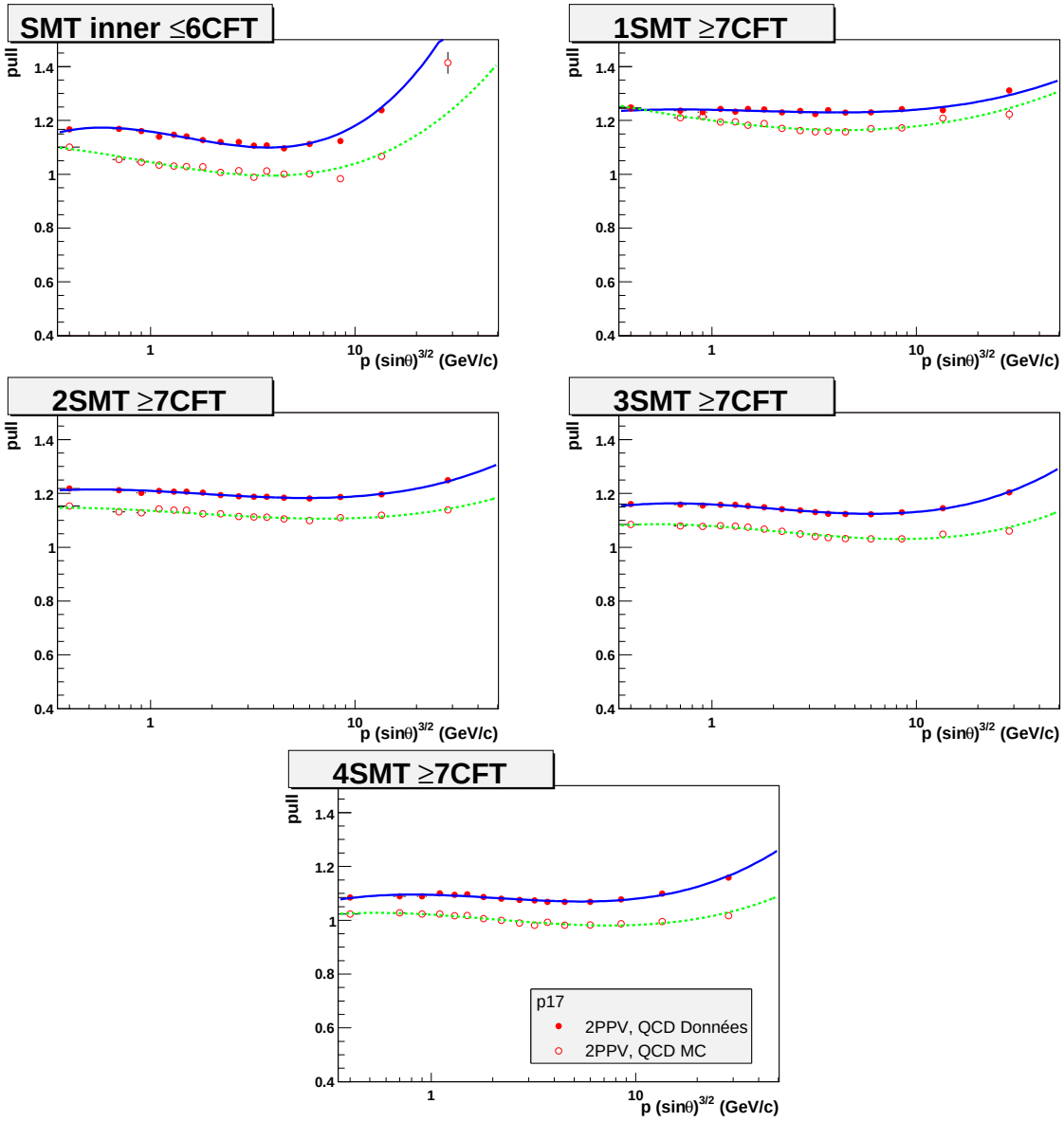

FIG. 3.10: Comparaison entre données (cercles rouges pleins) et simulation QCD p17 (cercles rouges vides) des écart-types des gaussiennes ajustant les distributions en $\mathcal{S}_{IP}$ précédemment déterminées. L'algorithme de reconstruction du vertex primaire utilisé ici est le 2-Pass Primary Vertex (2PPV). Chaque graphe correspond à la distribution de corrections d'erreurs en fonction de p_{scat} pour une super-catégorie de traces donnée. La paramétrisation correspondante est représentée en trait plein bleu pour les données, en trait pointillé vert pour la simulation.

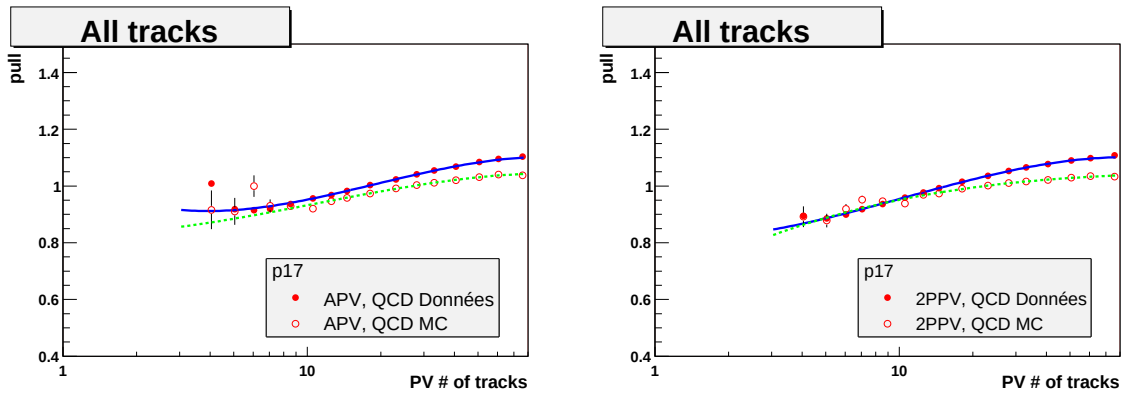

FIG. 3.11: Comparaison entre données (cercles rouges pleins) et simulation QCD p17 (cercles rouges vides) de la correction d'erreur due à la multiplicité en traces du vertex primaire. La paramétrisation correspondante est représentée en trait plein bleu pour les données, en trait pointillé vert pour la simulation. L'algorithme de reconstruction du vertex primaire utilisé est l'Adaptive Vertex (APV) pour la figure de gauche, et le 2-Pass Primary Vertex (2PPV) pour la figure de droite.

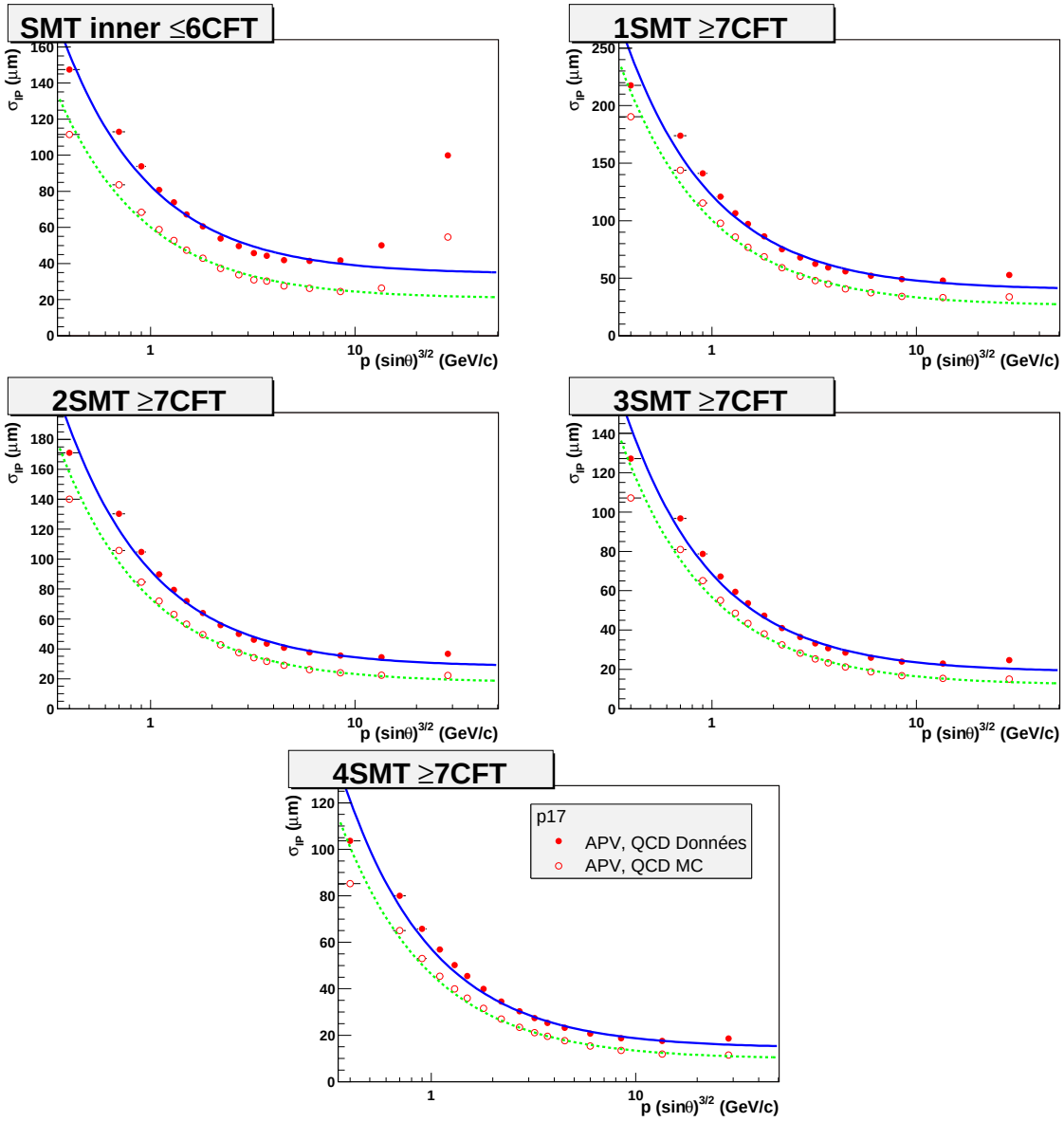

FIG. 3.12: Résolution sur le paramètre d'impact σ_{IP} corrigée pour chacune des cinq supercatégories, en fonction de p_{scat}. Les paramétrisations empirique de la correction sont superposées aux distributions. L'algorithme de reconstruction du vertex primaire utilisé ici est l'Adaptive Vertex (APV). Les données p17 sont en cercles rouges pleins, la simulation en cercles rouges vides et les paramétrisations correspondantes respectivement en trait plein bleu et en trait pointillé vert.

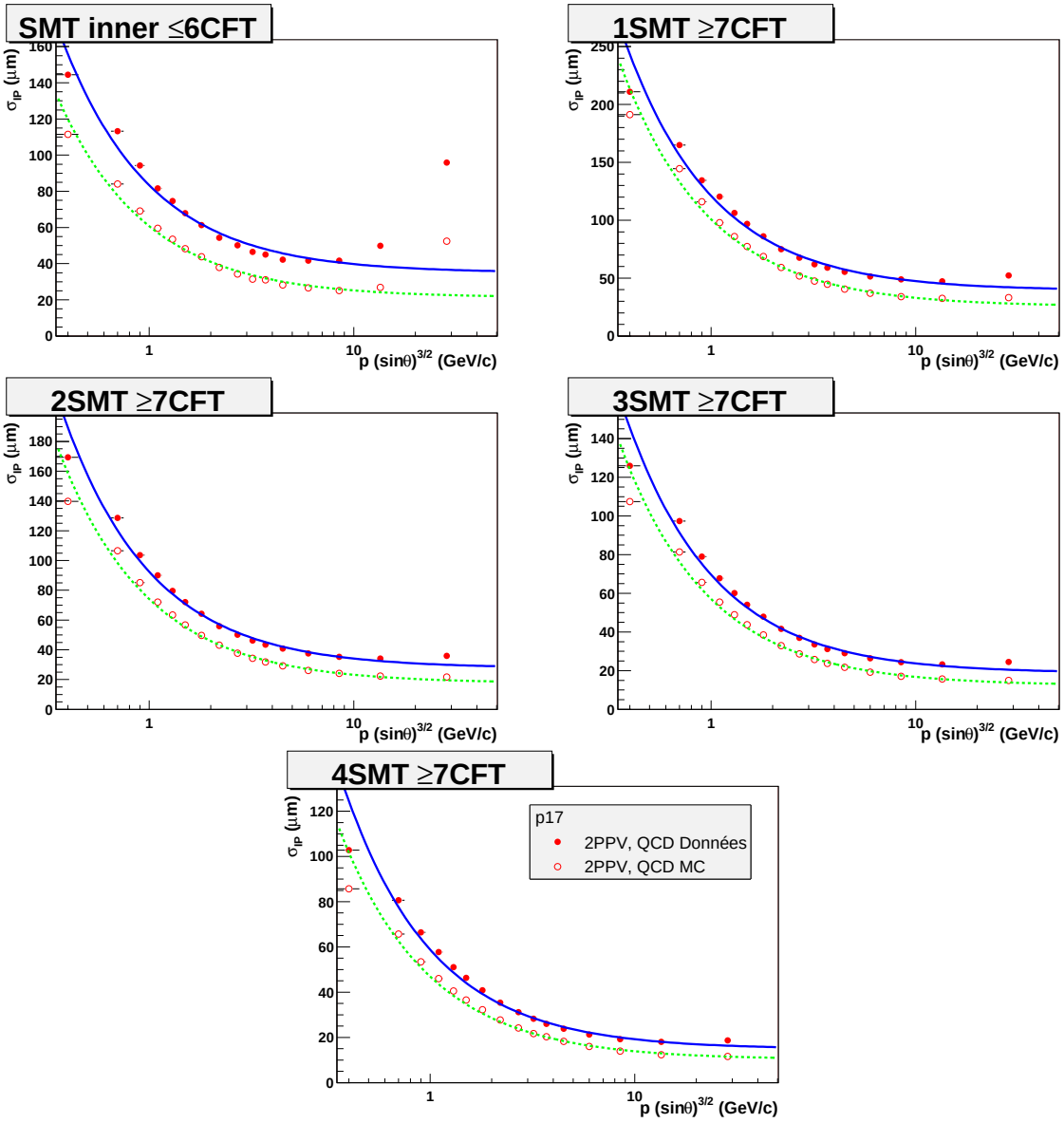

FIG. 3.13: Résolution sur le paramètre d'impact σ_{IP} corrigée pour chacune des cinq supercatégories, en fonction de p_{scat}. Les paramétrisations empirirques de la correction sont superposées aux distributions. L'algorithme de reconstruction du vertex primaire utilisé ici est le 2-Pass Primary Vertex (2PPV). Les données p17 sont en cercles rouges pleins, la simulation en cercles rouges vides et les paramétrisations correspondantes respectivement en trait plein bleu et en trait pointillé vert.

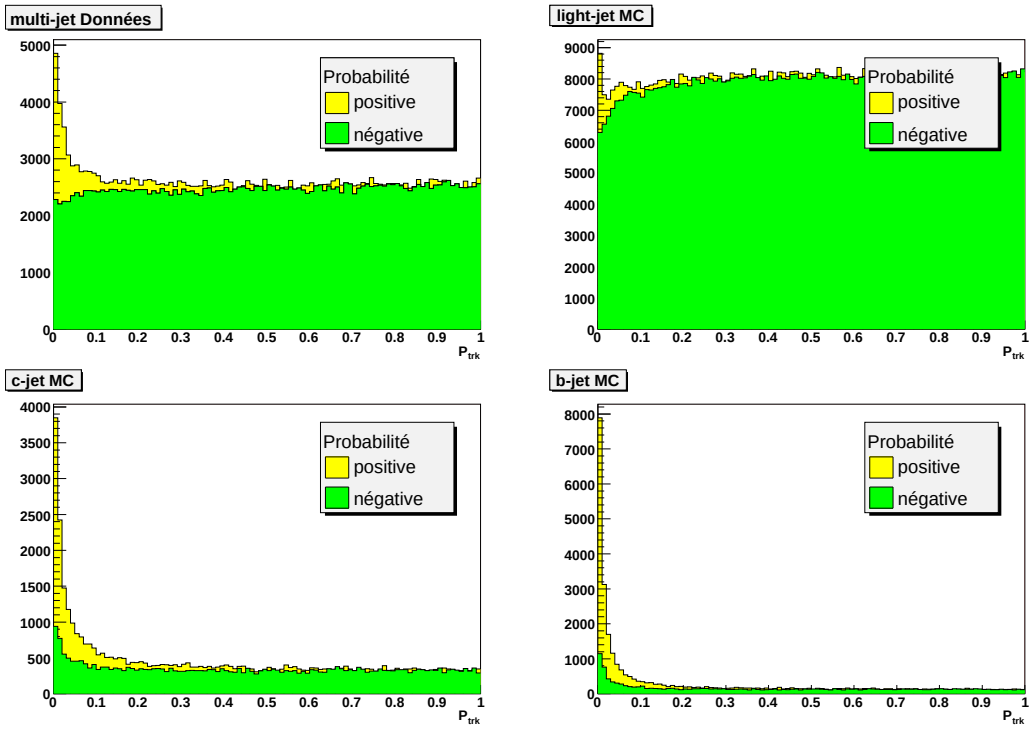

FIG. 3.14: Distributions de probabilité par trace, pour des traces de jets issus des données p20 multi-jets (figure haut gauche), et pour des traces de jets légers (haut droite), de jets c (bas gauche) et de jets b (bas droite) issus de la simulation QCD p20. La probabilité des traces de paramètre d'impact positif est en jaune, la probabilité négative en vert. On observe une nette remontée en 0 de la probabilité positive pour des traces issues de jets de saveurs lourdes, tandis que la probabilité est plate en dehors de ce pic à 0.

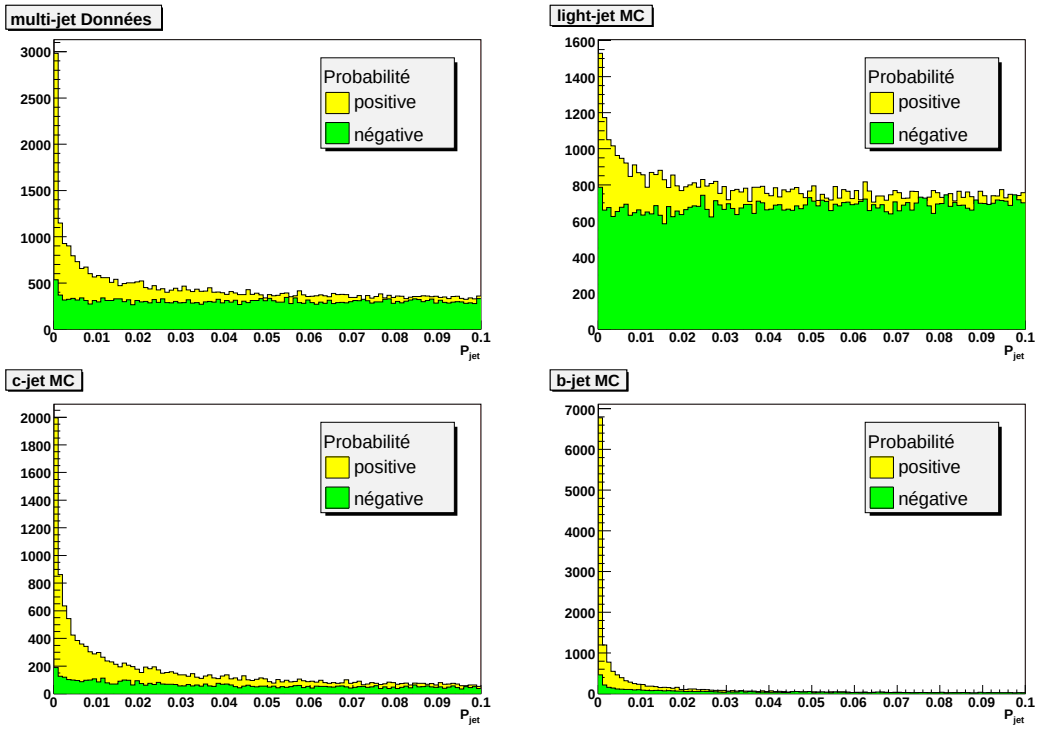

FIG. 3.15: Distributions de probabilité par jet pour des jets issus des données p20 multi-jets (figure haut gauche), et pour des jets légers (haut droite), de jets c (bas gauche) et des jets b (bas droite) issus de la simulation QCD p20. La probabilité positive est en jaune, la probabilité négative en vert et elles sont coupées à 0.1. On observe une nette remontée en 0 de la probabilité positive pour des jets de saveurs lourdes, et une probabilité plate en dehors de ce pic à 0, y compris dans la région non représentée.

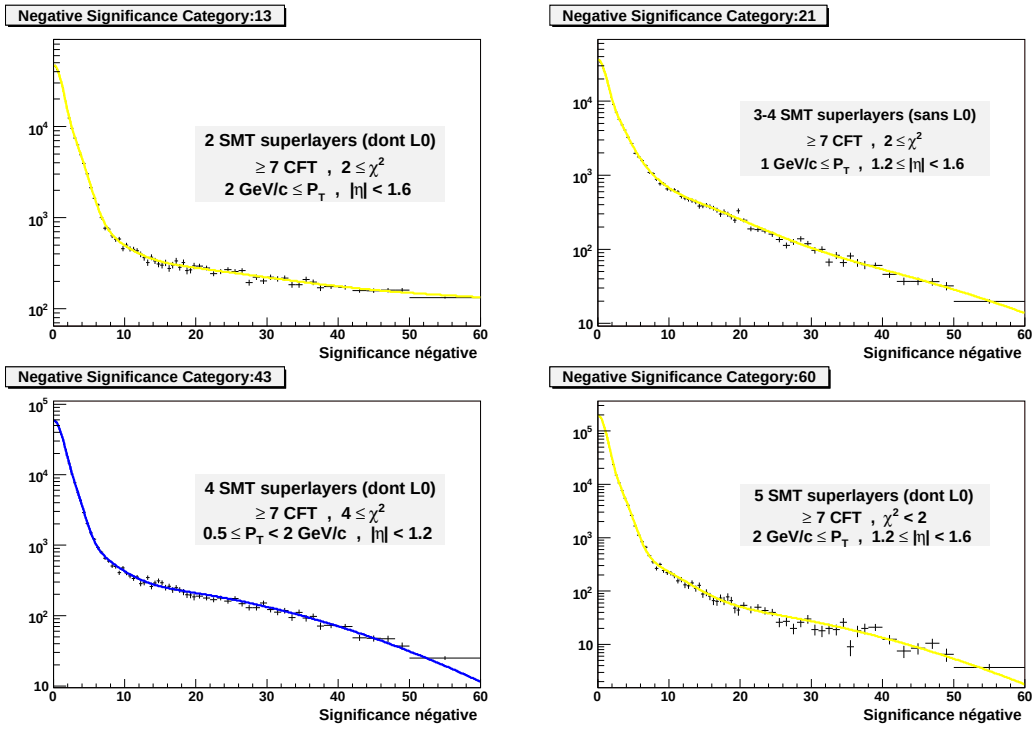

FIG. 3.16: Fonctions de résolution de 4 des 65 catégories déterminées dans les données QCD p20. Leur paramétrisation est réalisée au moyen de 4 gaussiennes (courbe d'ajustement bleue) ou de 5 gaussiennes (jaune).

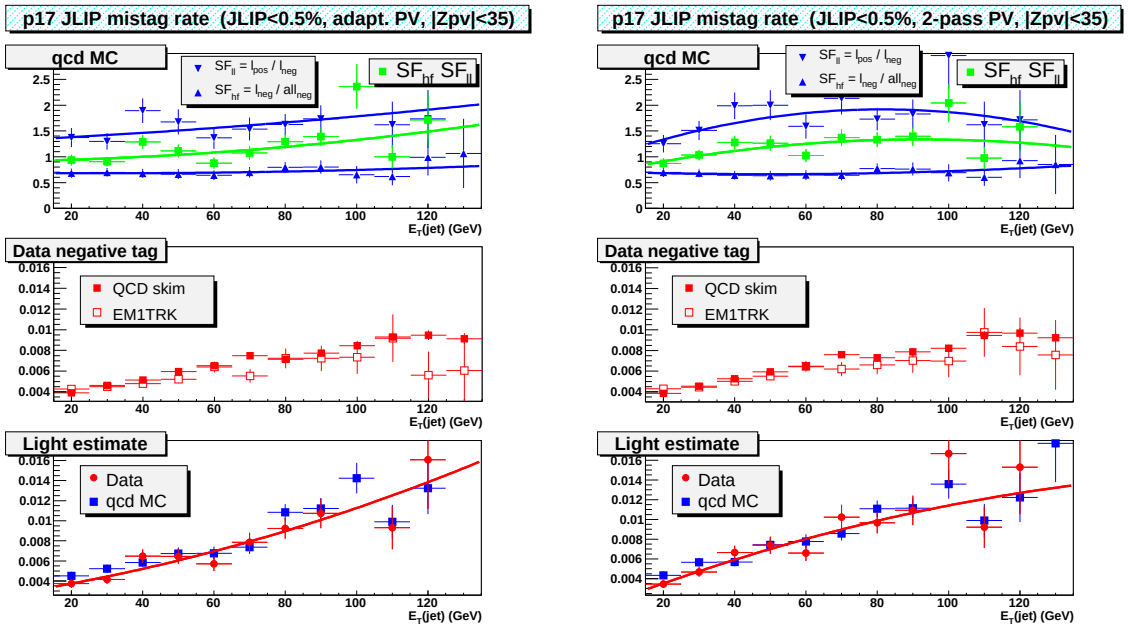

FIG. 3.17: Les figures de gauche correspondent aux résultats obtenus pour p17 avec une reconstruction du vertex primaire par l'algorithme *Adaptive Primary Vertex*, tandis que les figures de droite correspondent à une reconstruction par l'algorithme *2-Pass Primary Vertex*. La coupure appliquée sur la probabilité JLIP est de 0.5 %. Les graphes du haut représentent la distribution en fonction de l'énergie transverse des jets des facteurs d'échelle SF_{hf} et SF_{ll} et de leur produit. Les figures du milieu correspondent au taux ε_{-}^{data} de mauvaise identification avant correction déterminé dans les données QCD et EM1TRK, et montrent un très bon accord. Enfin, les graphes du bas comparent l'efficacité de mauvaise identification ε_{light} entre données et simulations.

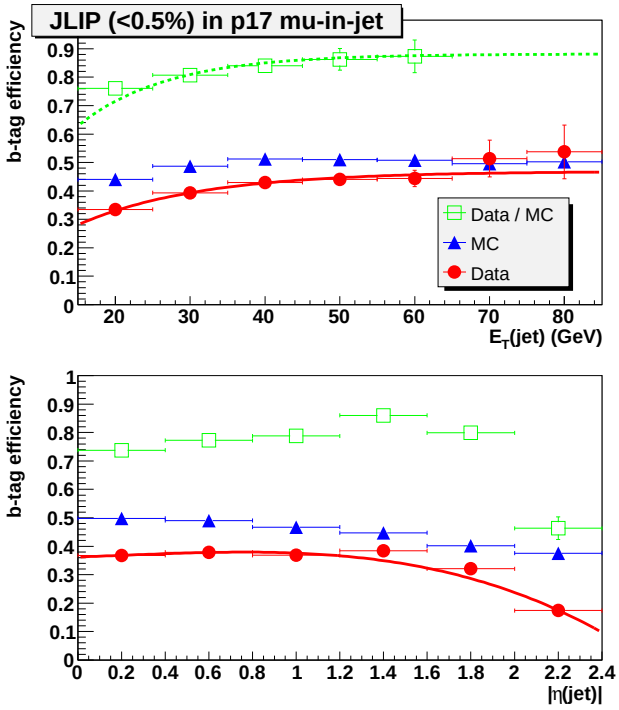

FIG. 3.18: Efficacité d'identification déterminée dans les données p17 possédant un muon associé au jet et dans la simulation pour une coupure sur la variable JLIP inférieure à 0.5 %, en fonction de l'énergie transverse du jet (figure du haut) ou de sa pseudo-rapidité (figure du bas). Les données sont représentées par des ronds rouges et la simulation par des triangles bleus. Les carrés verts correspondent au rapport données sur simulation. Le vertex primaire a été reconstruit avec l'algorithme APV

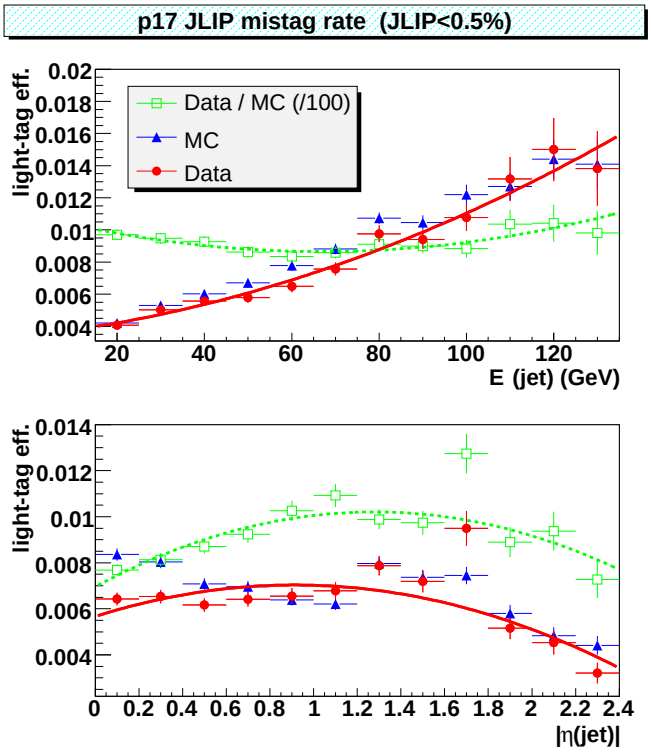

FIG. 3.19: Efficacité de mauvaise identification déterminée dans les données p17 possédant un muon associé au jet et dans la simulation pour une coupure sur la variable JLIP inférieure à 0.5 %, en fonction de l'énergie transverse du jet (figure du haut) ou de sa pseudo-rapidité (figure du bas). Les données sont représentées par des ronds rouges et la simulation par des triangles bleus. L'efficacité de mauvaise identification est assez bien représentée par le Monte-Carlo. Les carrés verts correspondent au rapport données sur simulation. Le vertex primaire a été reconstruit avec l'algorithme APV.

3.2 Identification des jets de quarks beaux par un réseau de neurones

Le réseau de neurone utilisé, qui sera désigné dans la suite par NN pour *Neural Network*, est un algorithme d'identification des jets de quarks beaux qui propose un point de fonctionnement continu. Pour ce faire, il combine les informations de trois algorithmes d'étiquetage existant dans DØ : SVT, CSIP, et JLIP. Parmi ces trois algorithmes, seul JLIP propose un point de fonctionnement continu. L'algorithme JLIP a été longuement présenté précédemment, les deux autres algorithmes vont être succintement décrits dans les paragraphes suivants.

3.2.1 Les algorithmes SVT et CSIP

3.2.1.1 L'algorithme d'identification SVT : Secondary Vertex Tag

Cet algorithme procède à la reconstruction de vertex secondaires en utilisant des traces d'un jet de traces associé à un jet calorimétrique *taggable*. Diverses quantités caractérisant ce vertex permettront ensuite d'identifier éventuellement ce jet comme étant un jet de quark beau.

Une fraction seulement des traces est utilisée, constituée de celles qui ont au moins 2 coups dans le SMT, un moment transverse $P_T > 3$ GeV/c et un paramètre d'impact inférieur à 1.5 mm. Toutes ces traces sont alors utilisées pour construire des vertex déplacés. L'efficacité de l'algorithme dépendra des contraintes appliquées. La version décrite ci-après correspond à la version *Super Loose* de l'algorithme, offrant une haute efficacité mais également un taux élevé de mauvaise identification. Les modifications entre parenthèses concernent, elles, la version *Loose*.

- (les traces doivent avoir une significance en paramètre d'impact dans le plan transverse $|\mathcal{S}_{IP}| > 3$) ;
- des vertex sont construits par un ajustement de Kalman sur toutes les paires de traces, et retenus si le χ^2 de l'ajustement est inférieur à 100 ;
- certaines des traces restantes sont ajoutées à la condition que l'accroissement $\Delta\chi^2$ soit inférieur à 15 (à 10) et que ce χ^2 ne dépasse pas 100. Une même trace peut donc contribuer à plusieurs vertex ;
- les vertex résultants sont conservés si l'angle α entre leur moment transverse reconstruit et la direction du vertex primaire au secondaire dans le plan transverse réalise la condition $\cos\alpha > 0.9$, car l'impulsion du hadron beau est parallèle à sa direction de vol ;
- (la significance du paramètre d'impact du vertex dans le plan transverse doit être supérieure à 5).

Plusieurs vertex secondaires sont ainsi reconstruits et associés à un même jet. Est retenu alors celui qui maximise la significance $\mathcal{S}_{xy}$ (cf. sa définition ci-après), qui est la variable la plus discriminante. Cinq variables discriminantes reliées à ce vertex vont servir de variables d'entrée au réseau de neurones. Le réseau de neurone tel qu'il est conçu nécessite une information sur les vertex déplacés. La version *Super Loose* de l'algorithme est alors utilisée pour assurer une bonne efficacité du NN. La version *Loose* est employée uniquement pour déterminer la variable N_{vtx} expliquée ci-dessous. Les cinq variables entrant dans le NN sont les suivantes :

$\mathcal{S}_{xy}$: la significance du paramètre d'impact dans le plan transverse du vertex secondaire par rapport au vertex primaire ;

χ^2_{dof} : le χ^2 par degré de liberté de l'ajustement des traces au vertex secondaire ;

N_{trk} : le nombre de traces considérées lors de la reconstruction du vertex secondaire ;

m_{vtx} : la masse invariante du vertex secondaire, calculée à partir de la combinaison des quadri-impulsions des traces ;

N_{vtx} : le nombre de vertex secondaires reconstruits pour un jet donné.

3.2.1.2 L'algorithme d'identification CSIP : Counting Signed Impact Parameter

De même que pour l'algorithme JLIP, les vertex secondaires ne sont pas reconstruits par cet algorithme. Celui-ci détermine une significance signée du paramètre d'impact $\mathcal{S}_{IP}$ des traces à partir de toutes les bonnes traces contenues dans un cône de 0.5 autour de l'axe du jet. Une bonne trace est définie ainsi pour cet algorithme :

- la différence selon l'axe z entre le paramètre d'impact et le vertex primaire doit être inférieure à 1 cm ;
- le paramètre d'impact de la trace doit être faible, inférieur à 2 mm ;
- l'ajustement de la trace doit être de bonne qualité : $\chi^2/N_{DF} < 3$;
- la trace doit comporter au moins 2 coups dans les couches internes du SMT ou au moins 3 coups dans des couches quelconques du SMT ;
- les traces comportant moins de 10 coups dans le CFT (traces dans la partie avant du détecteur) doivent avoir au moins 4 coups dans des couches stéréo du SMT.

Le signe de $\mathcal{S}_{IP}$ est parfois déterminé de manière incorrecte pour les traces proches de l'axe du jet. Un critère supplémentaire est alors défini, portant sur la différence d'angle azimutal $\Delta\phi$ entre la direction de la trace et celle du jet. Ainsi, quatre catégories de traces sont définies, l'algorithme réalisant un comptage de la population de chacune de ces catégories. Une combinaison particulière de ces populations sert ensuite de variable d'entrée au réseau de neurones. Les catégories sont les suivantes :

- traces telles que $\mathcal{S}_{IP}/1.2 > 3$ et $|\Delta\phi| > 20$ mrad, le nombre de traces de cette catégorie est notée N_{3s} ;

- traces telles que $2 < \mathcal{S}_{IP}/1.2 < 3$ et $|\Delta\phi| > 20$ mrad, N_{2s} traces composent cette catégorie ;
- traces telles que $\mathcal{S}_{IP}/1.2 > 3$ et $|\Delta\phi| < 20$ mrad (N_{3w} traces au total) ;
- traces telles que $2 < \mathcal{S}_{IP}/1.2 < 3$ et $|\Delta\phi| < 20$ mrad, (N_{2w} traces au total).

La variable d'entrée du réseau de neurones $\mathcal{N}_{CSIP}$ combine ces quatre variables (le nombre de traces de chaque catégorie), dans le double intérêt de réduire le nombre de variables d'entrées du NN et de mieux distribuer la variable. En effet, les quatre variables ci-dessus ont des valeurs faibles et discrètes et les réseaux de neurones préfèrent des variables continues étalées. La variable $\mathcal{N}_{CSIP}$ est définie comme suit :

$$\mathcal{N}_{CSIP} = 6 \times N_{3s} + 4 \times N_{2s} + 3 \times N_{3w} + 2 \times N_{2w}$$

Les poids appliqués ont été déterminés empiriquement afin d'optimiser les performances de cette variable dans le réseau de neurones.

3.2.2 Identification des jets de quarks beaux par un réseau de neurones

L'algorithme d'identification des jets issus de la désintégration de hadrons beaux est un réseau de neurones basé sur la combinaison d'un certain nombre de variables issues des algorithmes JLIP, SVT et CSIP définis précédemment. Ce faisant, les performances de ce nouvel algorithme se révèlent largement supérieures à celles de chacun des algorithmes classiques pris séparément.

Ce réseau de neurones a une structure de type $N_{input} : N_{hidden} : N_{output}$, avec une seule couche cachée (de 24 nœuds) afin de simplifier l'algorithme, sept nœuds d'entrée et un nœud de sortie. Les sept variables d'entrée ont été retenues pour leur pouvoir discriminant entre des jets de quarks b et les autres jets, et leur nombre a été fixé de manière à optimiser l'algorithme. Le tableau 3.5 liste ces variables, selon leur ordre d'importance. Cinq variables proviennent de l'algorithme SVT, une de JLIP et une de CSIP. La référence [81] présente le travail de construction et d'optimisation qui a été mené sur le réseau de neurones, et les références [82] et [83] ses performances déterminées respectivement dans les données p17 et p20.

L'algorithme est entraîné sur des lots d'événements Monte-Carlo QCD $b\bar{b}$ pour le signal et des lots d'événements QCD à jets légers pour le bruit de fond. Le nombre de cycles d'entraînement est fixé à 400. La figure 3.2.2 présente la distribution de la sortie du réseau de neurones (après optimisation) lorsqu'il est tourné sur le bruit de fond QCD de saveur légère et sur le signal $b\bar{b}$. La sortie du réseau de neurones est ainsi une variable continue comprise entre 0 et 1, qui tend vers la valeur 1 lorsque la probabilité que le jet soit un jet de quark b augmente.

Rang	Variable	Description
1	SVT_{SL} DLS	Significance de la longueur de désintégration du SV
2	CSIP Comb	Combinaison pondérée de la significance en IP des traces
3	JLIP Prob	Probabilité que le jet provienne du PV
4	$SVT_{SL}\chi^2_{dof}$	χ^2 par degré de liberté du SV
5	$SVT_L N_{Tracks}$	Nombre de traces utilisées pour la reconstruction du SV
6	SVT_{SL} Mass	Masse du SV
7	SVT_{SL} Num	Nombre de SV reconstruits pour un jet donné

TAB. 3.5: Variables d'entrée du réseau de neurones, classées par ordre d'importance.
PV : Vertex Primaire - SV : Vertex Secondaire - IP : Paramètre d'Impact.

Bien que permettant (à l'instar de JLIP et contrairement aux algorithmes SVT et CSIP) de définir un point de fonctionnement quelconque, les performances de ce réseau de neurones ont été évaluées pour douze points de fonctionnement en p17. Le tableau 3.6 liste ces points de fonctionnement et la coupure associée sur la variable de sortie du réseau de neurones. Les jets sont identifiés comme jets beaux si la variable de sortie du NN est supérieure à cette coupure. Le choix du point de fonctionnement optimal pour la mesure de l'hélicité du boson W faisant l'objet de cette thèse sera fait parmi ces points de fonctionnement.
La figure 3.21 compare les courbes de performance du réseau de neurones et de JLIP p17. Elle démontre une nette amélioration des performances de l'identification, augmentant jusqu'à 50 % l'efficacité d'identification pour un taux de mauvaise identification donné.

Nom	MegaTight	UltraTight	VeryTight	Tight	Medium	oldLoose
Coupure	0.925	0.9	0.85	0.775	0.65	0.5
Nom	Loose	L2	L3	L4	L5	L6
Coupure	0.45	0.325	0.25	0.2	0.15	0.1

TAB. 3.6: Points de fonctionnement du réseau de neurones p17.

3.2.3 Evolution future de l'algorithme d'identification des jets beaux dans DØ

Actuellement les deux expériences CDF et DØ auprès du Tevatron centrent la majeure partie de leurs efforts sur la recherche du boson de Higgs léger. Notons qu'un Higgs léger est nettement favorisé par la comparaison de l'ensemble des mesures dans le secteur électrofaible (effectuées au LEP, au SLC et au Tevatron) aux prédictions théoriques. La prédiction obtenue ainsi *via* la dépendance des corrections quantiques en la masse du

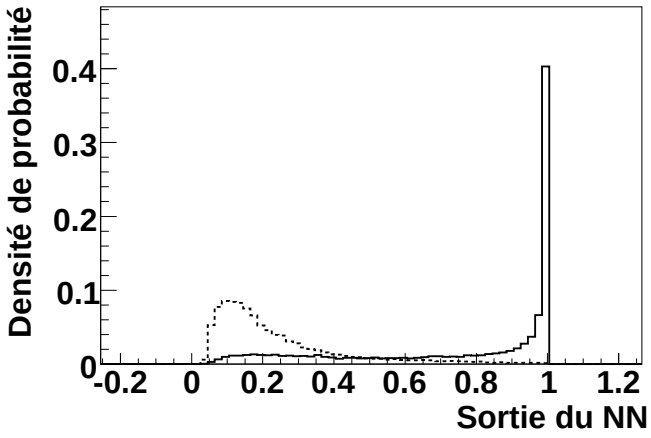

FIG. 3.20: Distributions normalisées de la sortie du réseau de neurones tourné sur du Monte-Carlo QCD à saveurs légères (distribution en pointillés) et sur du Monte-Carlo QCD $b\bar{b}$ (distribution en traits pleins).

Higgs est de $M_H = 87^{+36}_{-27}$ GeV/c^2 et $M_H < 160$ GeV/c^2 à 95 % de niveau de confiance (voir figure 3.22 [14]).

La luminosité délivrée par le Tevatron et son énergie de collision permettent de chercher le boson de Higgs du modèle standard en-dessous de 180 GeV/c^2 environ. Si le Higgs a une masse inférieure à 135 GeV/c^2, il se désintègre majoritairement en une paire de quarks b. L'algorithme d'identification des jets beaux est donc utilisé pour sélectionner les candidats Higgs. Une augmentation sensible des possibilités d'observation d'un Higgs léger est attendue d'une amélioration des performances de l'identification des jets beaux.

La complémentarité du Tevatron et du LHC pour découvrir un Higgs de masse inférieure à 135 GeV/c^2 est particulièrement importante. En effet les expériences CMS et ATLAS auprès du LHC n'auront pas une sensibilité suffisante pour découvrir le Higgs se désintégrant en une paire de quarks beaux et utiliseront le canal $H \to \gamma\gamma$ pour cette gamme de masse. Du côté du Tevatron, les études prospectives et les résultats déjà obtenus (voir figure 3.23 [84]) montrent qu'il y a une bonne probabilité d'exclure le Higgs léger avec les données du *Run II* (de l'ordre de 6 à 8 fb^{-1} par expérience d'ici fin 2009) ainsi que de le mettre en évidence à 3σ. Mais une découverte à 5σ semble largement hors de portée.

Plusieurs idées sont en cours d'étude dans DØ dans l'optique d'améliorer les performances de l'identification des jets beaux. Notamment le réseau de neurones dans sa version actuelle a été développé et optimisé de manière à discriminer les jets beaux des jets de saveurs légères (u, d, s et gluons). Il est peu discriminant envers les jets de quarks c, qui ont des propriétés intermédiaires entre les jets beaux et les jets légers. Le groupe DØ de l'IPHC s'attèle donc maintenant à la séparation b/c. Nous considérons plusieurs

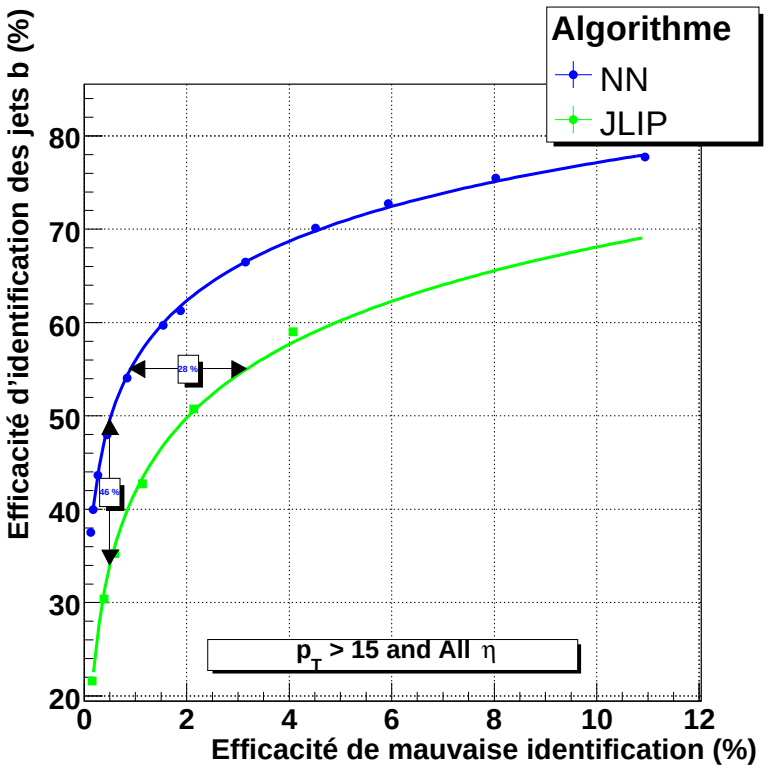

FIG. 3.21: Comparaison des courbes de performance du réseau de neurones et de JLIP, obtenues dans les données dans leur version p17. Le réseau de neurones améliore grandement les performances de l'identification, comparativement à JLIP, offrant jusqu'à 50 % d'augmentation du taux d'identification pour un taux de mauvaise identification donné. Pour une efficacité d'identification fixée, le taux de mauvaise identification est quant à lui réduit d'un tiers.

techniques avancées, comme les arbres de décision *boostés* ou les réseaux de neurones, pour exploiter les différences subtiles caractérisant les jets b des jets c (entre autres la masse, la multiplicité chargée et la rapidité des traces). Le but recherché est de garantir une efficacité de sélection des jets beaux de 90 % tout en rejetant 50 % des jets charmés en utilisant ce deuxième algorithme d'identification dédié, à la suite du premier (le réseau de neurones actuel).

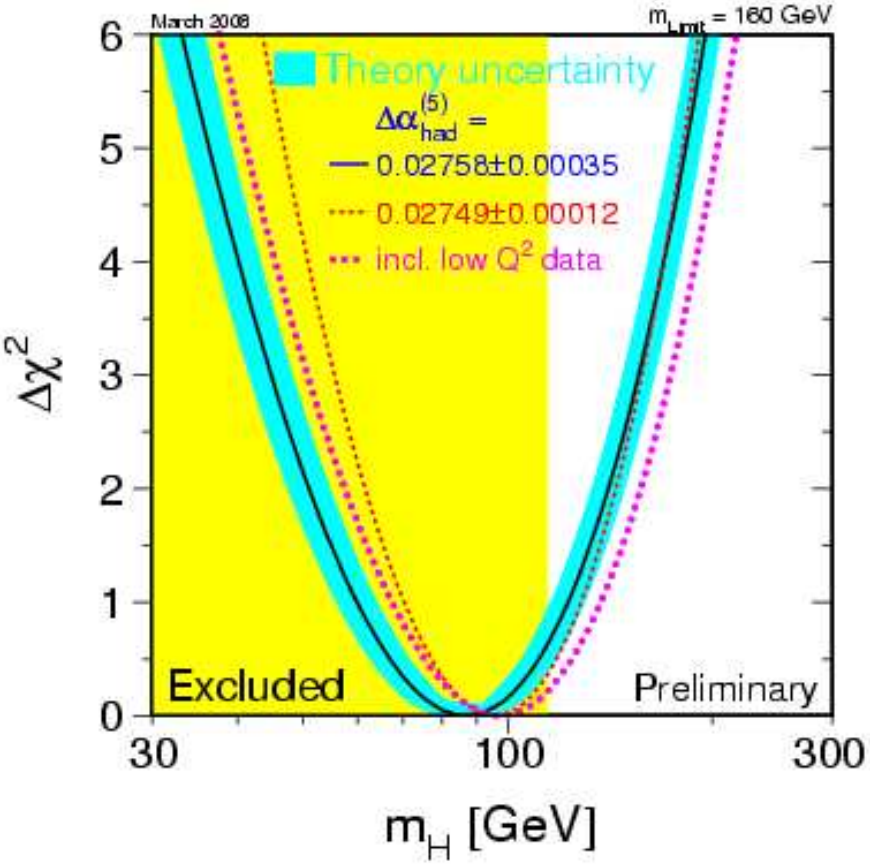

FIG. 3.22: Fonction du χ^2 de l'ensemble des mesures dans le secteur électrofaible comparées aux prédictions théoriques, dépendant de la masse du boson de Higgs *via* les corrections quantiques [14]. La valeur préférée est $M_H = 87^{+36}_{-27}$ GeV/c^2 et la limite supérieure à 95 % de niveau de confiance est $M_H < 160$ GeV/c^2.

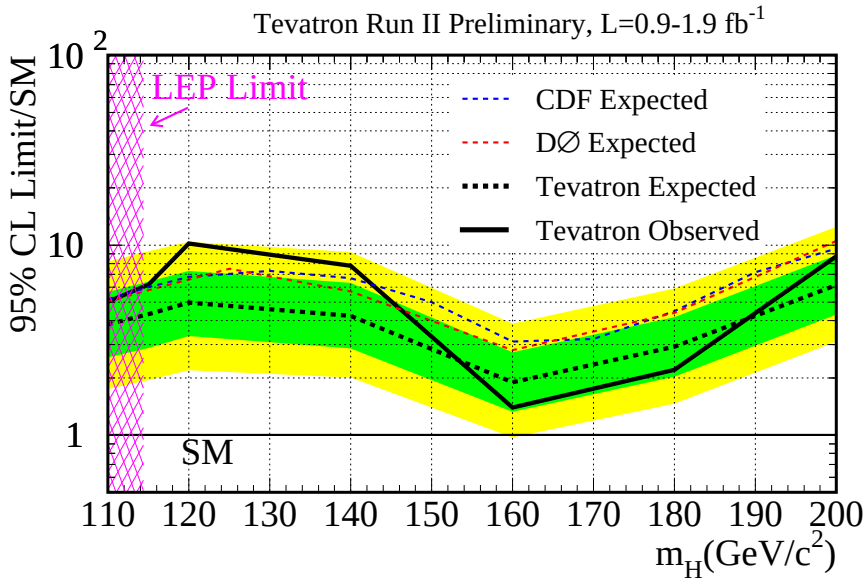

FIG. 3.23: Limites supérieures observées (trait plein) et attendues (trait en pointillés) à 95 % de niveau de confiance (obtenues avec l'hypothèse de bruit de fond sans signal) pour la section efficace de production d'un Higgs standard au Tevatron, rapportée à la prédiction du modèle standard et en fonction de la masse du Higgs [84]. Les bandes vertes et jaunes indiquent respectivement les régions à 68 et 95 % de probabilité dans lesquelles la limite supérieure peut fluctuer en l'absence de signal. Le Tevatron devient sensible à l'existence d'un boson de Higgs quand la limite supérieure atteint le taux de production prédit dans le cadre du modèle standard, à l'abscisse 1.

4

Sélection des paires de quarks top

Le collisionneur hadronique Tevatron est la seule source disponible de quarks top. Notre étude du couplage tWb est basée sur l'analyse des collisions $p\bar{p}$ ayant produit des paires de quarks top et enregistrées par le détecteur DØ auprès du Tevatron. Les données utilisées dans cette thèse sont celles enregistrées entre juillet 2002 et février 2006 au cours du *Run IIa* et correspondent à une luminosité intégrée de l'ordre de 1 fb^{-1}. A l'énergie du Tevatron, de nombreux processus sont produits en plus des paires de quarks top qui nous intéressent. Le taux de production d'une paire de quarks top par le Tevatron durant le *Run II* est de l'ordre de 3 paires par heure, à comparer à 7 millions par seconde pour l'ensemble des processus.
Ce chapitre présente les critères de sélection que nous avons mis au point pour extraire le signal recherché. Nous procédons dans un premier temps à une sélection des événements basée sur des coupures séquentielles sur des observables discriminantes ainsi que sur l'identification des jets de quark beau par le réseau de neurones présenté au chapitre précédent. Néanmoins, le lot sélectionné à l'issue de cette première étape est encore contaminé par du bruit de fond. Procéder à d'autres coupures séquentielles n'est plus efficace à ce stade de la sélection car le bruit de fond restant présente un état final visible dans le détecteur trop semblable au signal recherché. Nous avons alors optimisé le pouvoir discriminant d'un lot de variables en les combinant dans une nouvelle variable en utilisant le test de Neyman-Pearson. Nous montrons que le gain obtenu en coupant sur cette dernière variable n'est néanmoins pas très important. Mais son grand intérêt réside dans le fait que la distribution de cette variable nous permet d'estimer la composition en signal et bruit de fond du lot d'événements sélectionnés directement dans les données. Ceci permet de s'affranchir de l'incertitude importante qui découlerait

d'une estimation *via* les sections efficaces des processus de bruit de fond (qui sont mal connues), la luminosité et les efficacités de sélection déterminées dans la simulation.

4.1 Signal et bruit de fond

4.1.1 Le signal

Comme nous l'avons vu dans le chapitre 1, nous voulons différencier un fermion d'isospin $+\frac{1}{2}$ d'un fermion d'isospin $-\frac{1}{2}$ issus de la désintégration du W provenant de $t \to Wb$. Ceci est très simple si l'on considère la désintégration leptonique du W. Par ailleurs nous avons aussi vu que les quarks top sont produits majoritairement par paires. Nous pouvons donc utiliser les états finals $\ell^+\nu b\ell^-\bar{\nu}\bar{b}$ (dileptoniques) et $\ell\nu bq\bar{q}'\bar{b}$ (semi-leptoniques) des paires de quarks top. Nous avons choisi d'étudier le canal semi-leptonique ℓ+jets dont la cinématique est mieux connue car il y a moins de neutrinos dans l'état final. De plus, la statistique dans ce canal est plus importante. Comme lepton chargé, nous ne considérons que e et μ, ce qui correspond à 30 % des états finals de la paire $t\bar{t}$. La figure 4.1 illustre l'état final ℓ+jets d'une paire de quarks top : les deux quarks top se désintègrent chacun en un quark b et un boson réel W, l'un des bosons W donnant alors deux quarks, l'autre se désintégrant en un quark, un lepton chargé et son neutrino associé. Ainsi, ce type d'événement comporte toujours au moins quatre jets, un lepton isolé de haute impulsion transverse et un neutrino.

Le signal $t\bar{t}$ a été généré par ALPGEN avec une masse du top de 175 GeV/c^2 et les fractions standards d'hélicité du W issu de la désintégration du top (se référer à la section 2.5 pour plus d'information sur la simulation des données et les générateurs utilisés). Un autre lot avec un couplage $V + A$ pur a été simulé de la même manière. Notre mesure est basée sur la comparaison des données à la simulation par une méthode de vraisemblance. Dans ce cas, ces deux lots ne permettent que la mesure de f_+ en les combinant linéairement $\alpha \cdot (V - A) + (1 - \alpha) \cdot (V + A)$, car f_0 reste fixée à 0.7 quand α varie. Pour pouvoir mesurer f_0, des lots d'hélicité du W donnée ont été constitués dans DØ à partir du lot $V - A$. La procédure est décrite dans [85].
D'autres lots $t\bar{t}$ ont été utilisés pour l'étude des erreurs systématiques associées à la mesure de l'hélicité. Des lots avec une masse du top de 170 GeV/c^2 et de 180 GeV/c^2 ont permis d'étudier l'effet de la masse du top simulé sur la mesure. Une erreur systématique est également associée au modèle choisi pour simuler les désintégrations $t\bar{t}$. Un lot alternatif a été généré en utilisant PYTHIA, et un autre lot a été produit en changeant la paramétrisation de l'*underlying event* (lot "DW"). Les erreurs systématiques seront détaillées au chapitre 5.

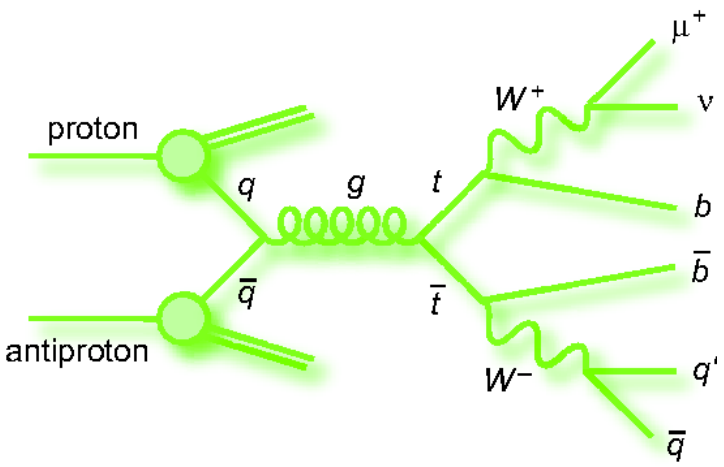

FIG. 4.1: Diagramme de Feynman dominant de production d'une paire de quarks top et de leur désintégration dans le canal semi-leptonique muon+jets. Ce diagramme correspond au signal recherché pour la mesure de l'hélicité du boson W. Un diagramme similaire est obtenu pour le canal électron+jets en remplaçant le muon par un électron et le neutrino muonique par un neutrino électronique.

4.1.2 Les processus de bruit de fond

Les processus physiques dont l'état final reconstruit a une topologie similaire à celle des événements $t\bar{t} \rightarrow \ell + jets$ et qui risquent donc d'être sélectionnés à la place de notre signal constituent un bruit de fond. Ces processus peuvent être de deux types. D'une part un bruit de fond physique constitué de processus qui ont un état final correspondant à celui recherché sans toutefois que ce dernier provienne d'une décroissance $t\bar{t}$. C'est le cas de la production d'un boson W associée à des jets. Et d'autre part, un bruit de fond instrumental dû à des processus multijets sélectionnés par le truchement de mauvaises reconstructions et identifications, qui simulent un faux lepton. La figure 4.2 illustre ces deux types de processus de bruit de fond pour notre analyse.

4.1.2.1 Le bruit de fond multijets QCD

Ce bruit de fond instrumental voit son origine dans l'identification par erreur d'un jet comme un lepton et une énergie transverse manquante mal mesurée. Un pion par exemple peut simuler un électron, s'il correspond à un jet reconstruit avec une grande fraction électromagnétique et associé à une faible multiplicité en particules chargées dont une seule sera reconstruite. Une conversion de photon en une paire e^+e^- dans la matière du détecteur dont un lepton n'est pas reconstruit, aussi. Un muon peut quant à lui provenir d'une décroissance semi-leptonique d'un quark b, d'un c, de pions ou de kaons simulant faussement un muon isolé. Ces erreurs de reconstruction et d'identifica-

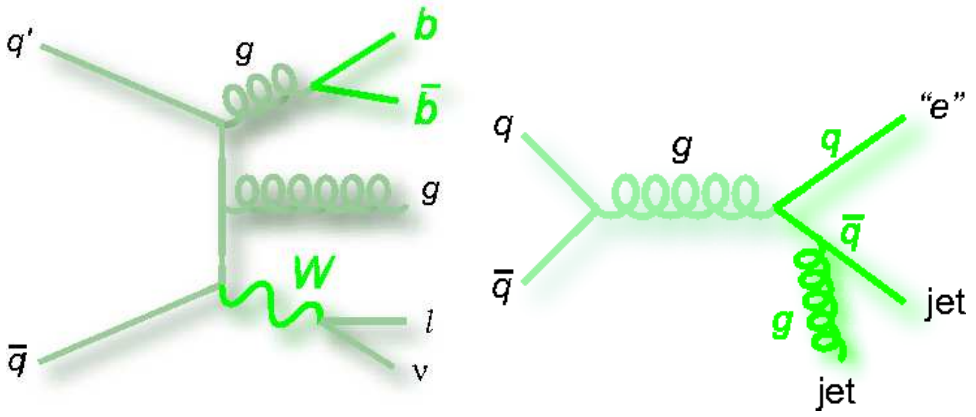

FIG. 4.2: Exemples de diagrammes de Feynman de processus de bruits de fond pour les analyses $t\bar{t} \to \ell + jets$. A gauche, un bruit de fond physique de production W+jets, à droite un bruit de fond instrumental de production QCD multijets.

tion ne sont pas fréquentes, mais les sections efficaces de ces processus multijets sont suffisamment importantes pour que ce bruit de fond ne soit pas négligeable. La sélection d'événements possédant des leptons de très bonne qualité ainsi que l'identification des jets beaux permet d'en rejeter une grande partie. Ce bruit de fond est estimé dans les données. En effet les événements de ce type subsistant après les sélections sont difficiles à simuler car il s'agit de queues de distribution.

4.1.2.2 Le bruit de fond W+jets

Ce bruit de fond regroupe plusieurs processus physiques, ayant tous en commun la production d'un boson W réel et la présence de jets associés. L'état final correspondant à ces processus est identique à celui du signal recherché et ces processus ont des sections efficaces telles qu'ils ne peuvent être négligés. Ce bruit de fond est généré par ALPGEN 2.05. Les multiplicités en jets sont controlées par une association parton-jets appelée *MLM matching scheme* [86], qui permet d'éviter le double comptage des radiations de gluons additionnels entre ALPGEN et PYTHIA. Les processus considérés sont les suivants :

- $Wjjjjj$: de 0 à 5 partons légers (u, d, s, gluon) accompagnent le W ;
- $Wc\bar{c}jjj$: le W est accompagné d'une paire $c\bar{c}$ et de 0 à 3 partons légers ;
- $Wb\bar{b}jjj$: le W est accompagné d'une paire $b\bar{b}$ et de 0 à 3 partons légers ou quarks c.

Les deux premiers types de bruit de fond seront fortement discriminés du signal par l'application de l'identification d'au moins un jet de quark b dans l'événement, en plus de la sélection d'événements à exactement 4 jets dans l'état final. Néanmoins, il reste toujours une probabilité de mal identifier un jet de quark c ou léger. Le dernier type de processus quant à lui représente un bruit de fond irréductible.

L'estimation de la contribution de ce bruit de fond à notre lot d'événements sélectionnés n'est pas faite à partir des sections efficaces de ces processus. En effet, les incertitudes théoriques qui leur sont associées sont très importantes (de l'ordre de 20 à 30 %). Les sections efficaces relatives entre tous les processus W+jets sont par contre mieux connues. Nous considérons ainsi ces processus, après renormalisations relatives, comme un seul bruit de fond. Puis nous estimerons sa contribution aux événements sélectionnés par une normalisation globale estimée directement dans les données (cf. section 4.3).
L'échelle de factorisation utilisée pour simuler ces événements est $Q^2 = M_W^2 + (P_T^W)^2$. Une erreur systématique sera calculée en utilisant d'autres lots de processus W+jets générés avec l'échelle de factorisation $Q^2 = M_W^2 + \sum (P_T^{jet})^2$.

4.1.2.3 Les autres bruits de fond

D'autres processus sont susceptibles de contribuer à notre lot de candidats $t\bar{t} \rightarrow \ell + jets$ sélectionnés : production du quark top par voie électrofaible (*single top*), processus Z+jets (mêmes diagrammes de production que les processus W+jets), paires de bosons vecteurs (dibosons[1] WW, WZ et ZZ), décroissance des paires $t\bar{t}$ en deux leptons. Néanmoins après application des sélections présentées par la suite, ces bruits de fond contribuent de façon totalement négligeable à notre lot sélectionné. Les processus Z+jets sont fortement éliminés par le véto sur le second lepton et le fait que ce sont des processus sans $\not\!\!E_T$ importante. Ne restent que les événements dont un lepton n'est pas reconstruit. En supposant que les processus Z+jets et W+jets ont les mêmes proportions en saveurs lourdes, la contribution Z+jets est prise en compte par le processus W+jets qui est normalisé aux données.

4.2 Sélection des événements

Dans cette section nous décrivons les critères utilisés pour des coupures séquentielles visant à sélectionner des candidats $t\bar{t}$ dans le canal ℓ+jets. Nous récapitulons en fin de section les efficacités de chaque coupure et le nombre de candidats sélectionnés dans les données du *Run IIa*.

4.2.1 Les données

Les données analysées correspondent à celles enregistrées par le détecteur DØ de juillet 2002 à février 2006. Elles couvrent l'ensemble de la période de prise de données appelée *Run IIa*. Les données ont subi une première phase de sélection préalable peu sévère, permettant de préparer un lot d'événements qui seront utilisés pour toutes les analyses de DØ en physique du top dans le canal ℓ+jets. Un lot pour le canal e+jets composé de

[1] Les processus dibosons ainsi que les processus de production de *single top* ont une section efficace de 2 à 3 pb^{-1}.

1346749 événements, et un lot pour le canal mu+jets composé de 234094 événements sont ainsi constitués. Les présélections appliquées sont les suivantes :

- l'événement doit comporter un jet d'impulsion transverse $P_T > 15$ GeV/c et de pseudo-rapidité $|\eta| < 2.5$;
- il doit comporter un objet électromagnétique *loose* de $P_T > 15$ GeV/c (canal e+jets) ;
- il doit comporter un muon, sans exigence d'isolation autre que $\Delta R(\mu, jet)$ (canal mu+jets, cf. section 2.6.5) ;
- l'énergie transverse manquante $\not\!\!E_T > 15$ GeV/c ;
- valider les critères de déclenchement correspondant au canal ;
- valider les critères de qualité.

La définition de ces sélections sera donnée par la suite.

4.2.2 Qualité des événements

La qualité des données enregistrées par DØ est vérifiée de manière à rejeter tous les événement pour lesquels les sous-détecteurs nécessaires n'étaient pas tous pleinement opérationnels, ou bien encore ceux présentant un bruit instrumental excessif dans le calorimètre. Dans le premier cas, ce sont des plages de prises de données complètes qui sont rejetées (*Runs*), lorsque le ou les sous-détecteurs défaillants sont nécessaires à notre analyse. Ils sont au nombre de quatre pour la mesure de l'hélicité : les deux détecteurs de traces et de vertex SMT et CFT, le calorimètre et le spectromètre à muons. En plus de la réjection complète de tels *Runs*, le fonctionnement des moniteurs de luminosité est vérifié. Les *Runs* sont découpés en blocs d'une minute de prise de données, appelés LBN (*Luminosity Bloc Number*). L'utilisation de ces petites structures permet de minimiser la quantité de bons événements rejetés. En cas de mauvais fonctionnement des luminomètres ou de mesures déviant trop par rapport à la luminosité moyenne, le LBN correspondant est rejeté.

Dans le cas d'un bruit dans le calorimètre, les événements sont rejetés au cas par cas. Les origines sont diverses, et dues soit à l'électronique du calorimètre lui-même, soit à celle des autres appareils électroniques l'entourant produisant des interférences. On distingue le "bruit cohérent" (ou *Coherent Noise*), la "caisse manquante" (ou *Empty Crate*), l'"anneau de feu" (ou *Ring of Fire*), le "bruit de midi" (ou *Noon Noise*) et enfin le "brouillard violet" (ou *Purple Haze*). Plus de détails et une explication de l'origine de ces bruits instrumentaux sont disponibles dans la référence [87].

Les *Runs* et LBN rejetés ne sont pas pris en compte pour le calcul de la luminosité intégrée totale (cf. tableaux 4.2 et 4.3). Par contre, dans le cas du bruit calorimétrique, les événements rejetés ne peuvent pas être exclus du calcul de la luminosité. Pour pallier à cela, l'efficacité correspondante a été determinée (97.14 ± 0.003 % des événements sont conservés), et la luminosité en est corrigée [88].

Concernant la simulation Monte-Carlo, lors de la génération des événements, certains ont été malheureusement dupliqués suite à une erreur logiciel des programmes de DØ. Une procédure de rejet de tels événements est appliquée au Monte-Carlo en amont de la sélection.

4.2.3 Critères de déclenchement

Ces critères de déclenchement, utilisés ici comme critères de tri, permettent une première réjection d'événements sans intêret pour les analyses de type ℓ+jets. On requiert un candidat lepton en plus d'un candidat jet avec une impulsion transverse supérieure à un certain seuil. Diverses périodes de prises de données sont considérées, correspondant à des versions successives des critères de déclenchement. Elles se distinguent par l'introduction de nouveaux critères de déclenchement en réponse à l'augmentation de la luminosité instantanée. Le tableau 4.1 liste les versions des listes de critères de déclenchement et la période de prise de données associée, pour les données du *Run IIa* utilisées dans cette thèse.

Version	Période	Runs
v8.0	juillet - octobre 2002	160582 - 167015
v9.0	octobre 2002 - Janvier 2003	167019 - 170246
v10.0	janvier - mars 2003	170247 - 174802
v11.0	mars - juillet 2003	174845 - 178721
v12.0	juillet 2003 - juin 2004	178722 - 194566
v13.0 - v13.2	juin - décembre 2004	194567 - 202024
v13.3 - v13.9	décembre 2004 - janvier 2005	202025 - 207680
v13.9 - v14.0 - v14.1	juin - juillet 2005	207719 - 208144
v14.2	juillet 2005	208168 - 208727
v14.3 - v14.7	juillet - octobre 2007	208728 - 211482
v14.8	novembre 2005 - février 2006	211513 - 215670

TAB. 4.1: Périodes de prise de données et version des listes de critères de déclenchement correspondantes.

4.2.3.1 Canal électron+jets

Le tableau 4.2 référence les critères de déclenchement appliqués et la luminosité correspondante. Cette analyse exploite ainsi, dans le canal e+jets, 1037.82 pb^{-1} de données après la selection opérée sur leur qualité.

Version des critères de déclenchement	Nom des critères de déclenchement	Luninosité délivrée (pb^{-1})	Luminosité enregistrée (pb^{-1})	Luminosité après critères de qualité (pb^{-1})
v8.0 - v12.0	EM15_2JT15	188.59	158.51	121.66
v12.0 - v13.0	E1_SHT15_2J20	272.64	250.76	227.14
v13.0 - v13.3	E1_SHT15_2J_J25	80.43	73.02	54.81
v13.3 - v14.0	E1_SHT15_2J_J30	353.70	324.69	294.27
v14.0 - v15.0	E1_SHT15_2J_J25	416.76	388.84	334.20
Total		1312.12	1195.82	1037.82

TAB. 4.2: Critères de déclenchement appliqués dans le canal e+jets selon la période de prise de données et luminosités intégrées délivrée, enregistrée, et après sélection sur la qualité des données.

Voici une description succinte des différents critères appliqués :

• Critère **EM15_2JT15**, qui exige aux différents niveaux de déclenchement :

- **niveau L1** : une tour calorimétrique d'énergie telle que $E_T > 10$ GeV/c et deux tours d'énergies telles que $E_T > 5$ GeV/c ;
- **niveau L2** : un objet électromagnétique tel que $E_T > 10$ GeV/c et sa fraction électromagnétique $Emf > 0.85$, et deux candidats jets tels que $E_T > 15$ GeV/c ;
- **niveau L3** : un candidat électron d'énergie telle que $E_T > 10$ GeV/c et deux candidats jets tels que $E_T > 15$ GeV/c.

• Critère **E1_SHT15_2J20**, qui exige aux différents niveaux de déclenchement :

- **niveau L1** : une tour calorimétrique d'énergie telle que $E_T > 11$ GeV/c ;
- **niveau L2** : aucune contrainte ;
- **niveau L3** : un candidat électron tel que $E_T > 15$ GeV/c et validant une contrainte sur la forme de la gerbe électromagnétique, et deux candidats jets tels que $E_T > 20$ GeV/c.

• Critère **E1_SHT15_2J_J25** : identique à E1_SHT15_2J20 à ceci près qu'au niveau L3 l'un des jets doit vérifier $E_T > 25$ GeV/c.

• Critère **E1_SHT15_2J_J30** : identique à E1_SHT15_2J20 à ceci près qu'au niveau L3 l'un des jets doit vérifier $E_T > 30$ GeV/c.

4.2.3.2 Canal muon+jets

Le tableau 4.3 référence les critères de déclenchement appliqués et la luminosité correspondante. Cette analyse exploite ainsi, dans le canal mu+jets, 996.27 pb^{-1} de données après la sélection opérée sur leur qualité.

Version des critères de déclenchement	Nom des critères de déclenchement	Luninosité délivrée (pb^{-1})	Luminosité enregistrée (pb^{-1})	Luminosité après critères de qualité (pb^{-1})
v8.0 - v12.0	MU_JT20_L2M0	191.11	163.41	126.32
v12.0 - v13.0	MU_JT25_L2M0	277.02	254.96	230.93
v13.0 - v13.2	MUJ2_JT25	55.74	39.47	31.43
v13.2 - v13.3	MUJ2_JT25_LM3	26.27	22.26	16.10
v13.3 - v14.0	MUJ2_JT30_LM3	382.01	277.12	252.17
v14.0 - v14.2	MUJ1_JT25_LM3	0.01	0.01	0.01
v14.2 - v14.3	MUJ1_JT25_ILM3	24.83	22.81	20.84
v14.3 - v15.0	MUJ1_JT35_LM3	392.19	366.28	312.55
Total		1349.19	1146.31	996.27

TAB. 4.3: Critères de déclenchement appliqués dans le canal mu+jets selon la période de prise de données et luminosités intégrées délivrée, enregistrée, et après sélection sur la qualité des données.

Voici la description succinte des différents critères appliqués :

• Critère **MU_JT20_L2M0**, qui exige aux différents niveaux de déclenchement :

- **niveau L1 :** un coup dans un des scintillateurs à muons et une tour du calorimètre telle que $E_T > 3$ GeV/c ;
- **niveau L2 :** un candidat muon et un candidat jet tel que $E_T > 10$ GeV/c ;
- **niveau L3 :** un candidat jet tel que $E_T > 20$ GeV/c.

• Critère **MU_JT25_L2M0** : identique à MU_JT20_L2M0 à ceci près qu'au niveau L3 le jet doit vérifier $E_T > 25$ GeV/c.

• Critère **MUJ2_JT25** :

- **niveau L1 :** un coup dans un des scintillateurs à muons ou une des chambres à fils et une tour du calorimètre telle que $E_T > 5$ GeV/c ;
- **niveau L2 :** un candidat muon et un candidat jet tel que $E_T > 8$ GeV/c ;

- **niveau L3 :** un candidat jet tel que $E_T > 25$ GeV/c.

• Critère **MUJ2_JT25_LM3** : identique à MUJ2_JT25 à ceci près qu'un candidat muon de $P_T > 3$ GeV/c est requis aux niveaux L1 et L3.

• Critère **MUJ2_JT30_LM3** : identique à MUJ2_JT25_LM3 à ceci près qu'au niveau L3 le jet doit vérifier $E_T > 30$ GeV/c.

• Critère **MUJ1_JT25_LM3** : identique à MUJ2_JT25_LM3.

• Critère **MUJ1_JT25_ILM3** :

- **niveau L1 :** un coup dans un des scintillateurs à muons ou une des chambres à fils et une tour du calorimètre telle que $E_T > 20$ GeV/c ;
- **niveau L2 :** un candidat muon de $P_T > 3$ GeV/c et un candidat jet tel que $E_T > 20$ GeV/c ;
- **niveau L3 :** un candidat muon de $P_T > 3$ GeV/c et un candidat jet tel que $E_T > 20$ GeV/c.

• Critère **MUJ1_JT35_LM3** : identique à MUJ2_JT25_LM3 à ceci près que le candidat jet doit vérifier $E_T > 35$ GeV/c au niveau L3.

4.2.3.3 Simulation du déclenchement

Les critères de déclenchement ne sont pas simulés événement par événement. De la même manière que pour l'identification des jets beaux dans la simulation (cf. chapitre 3), des paramétrisations des probabilités qu'un événement passe un critère donné sont établies dans les données réelles (ces courbes sont dites de *turn-on*). Ces courbes de calibration ont été établies dans les données réelles pour chaque critère et pour chaque niveau de déclenchement, et ceci pour chaque version de liste. A chaque objet (électron, muon, jet...) correspond une paramétrisation bien distincte. Leur combinaison fournit alors la probabilité qu'un événement simulé valide tel ou tel ensemble de critères, et cette probabilité servira de base à une repondération de l'événement.

4.2.4 Sélection du Vertex Primaire

Les critères de qualité requis pour sélectionner un vertex primaire sont les suivants :
– la position du vertex primaire le long du faisceau $|z_{PV}|$ doit être inférieure à 60 cm ;

– au moins trois traces doivent êtres associées au vertex primaire ;
– la distance le long du faisceau entre l'électron (ou le muon, selon le canal d'étude) et le vertex primaire doit être inférieure à 1 cm. Ceci permet de s'assurer que le lepton provient bien du vertex primaire.

4.2.5 Sélection des jets

La sélection des jets est identique dans les deux canaux e+jets et μ+jets. La topologie de nos événements comporte quatre quarks de haute impulsion transverse, mais le nombre de jets reconstruits peut être différent de ce chiffre, selon les gluons rayonnés dans l'état final, les jets supplémentaires assimilés à l'interaction, les jets reconstruits par erreur, les jets non reconstruits ou enfin les jets non sélectionnés. Néanmoins, nous avons utilisé les événements à exactement quatre jets pour reconstruire notre observable de mesure $\cos\theta$. Ces jets doivent tous les quatre avoir une impulsion transverse de plus de 20 GeV/c et le jet de plus haut P_T doit avoir une impulsion transverse de plus de 40 GeV/c. Enfin, tous les jets doivent être dans l'acceptance du détecteur, leur pseudo-rapidité doit vérifier $|\eta_{det}| < 2.5$. Ils ont été reconstruits avec l'algorithme de cône JCCB et un cône de $\Delta R < 0.5$.

4.2.6 Sélection des leptons

Une sélection est également faite sur les caractéristiques des leptons présents dans l'événement. Seuls sont retenus les événements comportant un lepton isolé unique et de haute impulsion transverse. Les critères varient selon les canaux. Certaines coupures, correspondant à des vétos sur la présence d'autres leptons, sont appliquées entre autres dans le but de rendre les analyses lepton+jets orthogonales aux analyses dans le canal dileptons, de manière à combiner éventuellement tous les états finals ensuite (hadronique + semileptonique + dileptonique). Ces vétos sont par ailleurs efficaces contre certains bruits de fond, comme les processus Z+jets.

Les critères de qualité des électrons et des muons ont été définis au chapitre 2 dans la section traitant de la reconstruction des objets.

4.2.6.1 Canal e+jets

Un électron de qualité *tight* et d'impulsion transverse $P_T > 20$ GeV/c doit être présent dans la partie centrale $|\eta_{det}| < 1.1$. Deux vétos sont ensuite appliqués. L'événement ne doit pas comporter d'autre électron de qualité *tight* (donc isolé), de $P_T > 15$ GeV/c et de pseudo-rapidité $|\eta_{det}| < 2.5$, ni de muon isolé de $P_T > 15$ GeV/c et $|\eta| < 2.0$.

4.2.6.2 Canal mu+jets

L'événement doit comporter un muon de qualité *tight* d'impulsion transverse $P_T > 20$ GeV/c et de pseudo-rapidité $|\eta_{det}| < 2.0$. Deux vétos sont ensuite appliqués, comme dans le canal électron. L'événement ne doit pas comporter d'autre muon de qualité *tight*, de $P_T > 20$ GeV/c et de pseudo-rapidité $|\eta_{det}| < 2.0$, ni d'électron isolé de qualité *tight* et de $P_T > 15$ GeV/c et $|\eta_{det}| < 2.5$. De plus, l'événement ne doit pas avoir un autre muon formant avec celui retenu une masse invariante comprise entre 70 GeV/c^2 et 110 GeV/c^2, caractéristique du bruit de fond Z+jets.

4.2.7 Energie transverse manquante

Les événements lepton+jets comportent dans leur état final un neutrino de grande impulsion transverse provenant de la désintégration leptonique du boson W. Cependant, le neutrino ne pouvant être détecté par le détecteur DØ, il n'est pas possible de mesurer directement son énergie. Celle-ci est estimée de la somme vectorielle des quadri-impulsions de tous les jets calorimétriques sélectionnés dans l'événement, prises après application des corrections sur l'énergie des jets et de l'électron. Lorsqu'il y a présence d'un muon cependant, celui-ci étant une particule au minimum d'ionisation, il dépose une quantité d'énergie très faible dans le calorimètre et l'impulsion de la trace qui lui est associée doit être soustraite de cette somme vectorielle. L'énergie transverse manquante correspondra au final au vecteur opposé à cette somme dans le plan transverse.
Finalement, afin de sélectionner des événements ayant produit un neutrino, une coupure sur l'énergie transverse manquante est appliquée : $\not\!\!E_T > 20$ GeV/c dans le canal e+jets, $\not\!\!E_T > 25$ GeV/c dans le canal mu+jets. La coupure est un peu plus sévère dans le canal mu+jets pour s'affranchir d'un accord imparfait à faible $\not\!\!E_T$entre les données et la simulation.

4.2.8 Coupures triangulaires

Lorsqu'un jet est mal reconstruit, de l'énergie manquante artificielle peut être créée avec une direction proche de celle du jet mal reconstruit. Ce jet mal reconstruit peut être identifié par erreur comme un lepton et ainsi être sélectionné comme événement lepton+jets. La collaboration DØ a optimisé des coupures pour les études du quark top, pour rejeter efficacement le bruit de fond dû à cet effet, en gardant au mieux le signal $t\bar{t}$. Plutôt que de couper sur $\Delta\Phi(lepton, \not\!\!E_T)$, il est plus efficace de couper sur $(\Delta\Phi(lepton, \not\!\!E_T), \not\!\!E_T)$ pour prendre en compte la corrélation entre ces deux variables.
Dans le canal e+jets, la coupure appliquée est la suivante :

$$\Delta\Phi(e, \not\!\!E_T) > 0.70\pi - 0.045 \cdot \not\!\!E_T$$

Et dans le canal mu+jets :

$$\Delta\Phi(\mu, \not\!\!E_T) > 0.67\pi - 0.011 \cdot \not\!\!E_T$$

Les coupures triangulaires tirent leur nom de la forme de la sélection dans le plan $(\Delta\Phi(lepton, \not\!\!E_T), \not\!\!E_T)$.

La figure 4.3 présente les distributions $(\Delta\Phi(lepton, \not\!\!E_T), \not\!\!E_T)$ d'événements sélectionnés dans le canal e+jets, issus de la simulation correspondant au signal $t\bar{t}$ et issus du bruit de fond QCD determiné dans les données, avant application de la coupure triangulaire. La droite en bleu sur les figures indique cette coupure. Ces événements ont déjà passé les critères de sélection précédents, en particulier la coupure sur l'énergie transverse manquante. On observe une accumulation du bruit de fond QCD en-deçà de la coupure choisie.

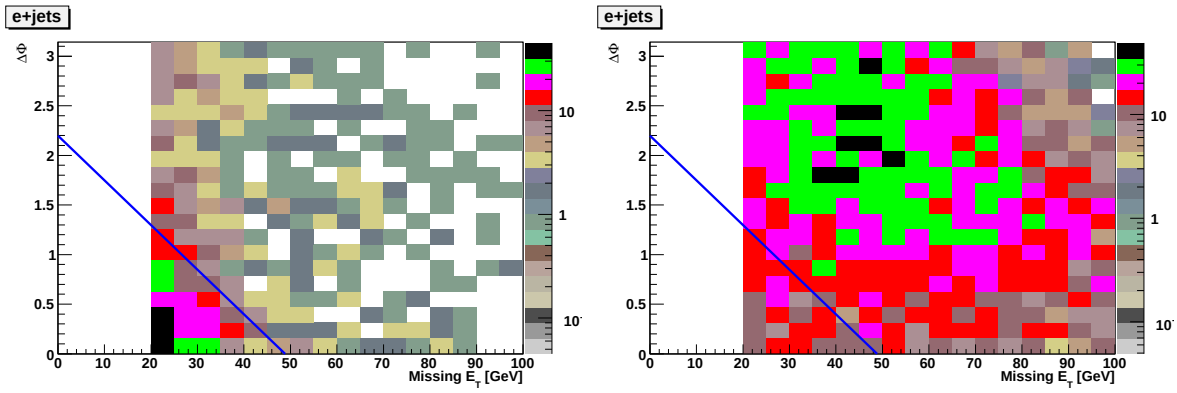

FIG. 4.3: Distributions $(\Delta\Phi(lepton, \not\!\!E_T), \not\!\!E_T)$ dans le canal e+jets pour des événements ayant passé la première partie des critères de sélection, en particulier la coupure sur l'énergie transverse manquante. La droite en bleu représente la coupure triangulaire appliquée. La figure de gauche correspond à la distribution obtenue dans le bruit de fond QCD issu des données et la figure de droite correspond à la distribution issue de la simulation $t\bar{t}$.

4.2.9 Identification des jets de quarks beaux

Les bruits de fond multijets et W+jets sont majoritairement composés de jets légers. Ils sont efficacement rejetés par une coupure sur la variable de sortie du réseau de neurones d'identification des jets beaux, tout en conservant le signal $t\bar{t} \to \ell + jets$ qui comporte deux jets beaux. Le tableau 4.4 illustre le nombre d'événements sélectionnés après coupure sur la variable de sortie du NN, pour les 6 points de fonctionnement étudiés. Ces événements comportent au moins un jet b identifié. Le choix du point de fonctionnement, c'est-à-dire de la coupure sur la variable de sortie du réseau de neurones, définissant l'efficacité et la pureté de la sélection du signal, est présentée dans la section 4.3.2.4. Ce choix a été fait de manière simultanée avec l'optimisation du discriminant qui permettra d'améliorer éventuellement la sélection et de connaître les proportions de signal et de bruit de fond. Nous verrons que les points de fonctionnement utilisés sont les

suivants : *L2* dans le canal e+jets et *Medium* dans le canal mu+jets. Peu d'événements présentent plus de deux jets identifiés comme b. Par exemple pour une coupure *Medium* appliquée dans le canal μ+jets, sur les 105 événements sélectionnés, 1 seul présente 3 jets identifiés comme issus de quarks b, 26 comportent 2 jets identifiés, et les 78 événements restants n'ont qu'un seul jet identifié par le NN.

Coupure sur le réseau de neurones	L4	L3	L2	Loose	Medium	Tight
Nombre d'événements, canal e+jets	162	151	132	119	100	87
Nombre d'événements, canal mu+jets	158	144	133	122	105	89

TAB. 4.4: Nombre final d'événements sélectionnés dans les données selon la coupure appliquée sur le réseau de neurones d'identification des jets de quarks beaux.

Le réseau de neurone permet d'identifier directement les jets de hadrons beaux dans les données. Dans la simulation par contre, les jets beaux ne sont pas directement identifiés par le réseau de neurones, mais un poids est attribué à chaque événement. Cette pondération traduit la probabilité que l'événement comporte un nombre donné de jets beaux identifiés. Elle est modélisée à partir des TRF correspondant au point de fonctionnement choisi (cf. section 3.1.7). Cependant, cette méthode ne permet pas de savoir quels jets de l'événement sont identifiés. Techniquement, l'événement est clôné autant de fois qu'il y a de combinaisons d'identification possibles (par exemple, 16 combinaisons de 0, 1, 2, 3 ou 4 jets beaux identifiés pour un événement à 4 jets). Chaque clône de l'événement est repondéré par sa probabilité. La probabilité $P_{jet\ \alpha}$ qu'un jet de saveur α (b, c ou légère) soit identifié est le produit de l'efficacité de *taggabilité* par la TRF correspondant à la saveur α :

$$P_{jet\ \alpha}(P_T, \eta) = P^{taggabilite}(P_T, \eta) \times TRF_\alpha(P_T, \eta)$$

De cette probabilité par jet on déduit la probabilité qu'un événement comporte au moins un jet identifié comme b :

$$P_{evenement}(n \geq 1) = 1 - \prod_{jets}(1 - P_{jet\ \alpha})$$

4.2.10 L'ajustement cinématique contraint

Pour procéder à la mesure de l'hélicité du boson W, nous allons étudier la distribution de l'angle θ entre le lepton chargé dans le référentiel de repos du W et le W dans le référentiel de repos du top. Cet angle est illustré sur la figure 4.4. La présence d'au moins un neutrino dans la chaîne de décroissance du quark top considéré, à laquelle s'ajoute la résolution sur les énergies et les impulsions reconstruites du jet et du lepton chargé, ne permet pas de mesurer directement l'angle θ à partir uniquement des quantités mesurées.

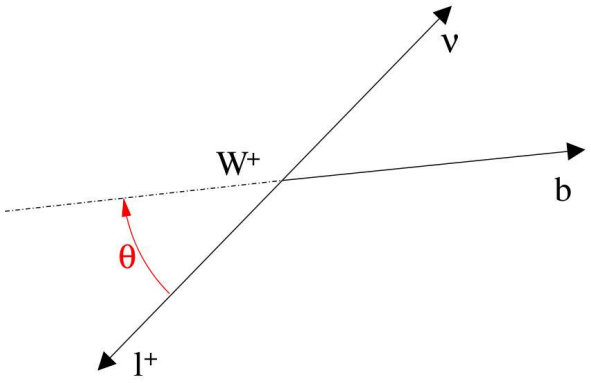

FIG. 4.4: Angle θ entre le lepton chargé dans le référentiel de repos du W et le W dans le référentiel de repos du quark top.

La résolution sur cet angle peut néanmoins être améliorée en utilisant les deux quarks top de l'événement et les contraintes cinématiques suivantes :

- l'impulsion transverse totale de l'état final doit être égale à 0 par conservation ;
- deux jets doivent correspondre aux produits de désintégration d'un W, le lepton et l'énergie manquante correspondent aux produits de désintégration de l'autre W. Les masses des deux W reconstruits sont contraintes à être égales à 80.4 GeV/c^2 (c'est-à-dire qu'on néglige la largeur du boson W) ;
- un W reconstruit combiné avec un jet doit correspondre à un quark top, l'autre W reconstruit combiné avec le jet restant doit correspondre à l'autre quark top. Les masses des deux quarks top sont contraintes à 172.6 GeV/c^2.

Les quadri-impulsions reconstruites du lepton chargé, des quatre jets et de l'énergie manquante sont modifées lors d'un tel ajustement contraint en utilisant une méthode de multiplicateurs de Lagrange. L'algorithme de l'ajustement contraint n'a pas été développé dans le cadre de cet thèse. La collaboration DØ utilise un algorithme mis au point à l'occasion de la mesure de la masse du quark top durant le *Run I* du Tevatron [89]. Dans la suite, il sera fait référence à cet algorithme d'ajustement contraint de DØ par son nom, qui est Hitfit. La modification des valeurs mesurées des quadri-impulsions se fait suivant la résolution correspondant à chaque objet mesuré, les directions étant généralement mieux mesurées que les modules. Nous utilisons par la suite ces quadri-impulsions modifiées.

Il y a 12 combinaisons possibles des quatre jets et du lepton chargé en deux W puis en deux top : chacun des quatre jets peut être le jet de quark beau du top se désintégrant leptoniquement, et du côté du top hadronique chacun des trois jets restants peut être le jet de quark b. Nous avons décidé de choisir une seule des 12 combinaisons plutôt que de les garder toutes en les pondérant, par souci de simplicité. Pour choisir la bonne com-

binaison parmi les 12, nous avons essayé deux méthodes. La première méthode consiste à ne garder que la solution ayant plus petit χ^2 de l'ajustement contraint. La deuxième méthode consiste à utiliser l'information issue de l'identification des jets de quarks beaux par le réseau de neurones pour réduire le nombre de combinaisons. Lorsqu'un seul jet est identifié, il reste 6 combinaisons possibles. Lorsque deux jets sont identifiés il n'y a plus que deux combinaisons possibles. On peut alors choisir parmi ces combinaisons restantes la combinaison de plus petit χ^2. Nous avons observé que la première méthode donnait de meilleurs résultats en terme de résolution sur l'angle θ, et c'est finalement ainsi que nous avons construit l'estimateur pour notre mesure.

Nous appliquons par ailleurs une sélection supplémentaire sur nos événements en rejetant tous ceux pour lesquels l'ajustement ne converge pas, empêchant la reconstruction cinématique de l'événement. Comme nous le verrons plus tard à la section 4.3.2, le χ^2 de l'ajustement contraint ($HitfitChi2$) sera de plus utilisé pour discriminer le signal $t\bar{t}$ du bruit de fond W+jets.

La contrainte correspondant à $M_{\mathrm{top},1} = M_{\mathrm{top},2} = 172.6$ GeV/c^2 améliore d'environ 10 % la résolution sur l'angle θ reconstruit et augmente de 55 % à 65 % le choix de la bonne combinaison de jets.

4.2.11 Bilan de la sélection

Les tableaux 4.5 et 4.6 récapitulent, respectivement pour le canal e+jets et le canal mu+jets, le nombre d'événements sélectionnés après application de chacun des critères que nous avons présentés, ainsi que l'efficacité correspondante. Ces nombres ne serviront pas à la mesure et sont donnés à titre d'illustration. Les lots de données utilisés ont déjà subi une sélection préalable des jets, des leptons et des critères de déclenchement (cf. section 4.2.1).

4.3 Détermination des contributions du signal $t\bar{t}$ et des bruits de fond

Cette section est consacrée à la détermination des contributions de signal $t\bar{t} \to \ell + jets$, de W+jets et de production multijets QCD aux événements que nous avons sélectionnés avec les coupures séquentielles présentées précédemment. Nous avons donc trois fractions, soit deux inconnues à déterminer. Pour ce faire, nous allons procéder en deux étapes. Tout d'abord, une première méthode nommée *Matrix Method* va nous permettre de déterminer la fraction d'événements QCD présents dans nos données. Ensuite, connaissant la fraction d'événements QCD, l'ajustement de la distribution d'une fonction discriminante permettra d'extraire la fraction d'événements W+jets.

Coupure	Evén.	Efficacité (%)	Eff. cumulée (%)
Lot initial e+jets	1346749		
Identification des jets	1329357	98.71 ± 0.01	98.71 ± 0.01
Electron associé à un jet	1319659	99.27 ± 0.01	97.99 ± 0.01
Exactement 4 jets	3986	0.30 ± 0.004	0.30 ± 0.005
Jet de plus haut $P_T > 40$ GeV/c	3345	83.92 ± 0.58	0.25 ± 0.004
Electron $P_T > 20.0$ GeV	2710	81.02 ± 0.68	0.20 ± 0.004
Premier veto sur muons	2609	96.27 ± 0.36	0.19 ± 0.004
Au plus 1 électron	2595	99.46 ± 0.14	0.19 ± 0.004
$\not\!\!E_T \geq$ 20 GeV	1075	41.43 ± 0.97	0.08 ± 0.002
Coupure triangulaire	602	56.00 ± 1.51	0.045 ± 0.002
Exactement 1 électron *tight* et veto sur second électron et sur muon	312	51.83 ± 2.04	0.023 ± 0.001
Identification d'un b par NN L2	133	43.89±2.85	0.010 ± 0.001
Convergence de *Hitfit*	132	99.25 ± 0.75	0.010 ± 0.001

TAB. 4.5: Coupures séquentielles appliquées pour sélectionner le signal $t\bar{t}$ dans les données pour le canal *e+jets*. Le nombre d'événements sélectionnés est indiqué, ainsi que les efficacités correspondantes.

4.3.1 Contribution du bruit de fond QCD multijets

Afin de déterminer la proportion du bruit de fond multijets QCD, une méthode dite *Matrix Method* est mise en œuvre [90]. Cette technique est utilisée dans plusieurs analyses de DØ (physique du top et recherche du boson de Higgs) pour évaluer la contribution de ce bruit de fond à partir des données uniquement.

4.3.1.1 Principe

Pour déterminer la proportion de bruit de fond QCD, nous utilisons le fait que ce bruit de fond ne comporte pas de lepton isolé. Les nombres de faux et de vrais leptons isolés, correspondant respectivement au bruit de fond QCD et au signal de type W+jets (ici nous appelons signal les topologies de type W+jets, incluant $t\bar{t}$ et *single top*, qui comportent un vrai lepton isolé), peuvent être déterminés dans les données en relâchant une coupure de qualité du lepton. Nous exploitons le fait que la sélection suivant le critère d'isolation a une efficacité différente sur des vrais et des faux leptons. Deux ensembles d'événements sont utilisés. Le premier est appelé *Loose*, il contient le second qui est dit *Tight*. L'ensemble *Tight* est celui qui servira à mesurer l'hélicité du W. Le passage du lot *Tight* au lot *Loose* est fait en relâchant le critère d'isolation du lepton. Dans le canal e+jets, ce critère correspond à la coupure sur la fonction de vraisemblance, tandis que dans le canal mu+jets il correspond à la sélection sur les halos de traces et le halo ca-

Coupure	Evén.	Efficacité (%)	Eff. cumulée (%)
Lot initial mu+jets	234094		
Identification des jets	229605	98.08 ± 0.03	98.08 ± 0.03
Muon associé à un jet	229491	99.95 ± 0.01	98.03 ± 0.03
Exactement 4 jets	1631	0.71 ± 0.02	0.70 ± 0.017
Jet de plus haut $P_T > 40$ GeV/c	1538	94.30 ± 0.57	0.66 ± 0.017
Muon $P_T > 20.0$ GeV	856	55.66 ± 1.27	0.37 ± 0.012
Au moins 1 muon *tight*	852	99.53 ± 0.23	0.36 ± 0.012
Critères de déclenchement	834	97.89 ± 0.49	0.36 ± 0.012
Au plus 1 muon	821	98.44 *pm* 0.43	0.35 ± 0.012
Premier veto sur électrons	819	99.76 ± 0.17	0.35 ± 0.012
$\not\!\!E_T \geq$ 25 GeV	470	57.39 ± 1.73	0.20 ± 0.009
Coupure triangulaire	370	78.72 ± 1.89	0.16 ± 0.008
Exactement 1 muon *tight* et veto sur second muon et sur électron	287	77.57 ± 2.17	0.12 ± 0.007
Identification d'un b par NN *Medium*	106	36.93 ± 2.85	0.045 ± 0.004
Convergence de *Hitfit*	105	99.06 ± 0.94	0.045 ± 0.004

TAB. 4.6: Coupures séquentielles appliquées pour sélectionner le signal $t\bar{t}$ dans les données pour le canal *mu+jets*. Le nombre d'événements sélectionnés est indiqué, ainsi que les efficacités correspondantes.

lorimètrique (cf. sections 2.6.4.2 et 2.6.5.2 pour la définition de ces variables d'isolation leptonique).

Le nombre total d'événements de chaque lot est respectivement noté N^{Loose} pour le lot *Loose* et N^{Tight} pour le lot *Tight*. Le nombre d'événements de type signal dans chacun de ces deux lots est noté respectivement N_{sig}^{Loose} et N_{sig}^{Tight}, le nombre d'événements de type QCD N_{qcd}^{Loose} et N_{qcd}^{Tight}. Nous avons ainsi le système d'équations suivant :

$$N^{Loose} = N_{sig}^{Loose} + N_{qcd}^{Loose}$$

$$N^{Tight} = N_{sig}^{Tight} + N_{qcd}^{Tight}$$

$$\text{avec} \quad N_{sig}^{Tight} = N_{sig}^{Loose} \cdot \epsilon_{sig} \quad \text{et} \quad N_{qcd}^{Tight} = N_{qcd}^{Loose} \cdot \epsilon_{qcd} \quad \text{soit :}$$

$$N^{Tight} = N_{sig}^{Loose} \cdot \epsilon_{sig} + N_{qcd}^{Loose} \cdot \epsilon_{qcd}$$

où ϵ_{sig} et ϵ_{qcd} sont respectivement les efficacités de sélection des vrais et des faux leptons suivant le critère d'isolation du lepton.
La résolution de ce système de deux équations à deux inconnues, N_{sig}^{Tight} et N_{qcd}^{Tight}, nous

conduit aux contributions suivantes :

$$N_{sig}^{Tight} = \epsilon_{sig} \cdot \frac{N^{Tight} - \epsilon_{qcd} \cdot N^{Loose}}{\epsilon_{sig} - \epsilon_{qcd}}$$

$$N_{qcd}^{Tight} = \epsilon_{qcd} \cdot \frac{\epsilon_{sig} \cdot N^{Loose} - N^{Tight}}{\epsilon_{sig} - \epsilon_{qcd}}$$

qui correspondent respectivement aux contributions de type signal W+jets et de type bruit de fond QCD dans le lot d'événements sélectionnés.

4.3.1.2 Calcul des erreurs associées

En introduisant les variables suivantes, qui assurent que les lots soient statistiquement indépendants :

$$N_1 = N^{Loose} - N^{Tight}$$

$$N_2 = N^{Tight}$$

nous pouvons déterminer les erreurs correspondantes :

$$\Delta N_{sig}^{T} = \sqrt{\left(\frac{dN_{sig}^{T}}{d\epsilon_{sig}}\right)^2 \cdot \Delta\epsilon_{sig}^2 + \left(\frac{dN_{sig}^{T}}{d\epsilon_{qcd}}\right)^2 \cdot \Delta\epsilon_{qcd}^2 + \left(\frac{dN_{sig}^{T}}{dN_1}\right)^2 \cdot \Delta N_1^2 + \left(\frac{dN_{sig}^{T}}{dN_2}\right)^2 \cdot \Delta N_2^2}$$

$$\Delta N_{qcd}^{T} = \sqrt{\left(\frac{dN_{qcd}^{T}}{d\epsilon_{sig}}\right)^2 \cdot \Delta\epsilon_{sig}^2 + \left(\frac{dN_{qcd}^{T}}{d\epsilon_{qcd}}\right)^2 \cdot \Delta\epsilon_{qcd}^2 + \left(\frac{dN_{qcd}^{T}}{dN_1}\right)^2 \cdot \Delta N_1^2 + \left(\frac{dN_{qcd}^{T}}{dN_2}\right)^2 \cdot \Delta N_2^2}$$

$\Delta\epsilon_{sig}$ et $\Delta\epsilon_{qcd}$ sont mesurées, et $\Delta N_1 = \sqrt{N_1}$ et $\Delta N_2 = \sqrt{N_2}$ de par la statistique élevée.

4.3.1.3 Détermination des efficacités ϵ_{sig} et ϵ_{qcd}

La mesure de l'efficacité ϵ_{sig} est realisée à partir des lots de simulation $t\bar{t}$ et W+jets, corrigée par un facteur de correction de la simulation.

L'efficacité ϵ_{sig} ne dépend pas, au premier ordre, de la composition en saveur des événements W+jets et peut ainsi être utilisée avant ou après enrichissement en jets beaux. L'efficacité ϵ_{qcd} est elle déterminée directement dans les données, en faisant l'hypothèse que la région à basse énergie transverse manquante ($\not\!\!E_T < 10$ GeV) est dominée par le bruit de fond QCD. Avec ces événements, ϵ_{qcd} est mesurée comme étant le rapport entre le nombre d'événements *Tight* sur le nombre d'événements *Loose* [91].

De plus amples détails, en particulier sur la dépendance de ces efficacités en la multiplicité de jets suivant les critères de déclenchement appliqués ainsi que sur le calcul des erreurs systématiques associées, sont disponibles dans [92]. Le tableau 4.7 récapitule les valeurs des efficacités utilisées dans cette analyse, ainsi que leurs erreurs statistiques et systématiques.

Canal	ϵ_{sig}	ϵ_{qcd}
e+jets	0.838 ± 0.010 (stat.) ± 0.022 (syst.)	0.18 ± 0.01 (stat.) ± 0.03 (syst.)
mu+jets	0.849 ± 0.010 (stat.) ± 0.020 (syst.)	0.24 ± 0.01 (stat.) ± 0.03 (syst.)

TAB. 4.7: Efficacités de sélection d'un lepton isolé dans les contributions de type signal W+jets et de type bruit de fond QCD dans les canaux e+jets et mu+jets.

4.3.1.4 Mesure de la proportion de QCD

Le tableau 4.8 présente les contributions mesurées dans notre lot de données selectionnées après application de la coupure *L2* (resp. *Medium*) du réseau de neurones dans le canal e+jets (resp. mu+jets).

Canal	N^{Loose}	N^{Tight}	N_{sig}^{Tight}	N_{qcd}^{Tight}
e+jets	205	132	121.12 ± 12.16	10.88 ± 2.02
mu+jets	121	105	105.89 ± 10.94	-0.89 ± 1.47

TAB. 4.8: Contributions de type signal W+jets et de type bruit de fond QCD au lot de données selectionnées, pour une coupure sur le réseau de neurones *L2* dans le canal canal e+jets et *Medium* dans le canal mu+jets.

La contribution négative de QCD est compatible avec 0, et nous la fixerons à 0 par la suite. Des résultats analogues ont été obtenus dans les autres mesures effectuées en physique du top dans DØ avec les données du *Run IIa*, bien que les critères de sélection ne soient pas exactement les mêmes que ceux utilisés dans notre analyse.

4.3.2 Fonction discriminante

Connaissant le nombre d'événements QCD dans nos données, nous pouvons maintenant déterminer la contribution W+jets aux événements que nous avons sélectionnés. Pour ce faire, nous avons mis au point une fonction de vraisemblance discriminante qui fournira ces fractions relatives manquantes, en plus de permettre une réjection supplémentaire du bruit de fond.

4.3.2.1 Définition du discriminant

Le test de Neyman-Pearson permet de tester de manière optimale si un candidat sélectionné appartient à une certaine catégorie (hypothèse H_0) contre le fait qu'il n'appartienne pas à cette catégorie (hypothèse H_1 alternative). Dans le cas d'un événement sélectionné caractérisé par un ensemble de variables $\overrightarrow{x}$, ce test revient à couper sur le

rapport défini de la manière suivante :

$$\mathcal{L}(\overrightarrow{x}, H_0, H_1) = \frac{P(\overrightarrow{x}|H_0)}{P(\overrightarrow{x}|H_1)} = \frac{P_S(\overrightarrow{x})}{P_B(\overrightarrow{x})}$$

où $P_S(\overrightarrow{x})$ et $P_B(\overrightarrow{x})$ correspondent respectivement aux densités de probabilité de $\overrightarrow{x}$ pour le signal S et le bruit de fond B. Il a été montré qu'il s'agit du test optimal si le jeu de variables $\overrightarrow{x}$ est complet, c'est-à-dire contient toute l'information pertinente pour différencier le signal du bruit de fond (en particulier, les variables ne doivent pas être corrélées).
Il est pratique dans notre cas de nous ramener à une grandeur comprise entre 0 et 1, ce que l'on fait par une transformation monotone en utilisant le rapport suivant plutôt que $\mathcal{L}(\overrightarrow{x}, H_0, H_1)$:

$$\frac{\mathcal{L}(\overrightarrow{x}, S, B)}{\mathcal{L}(\overrightarrow{x}, S, B) + 1}$$

Ce rapport peut être réécrit comme suit :

$$\frac{\mathcal{L}(\overrightarrow{x}, S, B)}{\mathcal{L}(\overrightarrow{x}, S, B) + 1} = \frac{P_S(\overrightarrow{x})}{P_S(\overrightarrow{x}) + P_B(\overrightarrow{x})} = P(S|\overrightarrow{x})$$

La dernière égalité correspond à une interprétation bayesienne du test de Neyman-Pearson.

Pour tenter de sélectionner encore mieux les paires de quarks top, nous construisons ainsi un discriminant $D_S(\overrightarrow{x})$ défini comme :

$$D_S(\overrightarrow{x}) = P(S|\overrightarrow{x}) = \frac{P_S(\overrightarrow{x})}{P_S(\overrightarrow{x}) + P_B(\overrightarrow{x})}$$

Les événements compatibles avec l'hypothèse signal auront des valeurs du discriminant voisines de 1 tandis que ceux compatibles avec l'hypothèse bruit de fond auront des valeurs tendant vers 0. Si l'on suppose que les n variables x_i sont décorrélées, les densités de probabilité $P_S(\overrightarrow{x})$ et $P_B(\overrightarrow{x})$ à n dimensions se factorisent en un produit de n densités de probabilité à une dimension, par exemple pour le signal :

$$P_S(\overrightarrow{x}) = \prod_{i=1}^{n} P_S(\overrightarrow{x}_i)$$

Dans le cas de variables corrélées, il n'est pas incorrect d'utiliser cette factorisation, néanmoins le test n'est plus optimal, le pouvoir séparateur entre le signal et le bruit de fond est réduit.

Dans notre analyse, le signal correspond aux paires de quark top. Quant au bruit de fond, il en reste deux sources après la première étape de sélection : des événements de

type multijets (QCD) et des événements W+jets. Le bruit de fond W+jets est largement majoritaire et nous l'utilisons seul pour construire notre fonction discriminante. Dans ce cas la densité de probabilité du bruit de fond s'écrit :

$$P_B(\overrightarrow{x}) = P_{Wnjets}(\overrightarrow{x})$$

4.3.2.2 Variables étudiées pour construire le discriminant

Dans la pratique nous construisons notre discriminant ainsi :

$$D_S(\overrightarrow{x}) = \frac{\exp(\sum_{i=1}^{n}(\ln \frac{P_S(x_i)}{P_B(x_i)})_{ajust.})}{\exp(\sum_{i=1}^{n}(\ln \frac{P_S(x_i)}{P_B(x_i)})_{ajust.}) + 1}$$

Nous transformons les variables x_i par des fonctions arbitraires de manière à ce que la distribution $\ln \frac{P_S(x_i)}{P_B(x_i)}$ soit aisément ajustable par une fonction polynômiale.
Le choix du lot de variables utilisées est basé sur un certain nombre de critères. Premièrement, le pouvoir discriminant entre le signal $t\bar{t}$ et le bruit de fond doit être intéressant. Ensuite, les variables doivent être peu corrélées entre elles ainsi que vis-à-vis de l'observable $\cos\theta$ qui permet de mesurer les fractions d'hélicité du W. Enfin, la distribution simulée de cette variable doit décrire correctement celle observée dans les données.

Un grand nombre de variables topologiques et cinématiques ont été étudiées au départ. Après les avoir scrutées suivant les trois critères que nous venons de donner, nous nous sommes restreints aux sept variables dont la définition est donnée ci-dessous.

- H_T :

Cette variable correspond à la somme scalaire des impulsions transverses des quatre jets de l'événement :

$$H_T = \sum_{N_{jet}=1}^{N_{jet}=4} P_T(jet)$$

La variable H_T donne une mesure de l'énergie hadronique transverse de l'événement, qui sera plus élevée pour des événements contenant des quarks top à cause de sa masse élevée.

- **Aplanarité :**

On définit dans un premier temps le tenseur $\mathcal{M}$ de la façon suivante :

$$\mathcal{M}_{ij} = \frac{\sum_o p_i^o p_j^o}{\sum_o |\vec{p^o}|^2}$$

où $\vec{p^o}$ est la quadri-impulsion d'un objet reconstruit o, et i et j correspondent aux coordonnées cartésiennes. Par objet, on entend les jets et le lepton provenant de la

décroissance du W. On note λ_1, λ_2 et λ_3 les trois valeurs propres déterminées en diagonalisant cette matrice ($\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$ et $\lambda_1 + \lambda_2 + \lambda_3 = 1$).
L'aplanarité est alors définie comme suit :

$$\mathcal{A} = \frac{3}{2}\lambda_3$$

Des événements sphériques se traduisent par des valeurs élevées de $\mathcal{A}$, des événements plutôt plans par de faibles valeurs de l'aplanarité. Les événements de type $t\bar{t}$ étant isotropes, ils contribueront aux fortes valeurs de $\mathcal{A}$, contrairement aux événements de type W+jets et QCD, car les jets correspondants à ces événements proviennent essentiellement des radiations dans l'état initial et dans l'état final.

- **Centralité :**

Cette variable est définie comme le rapport de la somme H_T des impulsions transverses des quatre jets sur la somme H_E des énergies de ces jets :

$$\mathcal{C} = \frac{H_T}{H_E}$$

Les quarks top sont produits quasiment au repos et leur désintégration produit des événements sphériques avec une plus grande énergie transverse que les événements de bruit de fond.

- $Kt'_{min}Red$ **:**

Cette variable correspond à la distance minimale dans l'espace $\eta-\phi$ entre deux jets parmi toutes les combinaisons possibles de deux jets, multipliée par l'impulsion transverse minimale parmi ces deux jets :

$$K'_{min}Red = \Delta R_{jj}^{min} \cdot P_T^{min}$$

- Kt'_{min} **:**

Cette variable est définie par :

$$Kt'_{min} = \frac{Kt'_{min}Red}{p_T^W}$$

où p_T^W est l'impulsion transverse du boson W reconstruit par l'ajustement cinématique contraint.

- $MinDijetMass$ **:**

Cette variable correspond à la plus petite masse invariante entre deux jets parmi toutes les combinaisons de jets de l'événement. Dans le cas d'événements $t\bar{t}$ cette variable correspond généralement à la masse du W se désintégrant en paire $t\bar{t}$, et est plus élevée que

dans le cas du bruit de fond où les jets proviennent des radiations dans l'état initial et final.

- $HitfitChi2$:

Cette variable a été définie dans la section 4.2.10. Elle correspond au χ^2 de l'ajustement cinématique contraint visant à reconstruire l'événement en associant les jets aux deux bosons W et aux deux quarks top.

Chaque variable est transformée par une fonction bijective dans le domaine de définition de la variable, de manière à ajuster aisément par une fonction polynômiale le logarithme du rapport de la distribution de la variable déterminée dans le signal sur la distribution de la variable issue du bruit de fond W+jets (cf. la définition du discriminant D_S, section 4.3.2.2). Le bruit de fond QCD, minoritaire, n'est pas pris en compte pour construire la vraisemblance discriminante. Les fonctions de transformation utilisées sont définies dans le tableau 4.9, elles sont inspirées d'analyses précédentes effectuées dans DØ.

Variable	Variable transformée
H_T	$\ln(H_T)$
$\mathcal{A}$	$\exp(-11.0 \cdot \mathcal{A})$
$\mathcal{C}$	$\ln(\mathcal{C})$
$Kt'_{min}Red$	$Kt'_{min}Red$
Kt'_{min}	$\ln(Kt'_{min})$
$MinDijetMass$	$\sqrt{MinDijetMass}$
$HitfitChi2$	$HitfitChi2$

TAB. 4.9: Transformation des sept variables étudiées pour construire le discriminant.

4.3.2.3 Choix des variables du discriminant

Les figures 4.5, 4.6, 4.7 et 4.8 comparent les distributions des variables transformées étudiées pour former le discriminant. La comparaison est effectuée entre les distributions simulées du signal $t\bar{t}$ et du bruit de fond W+jets.

Nous avons vérifié les corrélations entre ces variables pour éviter d'utiliser une variable n'ajoutant pas d'information. Ces corrélations sont données dans le tableau 4.10.
Nous avons aussi vérifié la corrélation entre chacune de ces variables et l'observable $\cos\theta$ qui servira à mesurer l'hélicité du W (cf. tableaux 4.11 et 4.12). Couper sur une variable corrélée à l'observable mesurée biaise la distribution de cette observable. Comme notre

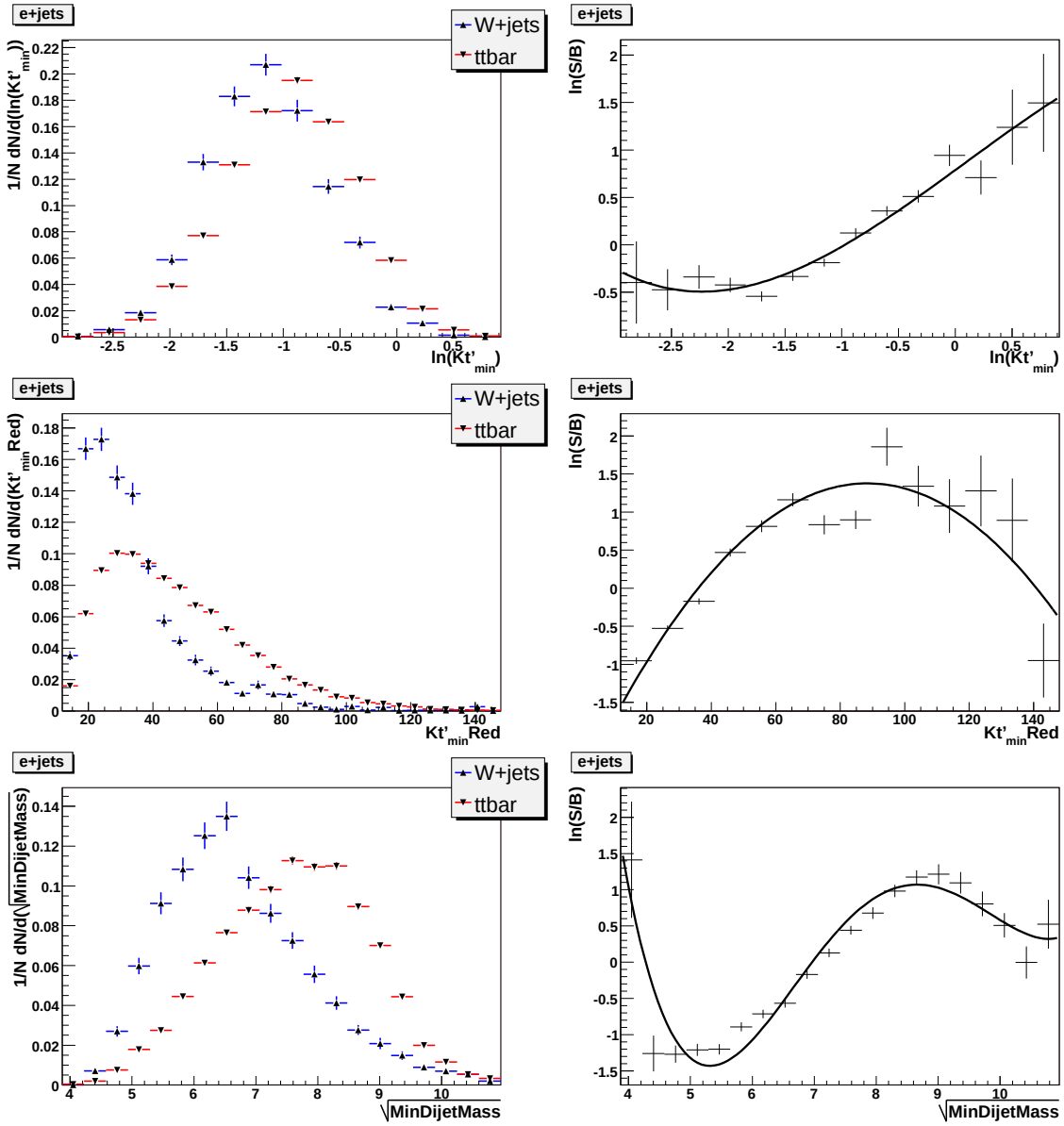

FIG. 4.5: Canal e+jets.
Figures de gauche : Distributions normalisées simulées pour le signal $t\bar{t}$ (triangles inversés, distribution rouge) et pour le bruit de fond W+jets (triangles droits, distribution bleue) des variables étudiées pour former la fonction discriminante. De haut en bas : Kt'_{min}, $Kt'_{min}Red$ et $MinDijetMass$.
Figures de droite : logarithme du rapport signal sur bruit de fond des distributions normalisées et ajustement par un polynôme.

mesure est basée sur la comparaison entre données et simulation, un tel biais n'a pas de conséquences sur la mesure si la simulation décrit correctement les données. Dans DØ, de nombreuses corrections sont appliquées à la simulation pour l'améliorer. Néanmoins il est plus prudent de choisir des variables peu corrélées à $\cos\theta$ pour former le discriminant. C'est le cas des sept variables que nous avons retenues pour notre étude, les variables corrélées, non définies dans ce chapitre, ont été rejetées dès le départ de l'étude.

e \ mu	$\mathcal{A}$	$\mathcal{C}$	H_T	Kt'_{min}	$HitfitChi2$	$MinDijetMass$	$Kt'_{min}Red$
$\mathcal{A}$	1	0.21	-0.16	0.17	-0.11	0.1	0.16
$\mathcal{C}$	0.18	1	0.32	-0.068	0.063	-0.021	-0.032
H_T	-0.17	0.36	1	0.11	0.13	0.46	0.26
Kt'_{min}	0.13	-0.098	0.09	1	-0.18	0.6	0.78
$HitfitChi2$	-0.072	0.077	0.11	-0.13	1	-0.042	-0.023
$MinDijetMass$	0.1	-0.013	0.46	0.61	-0.077	1	0.72
$Kt'_{min}Red$	0.14	-0.037	0.24	0.8	-0.048	0.72	1

TAB. 4.10: Corrélations entre les variables étudiées pour former le discriminant, déterminées dans la simulation $t\bar{t}$. La zone supérieure droite correspond au canal mu+jets, la zone inférieure gauche, au canal e+jets. Les corrélations déterminées dans les deux canaux montrent un bon accord.

e+jets	$\mathcal{A}$	$\mathcal{C}$	H_T	Kt'_{min}	$HitfitChi2$	$MinDijetMass$	$Kt'_{min}Red$
$\cos\theta$	0.018	-0.032	0.091	-0.036	0.02	0.067	0.038

TAB. 4.11: Corrélations entre l'observable $\cos\theta$ et les variables étudiées pour former le discriminant, déterminées dans la simulation $t\bar{t}$ (canal e+jets).

mu+jets	$\mathcal{A}$	$\mathcal{C}$	H_T	Kt'_{min}	$HitfitChi2$	$MinDijetMass$	$Kt'_{min}Red$
$\cos\theta$	-0.023	-0.026	0.066	-0.077	0.14	0.038	0.015

TAB. 4.12: Corrélations entre l'observable $\cos\theta$ et les variables étudiées pour former le discriminant, déterminées dans la simulation $t\bar{t}$ (canal mu+jets).

A partir des sept variables retenues, nous avons construit autant de fonctions discriminantes qu'il y a de combinaisons possibles de ces variables. Pour choisir une combinaison nous avons comparé les courbes de l'efficacité de sélection du signal $t\bar{t}$ en fonction de

l'efficacité de sélection du bruit de fond W+jets. Ces courbes sont obtenues en variant la coupure sur le discriminant. La figure 4.9 présente les courbes obtenues pour quatre combinaisons différentes de trois à cinq variables, pour les canaux e+jets et mu+jets. La combinaison finalement retenue, identique pour les deux canaux, est constituée des cinq variables *Aplanarité*, *Centralité*, *HitfitChi2*, H_T et $Kt'_{min}Red$. Les résultats obtenus sont similaires quel que soit le point de fonctionnement choisi pour l'identification des jets b par le réseau de neurones. Pour contrôler l'accord entre la simulation et les données pour chacune de ces variables, critère qui a été pris en compte pour le choix de ces variables, nous devons d'abord déterminer la proportion de bruit de fond W+jets.

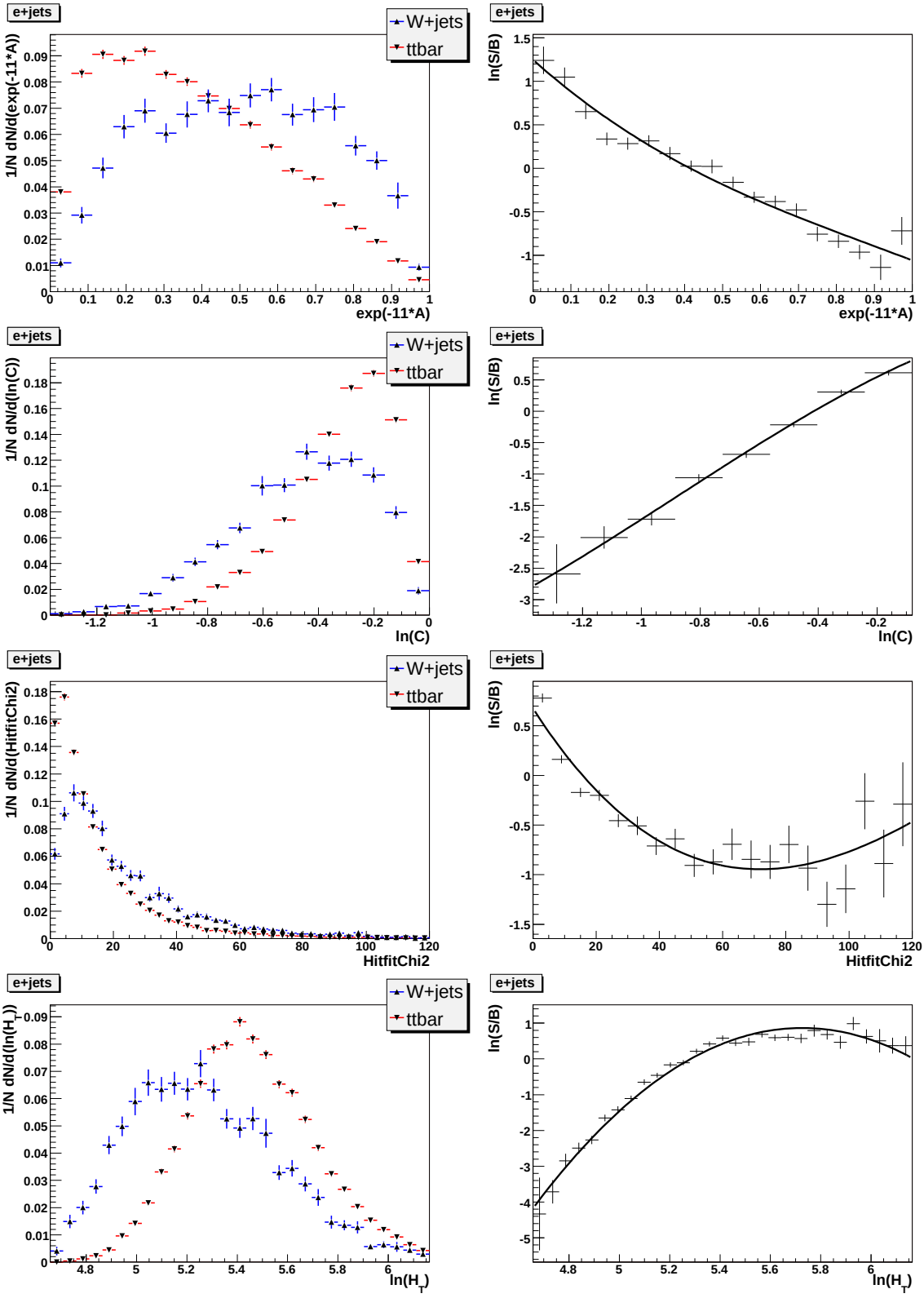

FIG. 4.6: Identique à la figure 4.5 avec, de haut en bas : *Aplanarity*, *Centrality*, $HitfitChi2$, et H_T, dans le canal e+jets.

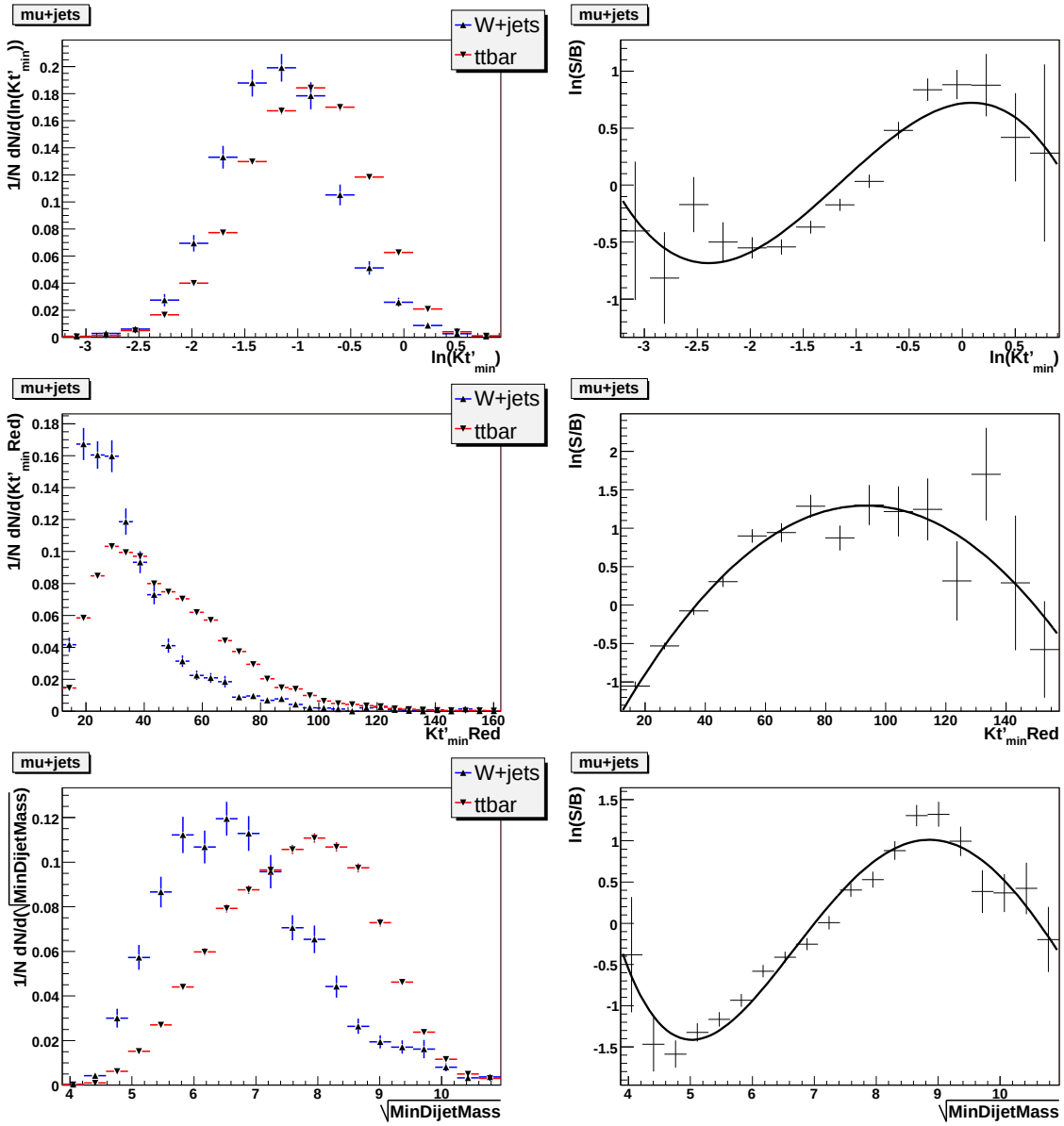

FIG. 4.7: Canal mu+jets.
Figures de gauche : Distributions normalisées simulées pour le signal $t\bar{t}$ (triangles inversés, distribution rouge) et pour le bruit de fond W+jets (triangles droits, distribution bleue) des variables étudiées pour former la fonction discriminante. De haut en bas : Kt'_{min}, $Kt'_{min}Red$ et $MinDijetMass$.
Figures de droite : logarithme du rapport signal sur bruit de fond des distributions normalisées et ajustement par un polynôme.

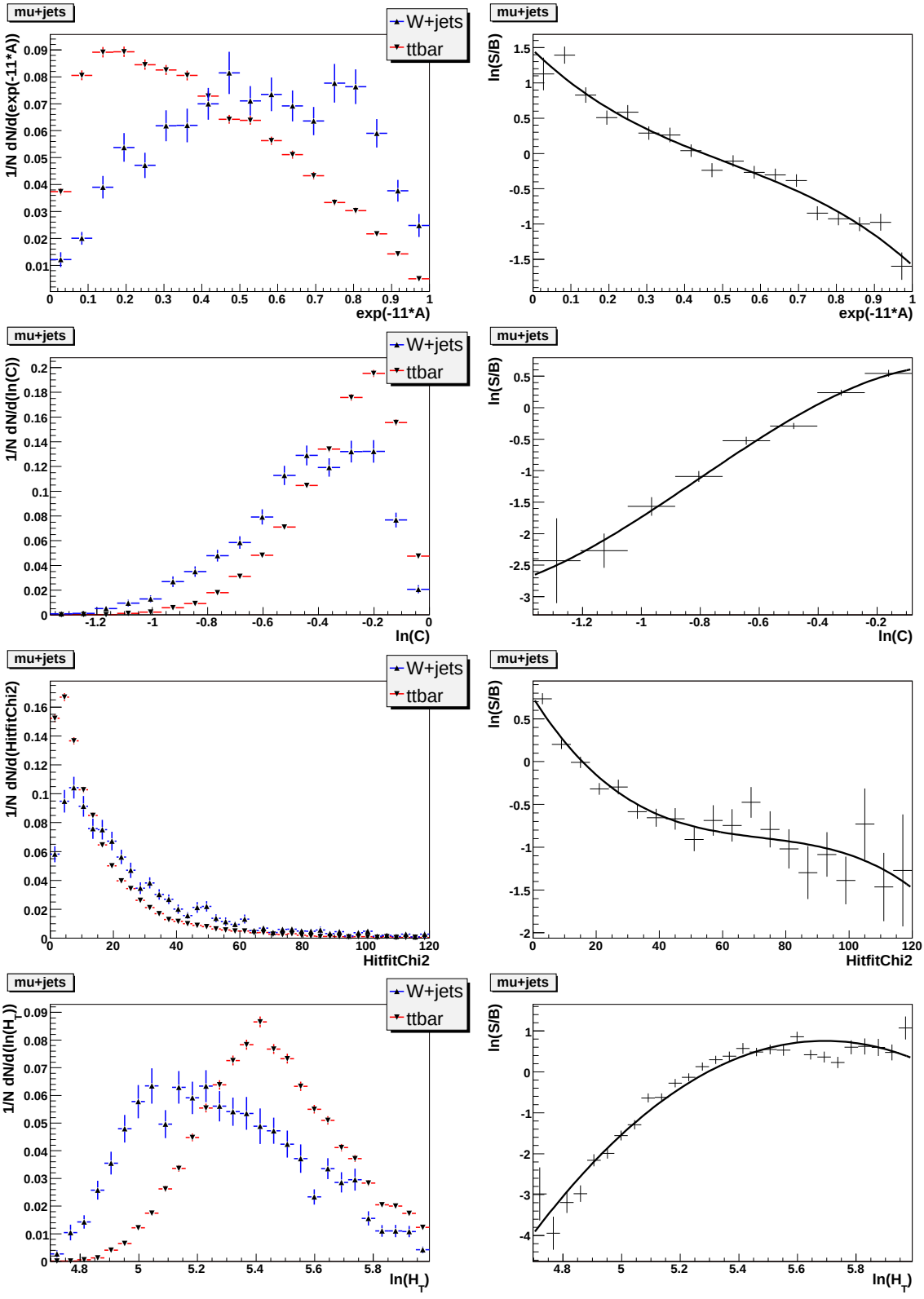

FIG. 4.8: Identique à la figure 4.7 avec, de haut en bas : *Aplanarity*, *Centrality*, *HitFitChi2* et H_T, dans le canal mu+jets.

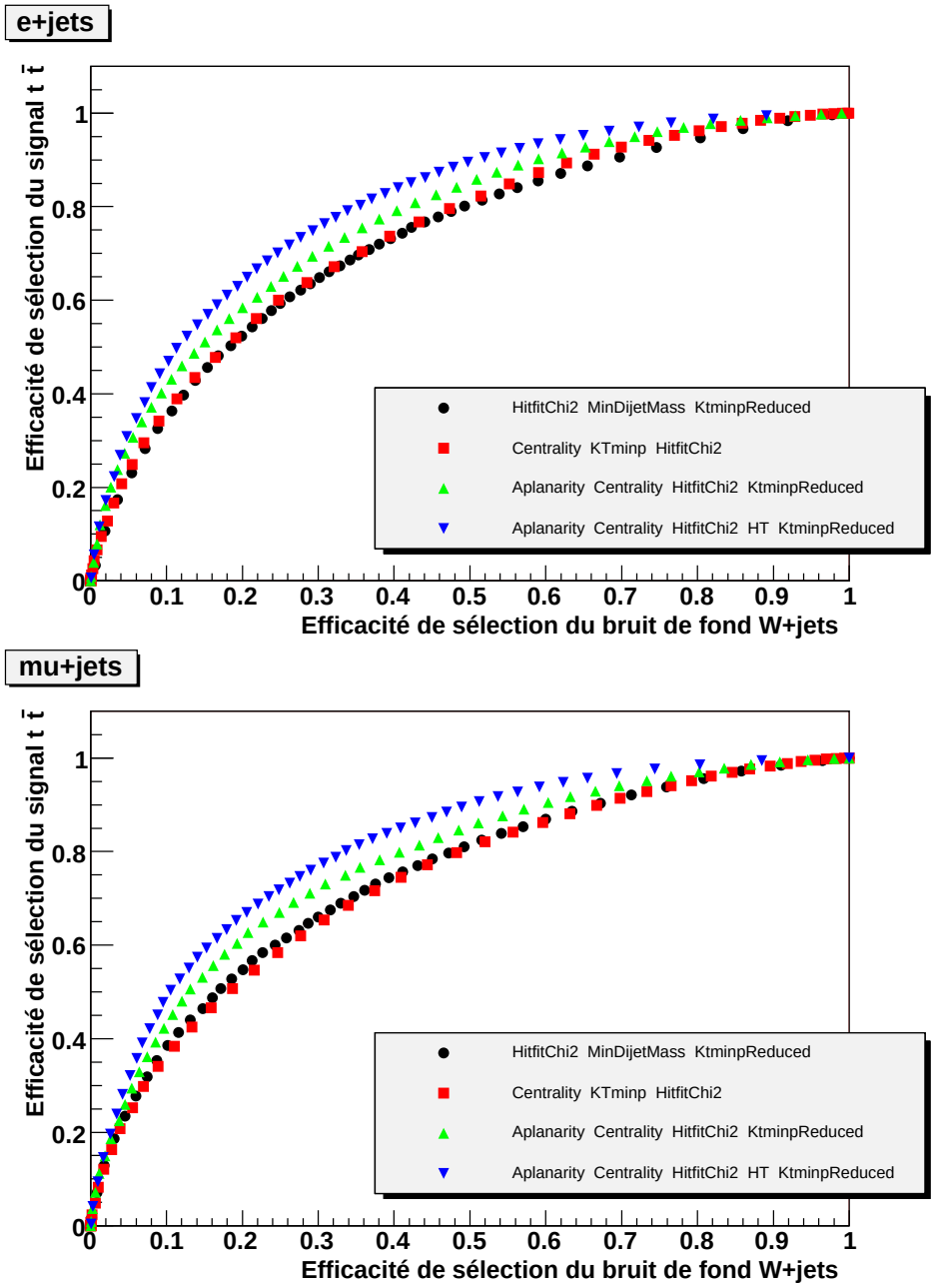

FIG. 4.9: Efficacité de sélection du signal $t\bar{t}$ en fonction de l'efficacité de sélection du bruit de fond W+jets, selon la coupure appliquée sur la fonction discriminante pour diverses combinaisons de variables. Le pas correspondant à la coupure sur le discriminant est de 0.02. La combinaison retenue dans les deux canaux correspond à la courbe représentée par les triangles inversés bleus. Figure du haut : canal e+jets. Figure du bas : canal mu+jets.

4.3.2.4 Point de fonctionnement du discriminant et du réseau de neurones

Pour optimiser la coupure sur le discriminant $D_S(\overrightarrow{x})$, la manière la plus pertinente serait d'estimer l'erreur attendue sur notre mesure d'hélicité du W en fonction de cette coupure. Ceci peut se faire par la méthode des *Ensembles Tests*, qui consiste à générer aléatoirement des événements ayant les mêmes caractéristiques que ceux utilisés pour notre mesure. En faisant cela un grand nombre de fois, nous construisons la distribution de probabilité de notre mesure, qui est une gaussienne dont la largeur correspond à l'erreur statistique attendue. Nous utiliserons cette méthode à plusieurs reprises dans la suite de l'analyse. Néanmoins dans le cas de l'optimisation de la sélection nous avons décidé d'utiliser un critère beaucoup plus simple à mettre en œuvre. En effet pour être vraiment correcte, l'étude de l'erreur attendue doit prendre en compte non seulement l'erreur statistique (estimée à partir des *Ensemble Tests*) mais aussi l'erreur systématique. En effet, par exemple en augmentant le nombre d'événements sélectionnés, l'erreur statistique peut diminuer car le signal augmente, mais le bruit de fond augmente aussi et donc l'erreur systématique correspondante peut augmenter. D'autre part l'efficacité de la coupure sur le discriminant $D_S(\overrightarrow{x})$ dépend aussi du point de fonctionnement choisi pour l'identification des jets de quarks beaux.
Il y a un compromis à trouver entre la sévérité de la coupure sur le discriminant et la qualité de l'identification des b. Nous avons ainsi choisi d'optimiser un point de fonctionnement simultané pour l'identification des b et la coupure sur $D_S(\overrightarrow{x})$. Le critère que nous avons utilisé pour cette optimisation est le rapport $\frac{S}{\sqrt{S+B}}$ où S désigne la quantité de signal $t\bar{t}$ dans les événements utilisés pour notre mesure, et B correspond à la quantité de bruit de fond W+jets. Ces deux grandeurs sont estimées par l'ajustement de la distribution de la variable $D_S(\overrightarrow{x})$, cette partie sera détaillée dans la section 4.3.3. Ce rapport $\frac{S}{\sqrt{S+B}}$ est calculé en fonction de la coupure appliquée sur la fonction discriminante, et est représenté sur la figure 4.10 pour les points de fonctionnement *L2*, *Loose* et *Medium* du réseau de neurones et séparement pour chaque canal e+jets et mu+jets. Concernant le canal e+jets, nous avons choisi le point de fonctionnement *L2*. Il permet en effet de conserver une statistique suffisante dans les données après coupure sur le NN, tout en assurant une bonne réjection du bruit de fond et ce, quel que soit la coupure appliquée sur le discriminant. Dans le cas du canal mu+jets par contre, le point de fonctionnement retenu est le point *Medium*.

Sur ces figures 4.10 nous observons qu'il est difficile d'augmenter encore le rapport $\frac{S}{\sqrt{S+B}}$ après les coupures séquentielles. Ce rapport atteint très rapidement un plateau maximum avant de chuter fortement. Nous avons appliqué une coupure sur le discriminant au départ du plateau maximal, pour ne pas perdre de statistique. Ainsi, des coupures à 0.05 dans le canal e+jets et 0.125 dans le canal mu+jets ont été retenues.

Le tableau 4.13 récapitule les efficacités de sélection obtenues dans les données et pour chaque processus de bruit de fond simulé après coupure sur le discriminant et sur

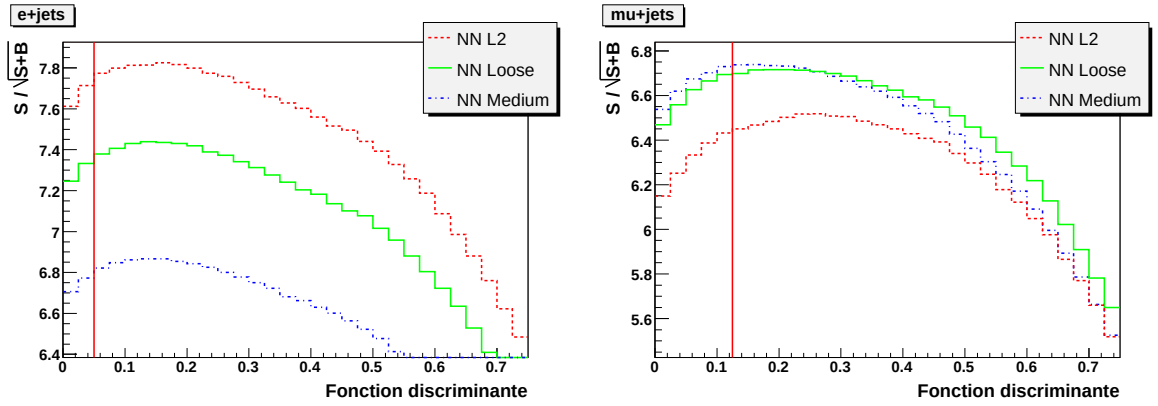

FIG. 4.10: Rapport Signal/$\sqrt{\text{Signal+Bruit}}$ en fonction de la coupure sur la fonction discriminante, pour les points de fonctionnement *L2*, *Loose* et *Medium* du réseau de neurones. Les points de fonctionnement retenus sont *L2* pour le canal e+jets et *Medium* pour le canal mu+jets. La barre verticale rouge correspond à la coupure appliquée sur la fonction discriminante. Figure de gauche : canal e+jets. Figure de droite : canal mu+jets.

	e+jets $\mathcal{L}_e > 0.05$	mu+jets $\mathcal{L}_{mu} > 0.125$
Données	0.924	0.819
$t\bar{t}$	0.987	0.959
W+jets	0.831	0.701
QCD	0.809	0.812

TAB. 4.13: Efficacité de la coupure appliquée sur la fonction discriminante, dans les canaux e+jets et mu+jets.

le réseau de neurones. Le point de fonctionnement du réseau de neurones correspondant à ces efficacités à été fixé à *L2* pour le canal e+jets et à *Medium* pour le canal mu+jets. Tous les résultats et toutes les figures qui seront présentés par la suite correspondront à ces points de fonctionnement.

4.3.3 Contribution du bruit de fond W+jets

Nous utilisons la fonction discriminante que nous venons de construire pour déterminer la proportion de bruit de fond W+jets dans les événements sélectionnés (avant coupure sur le discriminant $D_S(\overrightarrow{x})$) pour la mesure de l'hélicité du W. Nous comparons la distribution $D_{donnees}$ de la fonction discriminante observée dans les données à celle simulée, en utilisant une méthode de vraisemblance discrète (ou *binnée*).
La distribution simulée est construite à partir des distributions obtenues $D_{t\bar{t}}$ pour le

signal $t\bar{t} \rightarrow \ell + jets$, D_{W+jets} pour le bruit de fond W+jets et D_{QCD} pour le bruit de fond QCD (déterminé, lui, avec les données). Cette distribution est normalisée au nombre d'événements sélectionnés dans les données pour chacun des canaux e+jets et mu+jets. Les fractions $f_{t\bar{t}}$, f_{W+jets} et f_{QCD} des contributions sont les paramètres que nous voulons déterminer. La fraction f_{QCD} a été mesurée dans les données par la *Matrix Method*. Notre méthode de vraisemblance ajuste donc la fraction f_{W+jets} pour obtenir le meilleur accord avec les données, avec $f_{t\bar{t}} = 1 - f_{QCD} - f_{W+jets}$.
On définit la fonction de vraisemblance discrète ainsi :

$$-\ln \mathcal{L} = -\sum_{bin\ i} D^i_{donnees} \cdot \ln[N^{MM}_{QCD} \cdot D^i_{QCD} + x \cdot D^i_{W+jets} + (N_{donnees} - N^{MM}_{QCD} - x) \cdot D^i_{t\bar{t}}]$$

La distribution de cette fonction de vraisemblance est présentée sur la figure 4.11, pour chaque canal. Les fractions de W+jets mesurées sont $f_{W+jets} = 25.5^{+8.2}_{-7.7}$ % dans le canal e+jets et $f_{W+jets} = 36.2^{+8.7}_{-8.3}$ % dans le canal mu+jets. Le tableau 4.14 récapitule ces fractions relatives mesurées pour le signal $t\bar{t}$ et les bruits de fond W+jets et multijets QCD avant coupure sur la fonction discriminante.

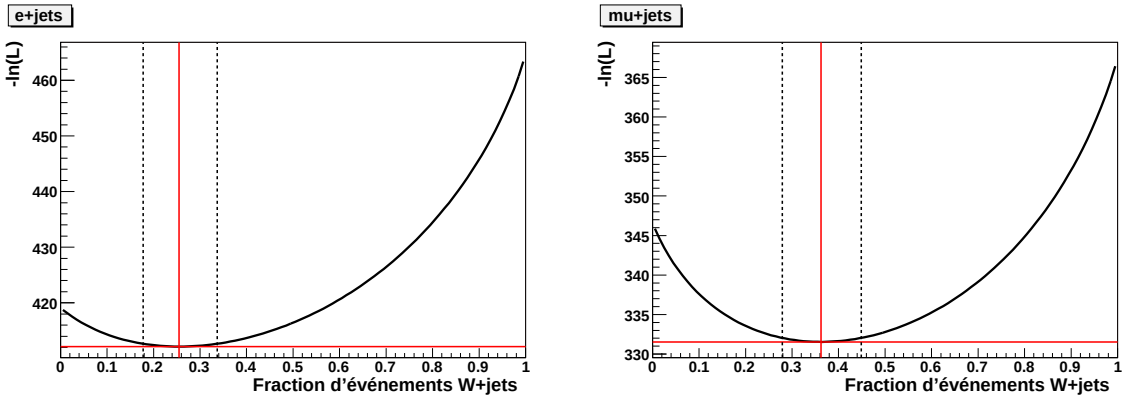

FIG. 4.11: Détermination de la contribution du bruit de fond W+jets par minimisation d'une fonction de vraisemblance. Les traits verticaux en pointillé correspondent à $\pm 1\sigma$. Figure de gauche : canal e+jets, $f_{W+jets} = 25.5^{+8.2}_{-7.7}$ %. Figure de droite : canal mu+jets, $f_{W+jets} = 36.2^{+8.7}_{-8.3}$ %.

La figure 4.12 présente la distribution de la fonction discriminante observée dans les données et décrite à partir des contributions simulées de signal $t\bar{t}$ et de bruit de fond W+jets ainsi que du bruit de fond QCD déterminé dans les données, en utilisant les normalisations relatives que nous venons de mesurer pour ces trois contributions avant coupure sur cette fonction discriminante. La figure 4.13 montre le rapport simulation sur données pour cette distribution pour illustrer la qualité de la description des données par la simulation.

	$t\bar{t}$	W+jets	QCD
e+jets (%)	66.3	25.5	8.2
mu+jets (%)	63.8	36.2	0.0

TAB. 4.14: Contributions relatives des différents processus de signal et de bruit de fond aux données sélectionnées avant coupure sur la fonction discriminante.

Ces contributions sont celles correspondant aux événement sélectionnés avant coupure sur la fonction discriminante. Pour obtenir les contributions aux données sélectionnées servant à la mesure de l'hélicité du W (c'est-à-dire après coupure sur la fonction discriminante), nous avons multiplié ces fractions par les efficacités de sélection par cette dernière coupure, qui ont été données dans le tableau 4.13. Notre lot est finalement constitué de 122 événements pour le canal e+jets et de 86 événements pour le canal mu+jets. Les contributions des différents processus après application de la coupure sur le discriminant sont données dans le tableau 4.15. La pureté des lots en signal $t\bar{t}$ est de 70.8 % et 74.7 % respectivement en e+jets et mu+jets.

	Données	$t\bar{t}$	W+jets	QCD
e+jets	122	86.38	27.97	8.76
mu+jets	86	64.24	26.65	0.00

TAB. 4.15: Contributions des différents processus de signal et de bruit de fond normalisées au nombre d'événements sélectionnés dans les données.

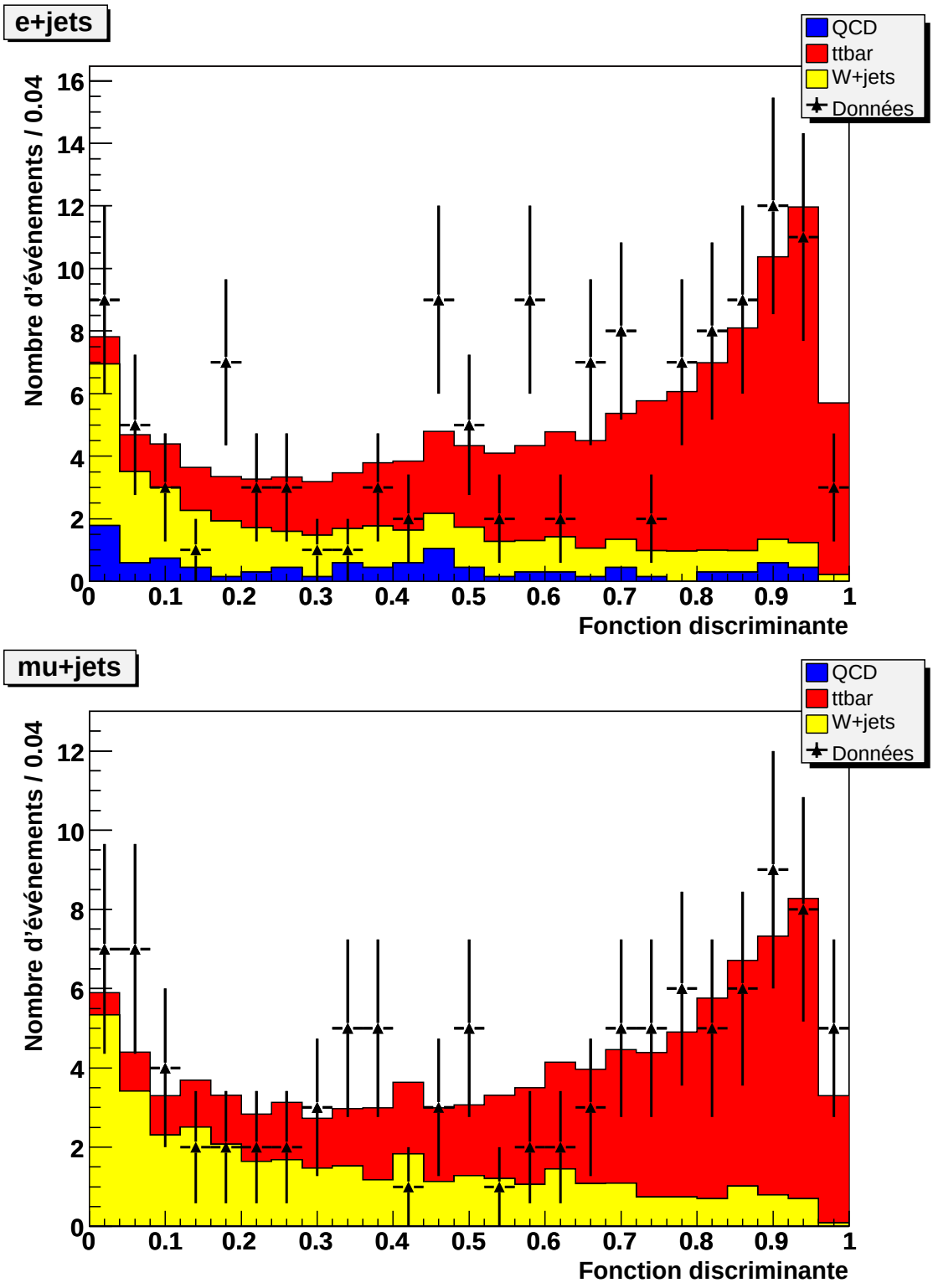

FIG. 4.12: Distributions de la fonction discriminante observées dans les données et décrites par la simulation. Figure du haut : canal e+jets. Figure du bas : canal mu+jets.

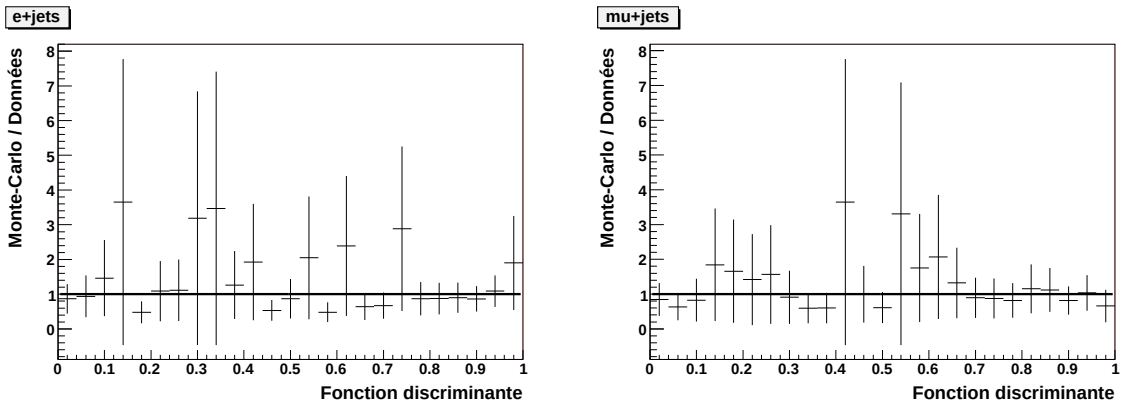

FIG. 4.13: Rapport entre la distribution de la fonction discriminante obtenue par la simulation et celle observée dans les données. Figure de gauche : canal e+jets. Figure de droite : canal mu+jets.

4.3.4 Contrôle de la description des données par la simulation

Les figures 4.14 et 4.16 présentent les distributions de contrôle de l'accord données / simulation pour les cinq variables utilisées pour construire la fonction discriminante. Les trois contributions $t\bar{t}$, W+jets simulées et QCD, ont été combinées à l'aide des fractions relatives déterminées précédemment. Ces figures correspondent aux distributions des variables avant application de la coupure sur la fonction discriminante. La coupure étant très faible, ces distributions changent très peu après coupure. L'accord entre les données et la simulation est correct, comme nous pouvons le voir également sur les figures 4.15 et figure 4.17 qui montrent le rapport des distributions simulées et réelles.

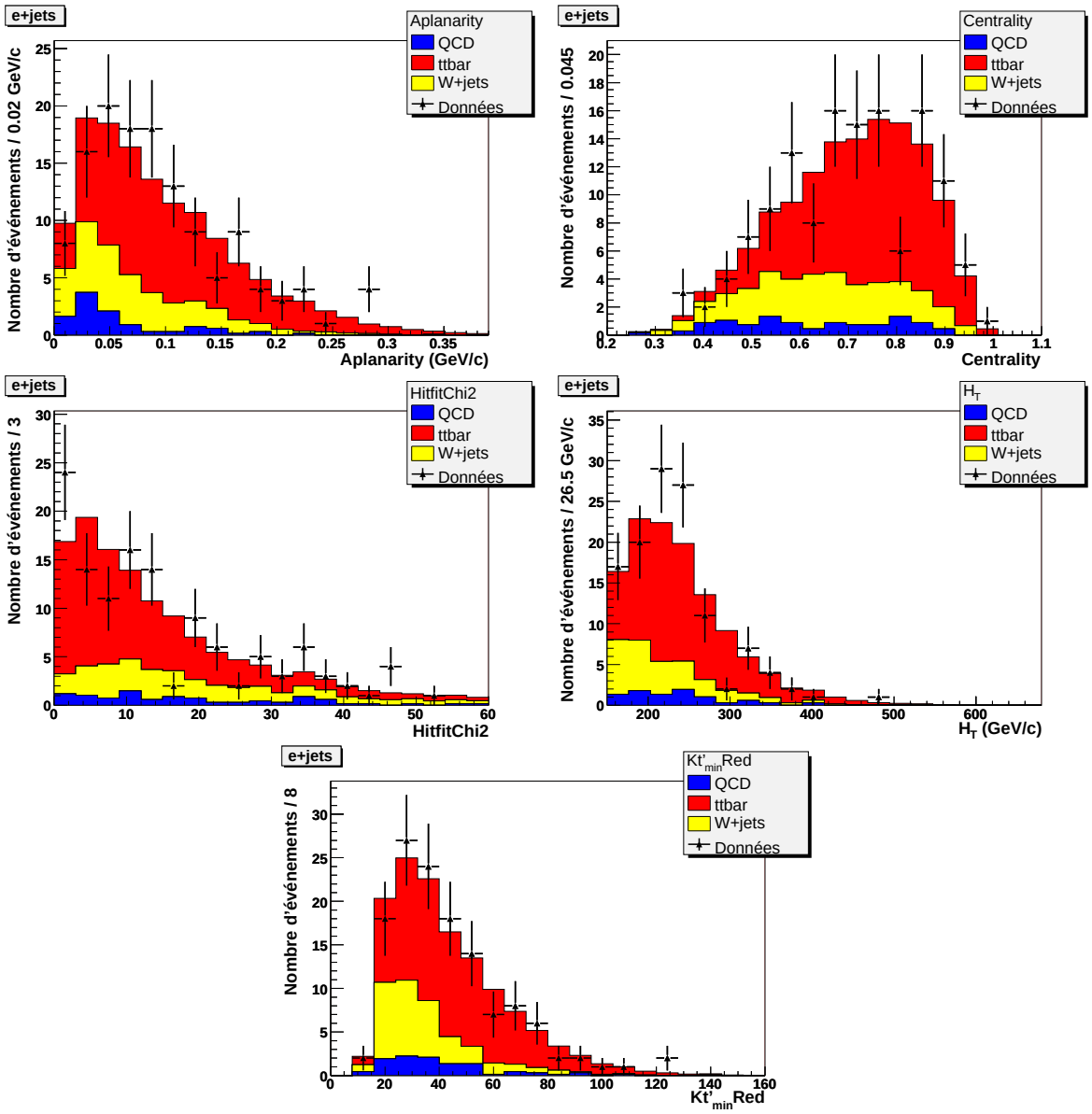

FIG. 4.14: Distributions de contrôle des variables utilisées pour construire la fonction discriminante. La distribution obtenue dans les données est comparée à la combinaison des trois composantes $t\bar{t} \to \ell + jets$, W+jets et QCD issues de la simulation pour les deux premières et des données pour le QCD. De gauche à droite et de haut en bas : *Aplanarité*, *Centralité*, *HitfitChi2*, H_T et $Kt'_{min}Red$. Ces distributions sont obtenues dans le canal e+jets avant coupure sur la fonction discriminante.

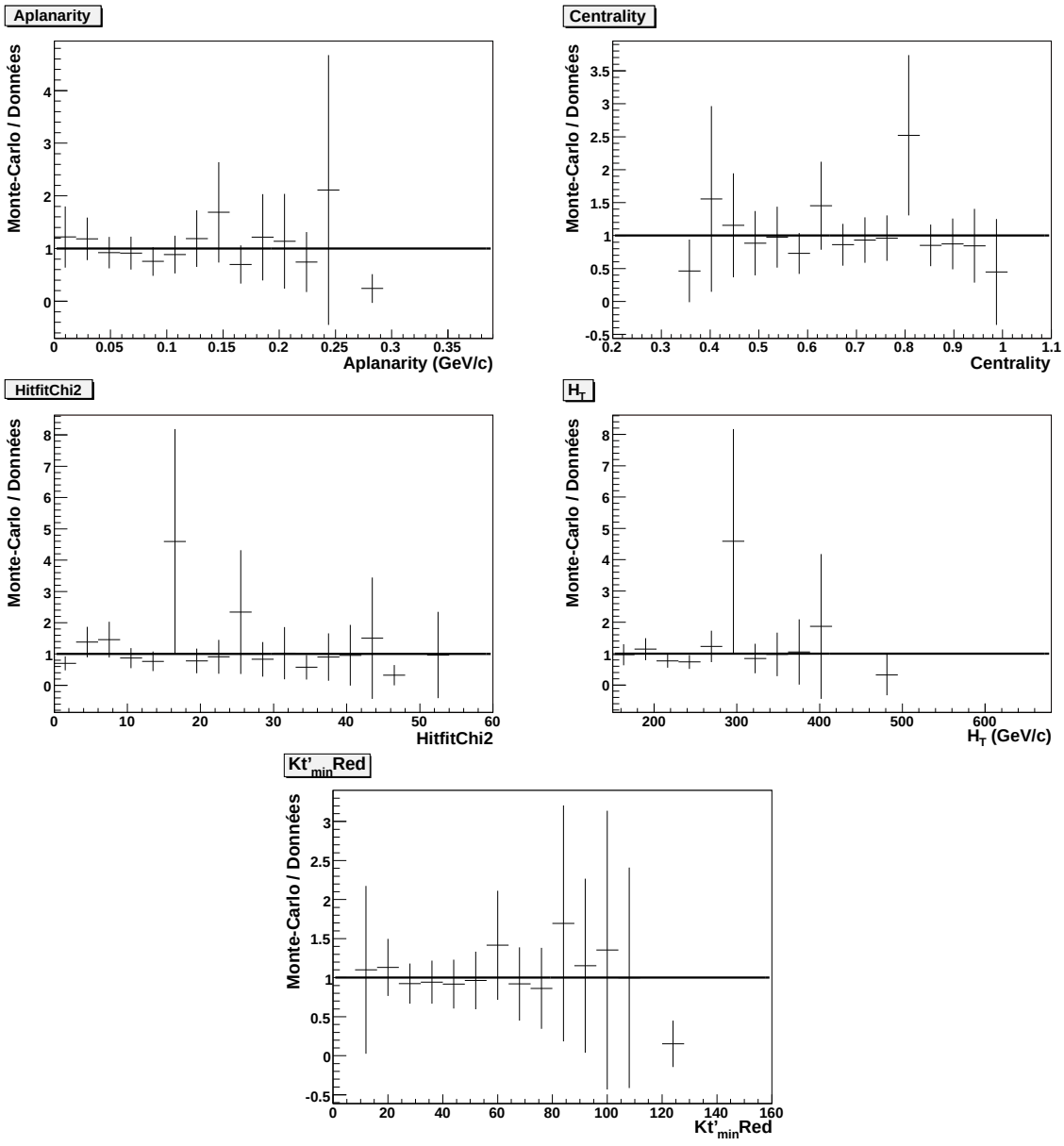

FIG. 4.15: Rapport de la distribution de contrôle issue de la simulation sur la distribution issue des données, pour chacune des variables utilisées pour construire la fonction discriminante. De gauche à droite et de bas en haut : *Aplanarité*, *Centralité*, *HitfitChi2*, H_T et $Kt'_{min}Red$. Ces distributions sont obtenues dans le canal e+jets avant coupure sur la fonction discriminante.

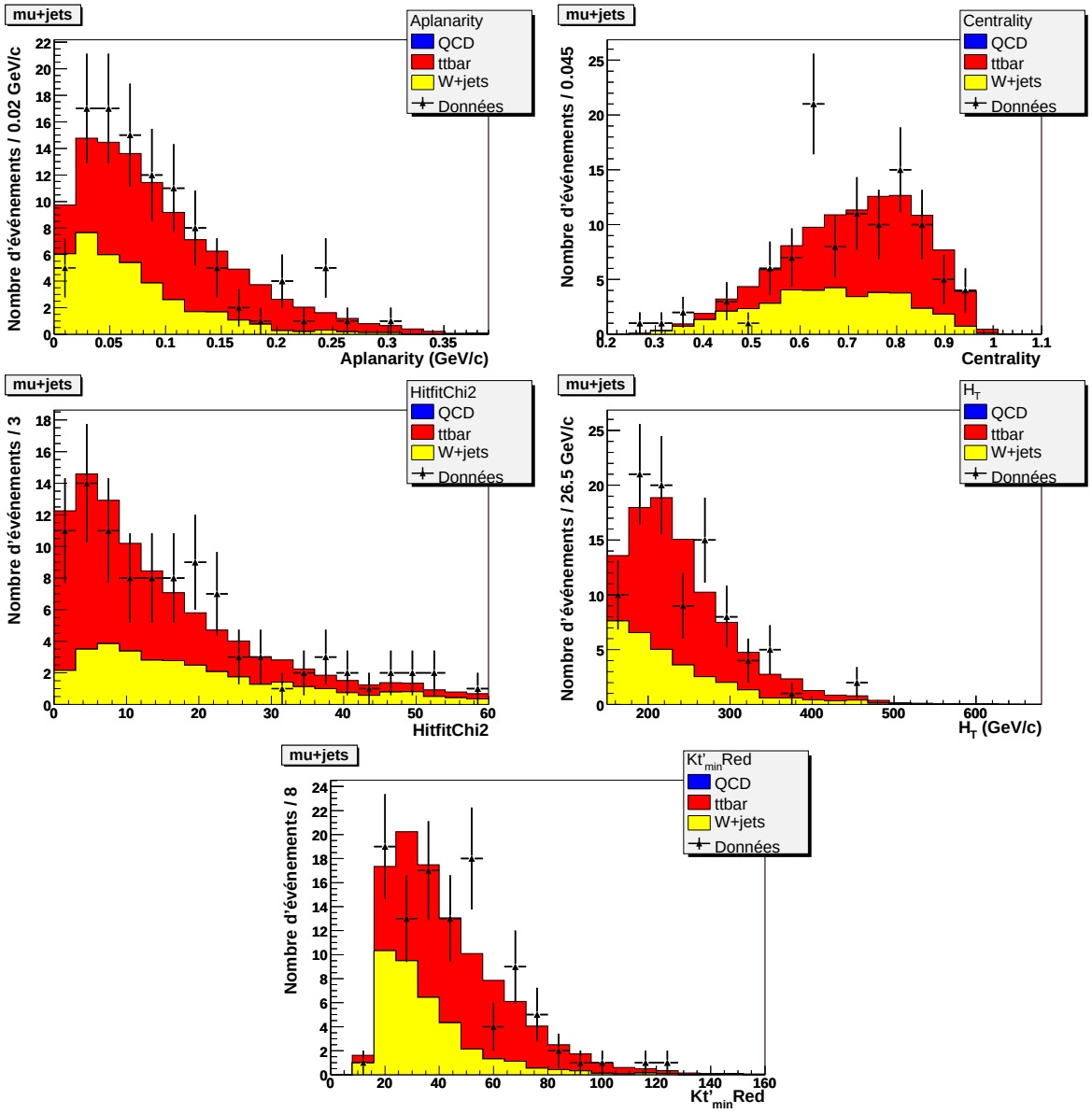

FIG. 4.16: Distributions de contrôle des variables utilisées pour construire la fonction de vraisemblance discriminante. La distribution obtenue dans les données est comparée à la combinaison des trois composantes $t\bar{t} \rightarrow \ell + jets$, W+jets et QCD issues de la simulation pour les deux premières et des données pour le QCD. De gauche à droite et de haut en bas : *Aplanarité*, *Centralité*, *HitfitChi2*, H_T et $Kt'_{min}Red$. Ces distributions sont obtenues dans le canal mu+jets avant coupure sur la fonction discriminante.

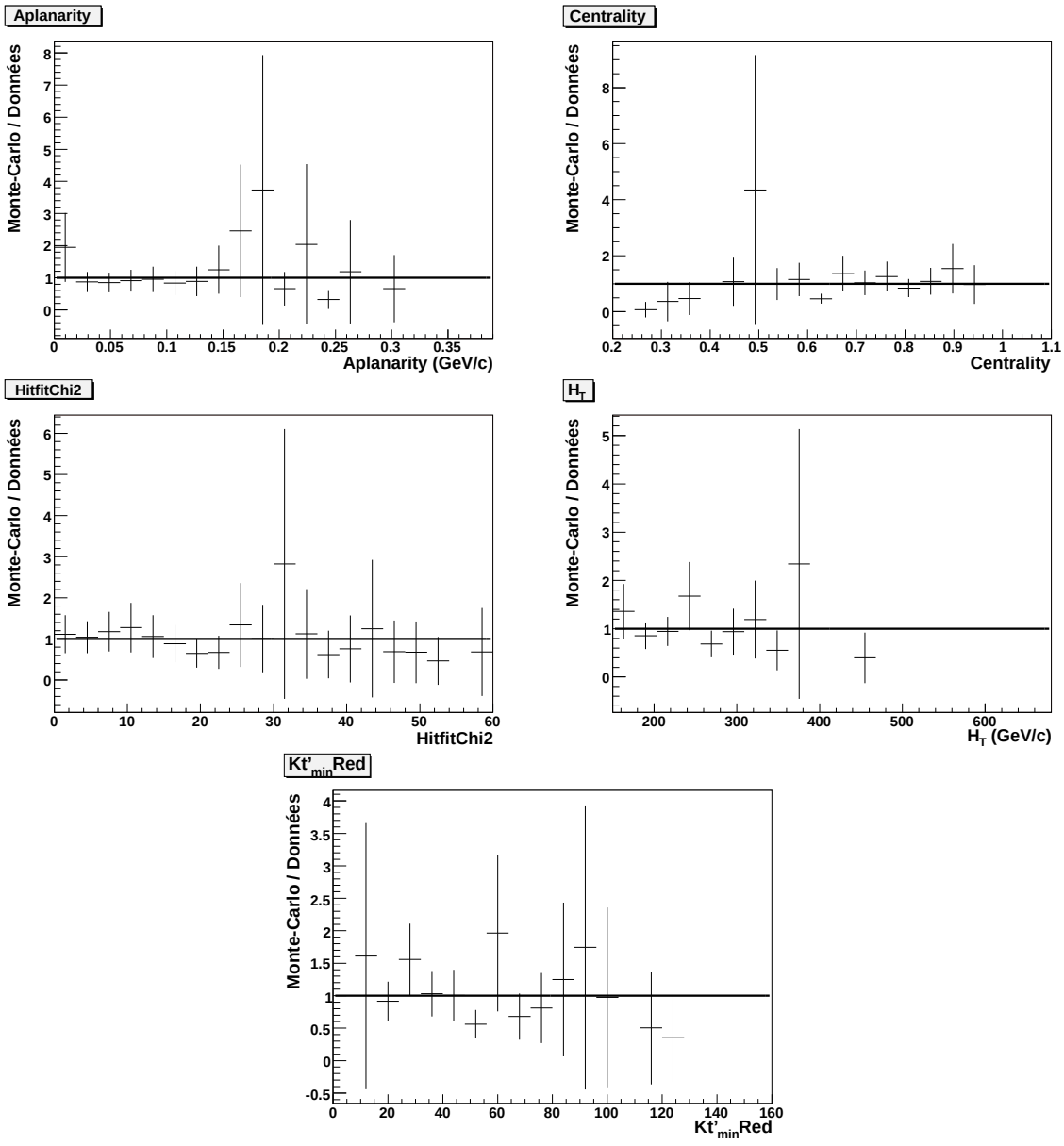

FIG. 4.17: Rapport de la distribution de contrôle issue de la simulation sur la distribution issue des données, pour chacune des variables utilisées pour construire la fonction discriminante. De gauche à droite et de bas en haut : *Aplanarité*, *Centralité*, *HitfitChi2*, H_T et $Kt'_{min}Red$. Ces distributions sont obtenues dans le canal mu+jets avant coupure sur la fonction discriminante.

5

Mesure de l'hélicité du boson W

Dans ce chapitre nous présentons la mesure de l'hélicité du boson W issu de la désintégration d'un quark top que nous effectuons avec les paires de quarks top que nous avons sélectionnées. Dans une première partie nous commentons les différents estimateurs de l'hélicité du W puis nous exposons le principe de notre mesure. Nous présentons ensuite les études de validation de notre mesure puis des sources d'erreurs systématiques et enfin le résultat final.

5.1 Estimateurs des fractions d'hélicité

5.1.1 Structure du vertex top-W-b et hélicité du W

Dans le Modèle Standard, le quark top se désintègre en un boson vecteur W^+ réel et un quark b. Ce quark b est fortement *boosté*, de telle manière que sa chiralité se confond avec son hélicité. Dans ce cas, pour conserver le spin lors de la désintégration et n'autoriser que des couplages V-A, les fractions f_0, f_+ et f_- estimées à l'ordre le plus bas dans le cadre du Modèle Standard pour les désintégrations du top en W_0^+ (W longitudinal), W_-^+ (W transverse gauche) et W_+^+ (W transverse droit) s'écrivent en fonction des masses M_t et M_W du top et du W :

$$\begin{cases} f_0 &= \frac{M_t^2}{M_t^2+2M_W^2} \\ f_- &= \frac{2M_W^2}{M_t^2+2M_W^2} \\ f_+ &= 0 \end{cases}$$

dans la limite de masse nulle du quark b. Ces fractions sont respectivement de 0.70, 0.30 et 0.00. Notons que la désintégration du quark top est la seule source de bosons W longitudinaux par un collisionneur hadronique. La seule autre source de bosons W réels est la production directe d'événements W+jets dont les W sont très majoritairement transverses.

Nous allons ainsi tester la nature standard du couplage tWb en mesurant les fractions d'hélicité du boson W pour vérifier si elles sont en accords avec ces prédictions. L'erreur théorique sur la prédiction $f_+ = 0$, due aux termes d'ordre supérieur négligés et à l'hypothèse de masse nulle pour le quark beau, est de l'ordre de 3.6×10^{-4}. La précision expérimentale qu'il faudrait donc atteindre à terme pour être sensible à une contribution non standard ne sera pas atteinte avec les données du Tevatron, il faudra attendre les données du collisionneur pp LHC. En effet, en plus de l'augmentation de la luminosité instantanée, la section efficace de production de paires de quarks top à l'énergie du LHC (14 TeV) est deux ordres de grandeur plus élevée qu'à l'énergie du Tevatron et une dizaine de millions de paires $t\bar{t}$ y sera produite chaque année dans la première phase de fonctionnement avec 10 fb^{-1} par an.

La mesure de la fraction d'hélicité f_+ permet de tester la contribution d'un couplage V+A dans la désintégration du quark top, qui diminue f_- au profit de f_+ en laissant f_0 inchangé. Cette mesure permet de tester de nombreuses théories au-delà du Modèle Standard, faisant par exemple intervenir un quark supplémentaire ayant une interaction V+A, un top composite ou des mécanismes de brisure dynamique de la symétrie électrofaible. Quant à la fraction f_0 de W longitudinaux, elle est proportionnelle au couplage de Yukawa du top. Une mesure de f_0 qui ne serait pas en accord avec la prédiction du Modèle Standard remettrait en cause le processus de brisure spontanée de la symétrie électrofaible par un mécanisme de Higgs.

5.1.2 Les différents estimateurs

L'hélicité ou le spin d'une particule instable sont des observables qui peuvent être mesurées à partir des distributions angulaires de ses produits de désintégration. Pour les produits de désintégration du W d'isospin $-\frac{1}{2}$ (lepton chargé et quarks d et s), cette distribution angulaire $\omega(\cos\theta)$ s'écrit en fonction de l'angle de désintégration θ dans le référentiel propre du W et des fractions de bosons W^+ d'hélicités longitudinale (f_0),

transverse gauche (f_-) et transverse droite (f_+) :

$$\omega(\cos\theta) = \frac{3}{8}(1+\cos\theta)^2 \cdot f_+ + \frac{3}{8}(1-\cos\theta)^2 \cdot f_- + \frac{3}{4}\sin^2\theta \cdot f_0$$

La distribution $\omega(\cos\theta)$ est représentée sur la figure 5.1.

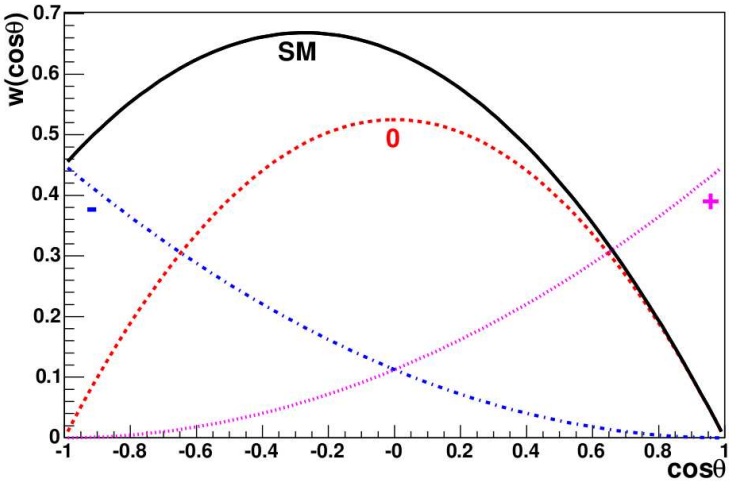

FIG. 5.1: Distribution angulaire $\omega(\cos\theta)$ de l'angle de désintégration θ entre le fermion de type *down* dans le référentiel propre du W, et le W dans le référentiel propre du top, lors de la désintégration $t \to W^+b$ avec $W^+ \to \bar{f}_{down}f_{up}$. En bleu (resp. violet) est représentée la distribution angulaire des produits de désintégration des W transverses gauches (resp. droits), et en rouge la distribution angulaire obtenue pour des W longitudinaux.

Pour mesurer cette distribution angulaire il faut être capable de différencier, parmi les produits de désintégration du W, les fermions d'isospin $+\frac{1}{2}$ de ceux d'isospin $-\frac{1}{2}$. C'est pourquoi nous utilisons la désintégration leptonique du W pour mesurer son hélicité. L'angle considéré est celui entre le lepton chargé dans le référentiel de repos du W et le W dans le référentiel de repos du top (voir figure 5.2).

Mesurer l'angle θ en toute rigueur suppose que l'on se place dans les référentiels de repos du quark top et du boson W. L'état final ℓ + jets que nous considérons dans notre analyse contient au minimum un neutrino, d'impulsion inconnue. Il faut donc soit trouver des contraintes pour être capable de calculer les quadri-impulsions du W et du top se désintégrant en lepton, soit faire des approximations, soit penser à une observable autre que l'angle θ dépendant de θ ou des fractions d'hélicité. Plusieurs observables sont utilisées dans les mesures actuelles de l'hélicité du boson W issu de la désintégration du quark top :

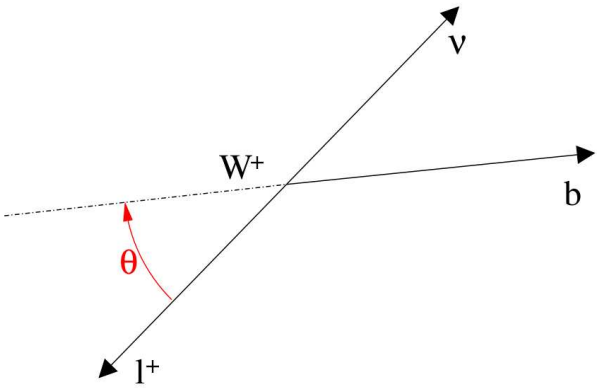

FIG. 5.2: Angle θ entre le lepton chargé dans le référentiel de repos du W et le W dans le référentiel de repos du top. La distribution de cet angle dépend des fractions d'hélicité $+$, $-$ et 0 du W.

• **L'angle θ reconstruit par un ajustement contraint :**

Les référentiels propres des deux quarks top et des deux bosons W sont reconstruits en utilisant un ajustement contraint des quantités cinématiques mesurées de l'événement (cf. section 4.2.10). L'angle θ peut ainsi être mesuré à partir des directions du lepton chargé dans le référentiel de repos du W et de l'impulsion du b dans le référentiel de repos du top.

• **L'approximation $\mathrm{M}^2_{\ell b}$:**

Si l'on ne veut pas utiliser un ajustement cinématique contraint pour reconstruire la paire de quarks top, ou encore dans le cas di-leptonique qui est sous-contraint et pour lequel la reconstruction complète des deux quarks top par un ajustement cinématique contraint est plus difficile à cause de la présence d'un neutrino dans les produits de désintégration de chaque quark top, on peut recourir à l'approximation suivante pour l'angle θ :

$$\cos\theta \sim \frac{2\mathrm{M}^2_{\ell b}}{M_t^2 - M_W^2} - 1$$

où M_t et M_W sont respectivement les masses du top et du W et $\mathrm{M}_{\ell b}$ est la masse invariante du lepton chargé et du quark beau issus de la désintégration du quark top. Lorsqu'un seul jet de quark beau est identifié, cette masse invariante est calculée une fois sur deux avec le mauvais jet. Lorsque les deux jets de quarks beaux sont identifiés, deux combinaisons de masse invariante peuvent être calculées, dont une seule est correcte.

Notons que cet estimateur de l'angle θ entraînera une incertitude systématique due à l'échelle d'énergie des jets de quarks beaux plus élevée que lors de la détermination de θ par un ajustement contraint.

• **L'impulsion transverse du lepton :**

A cause de la conservation du spin, le lepton chargé provenant de la désintégration d'un boson W longitudinal est préférentiellement émis dans la direction orthogonale à la direction de vol du W, alors que le lepton chargé venant d'un W transverse droit (resp. gauche) est préférentiellement émis dans la même direction (resp. opposée). Il en découle un spectre en impulsion transverse du lepton chargé plus dur pour les leptons de W transverses droits que pour les leptons chargés de W transverses gauches, et intermédiaire pour les leptons chargés de W longitudinaux (cf. figure 5.3).

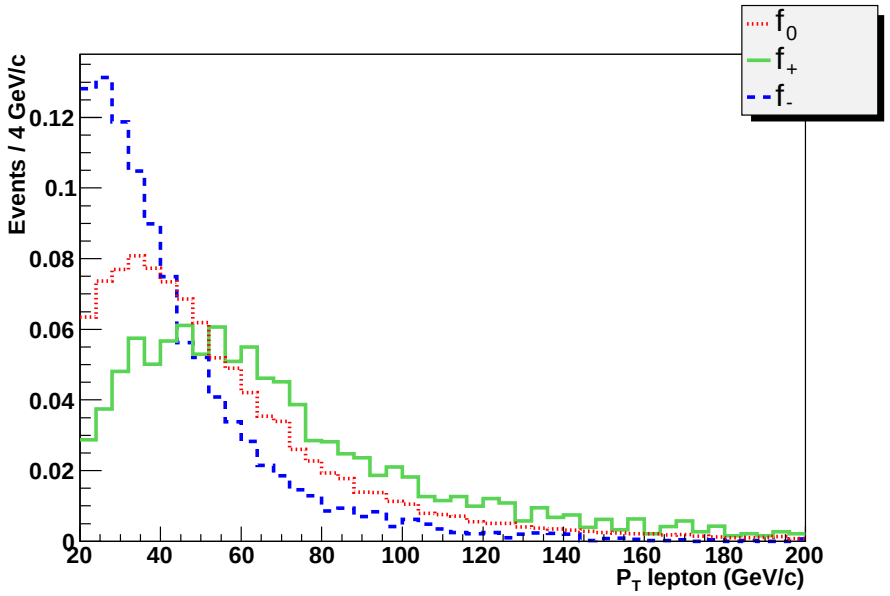

FIG. 5.3: Distribution de l'impulsion transverse du lepton chargé issu de la désintégration de bosons W longitudinaux, transverses droits et transverses gauches. Les W sont issus de la désintégration de quarks top.

Nous avons envisagé la possibilité de combiner deux estimateurs de l'angle θ pour mesurer les fractions d'hélicité du W, par exemple $\cos\theta$ et P_T^{lepton}, mais ces estimateurs sont très corrélés et nous n'avons pas poursuivi cette idée. Nous utilisons uniquement l'angle θ reconstruit suite à un ajustement contraint de l'événement.

5.2 Mesure de l'hélicité du boson W

5.2.1 Principe

La fraction d'hélicité mesurée dans les données est obtenue comme paramètre libre d'une fonction de vraisemblance discrète (construite à partir des *bins* d'histogrammes). Nous comparons la distribution de l'observable $\cos\theta$ déterminée dans les données à une distribution de référence constituée par la combinaison des trois composantes suivantes : un signal $t\bar{t}$ de fractions d'hélicité du W f_0, f_+ et f_- connues, et les bruits de fond W+jets et multijets QCD. La contribution de chacune de ces trois composantes a été déterminée auparavant (cf. section 4.3). La distribution de $\cos\theta$ du bruit de fond W+jets est issue de la simulation, et celle du bruit de fond QCD est déterminée dans les données. Pour la composante de signal $t\bar{t}$, nous utilisons trois lots distincts, respectivement d'hélicités droite, gauche et longitudinale. Les figures 5.4 et 5.5 présentent, respectivement pour le canal e+jets et le canal mu+jets, les distributions de référence de $\cos\theta$ normalisées à l'unité.

Nous minimisons la fonction de vraisemblance étendue (i.e. poissonnienne) discrète suivante :

$$-\ln\mathcal{L} = -\sum_{bin\ i} D^i_{donnees} \cdot \ln[D^i_{reference}] + \sum_{bin\ i} D^i_{reference}$$

avec :

$$D^i_{reference} = N_{QCD} \cdot D^i_{QCD} + N_{W+jets} \cdot D^i_{W+jets} + N_{t\bar{t}} \cdot (f_0 \cdot D^i_{t\bar{t}\ f_0} + f_+ \cdot D^i_{t\bar{t}\ f_+} + f_- \cdot D^i_{t\bar{t}\ f_-})$$

Les $D^i_{composante}$ correspondent à la population du *bin* i de la distribution de référence correspondante. Chaque composante est normalisée au nombre total d'événements dans les données, proportionnellement aux fractions relatives déterminées précédemment (termes N_{QCD}, N_{W+jets} et $N_{t\bar{t}}$). La composante de signal $t\bar{t}$ se décompose en trois sous constituants correspondant aux distributions de référence des trois états d'hélicité donnée du W. Nous ne procédons pas à une mesure simultanée des trois fractions d'hélicité car la statistique atteinte avec 1 fb^{-1} de données du Tevatron est trop faible. En fixant f_+ à sa valeur du Modèle Standard ($f_+ = 0.00$), le seul paramètre libre de la minimisation est f_0, avec $f_- = 1 - f_0$. De même, la mesure de la fraction d'hélicité droite f_+ est réalisée en fixant $f_0 = 0.70$, avec $f_- = 0.70 - f_+$. Nous permettons au paramètre libre de prendre des valeurs hors du domaine physique $[0.0, 1.0]$ (mesure non Bayesienne). La minimisation est réalisée par l'algorithme de minimisation MINOS du programme MINUIT.

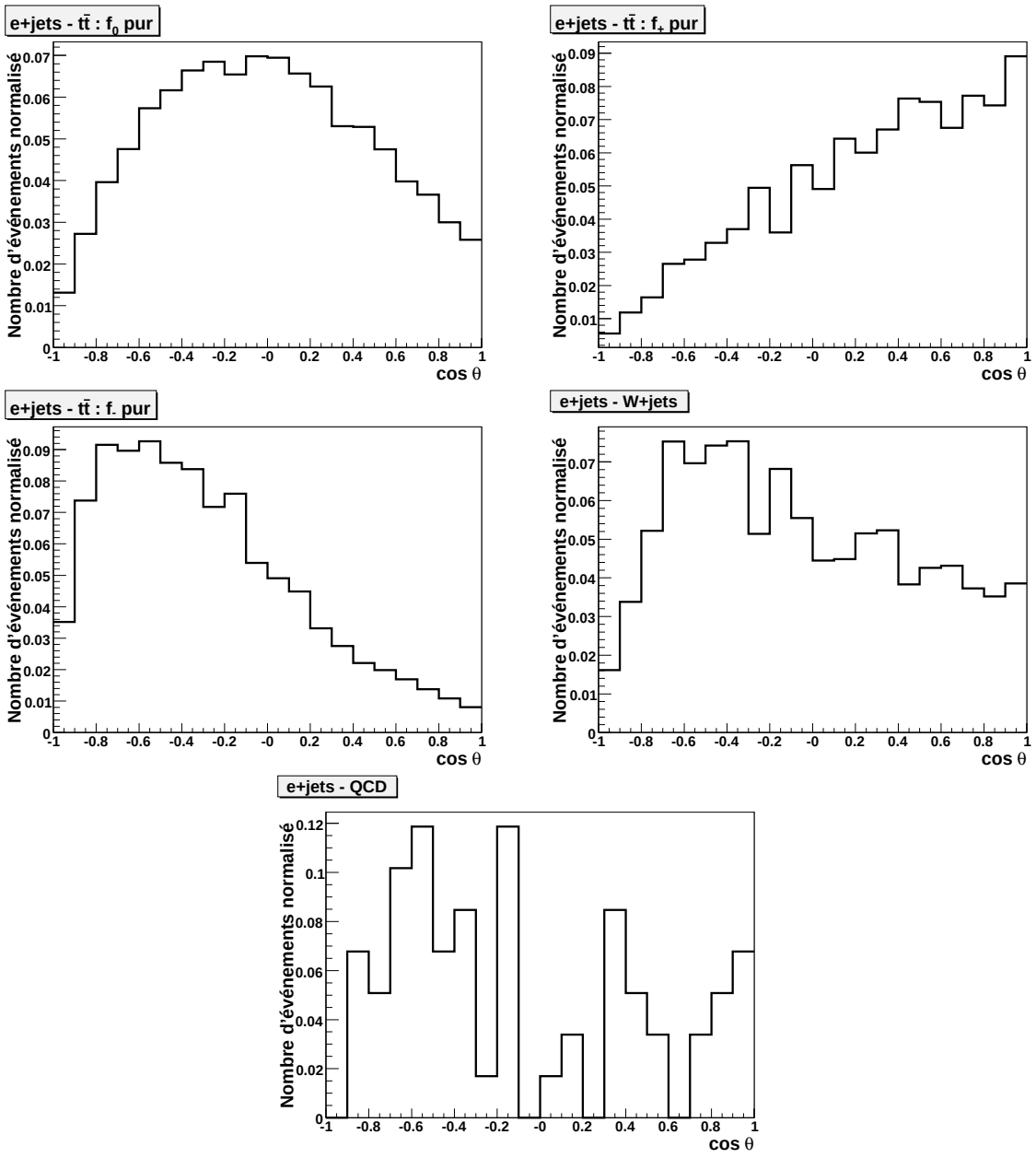

FIG. 5.4: Distributions de référence de l'observable $\cos\theta$ reconstruite dans le canal e+jets, normalisées à l'unité. De gauche à droite et de haut en bas : signal $t\bar{t}$ généré avec une hélicité du W uniquement longitudinale, $t\bar{t}$ avec des W droits, $t\bar{t}$ avec des W gauches, bruit de fond généré W+jets et bruit de fond QCD issu des données.

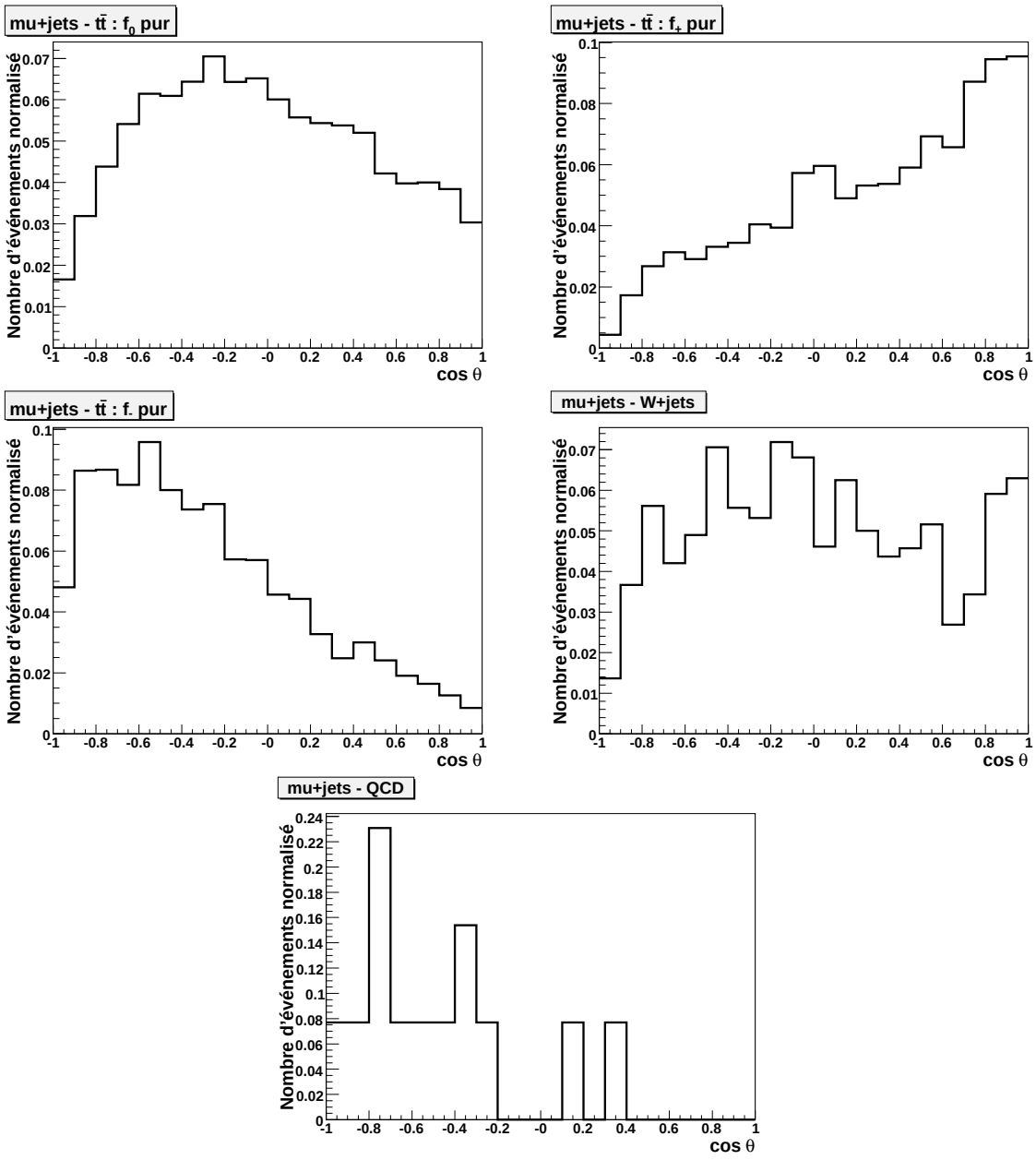

FIG. 5.5: Distributions de référence de l'observable $\cos\theta$ reconstruite dans le canal mu+jets, normalisées à l'unité. De gauche à droite et de haut en bas : signal $t\bar{t}$ généré avec une hélicité du W uniquement longitudinale, $t\bar{t}$ avec des W droits, $t\bar{t}$ avec des W gauches, bruit de fond généré W+jets et bruit de fond QCD issu des données.

5.2.2 Calibration de la méthode

Avant de procéder à la mesure dans les données réelles, nous nous assurons que notre méthode permet de mesurer correctement les fractions d'hélicité, c'est-à-dire qu'elle n'introduit pas de biais systématique. Dans le cas contraire il est nécessaire de corriger la mesure de ce biais pour obtenir une mesure correcte.

Le principe de mesure a été testé dans la simulation par la méthode dite des Ensemble Tests. Des données factices sont construites avec les mêmes caractéristiques que les données réelles à partir d'un tirage aléatoire (selon une loi de poisson) sur les distributions de référence simulées[1], en faisant varier les fractions f_0 et f_- (avec f_+ fixée à 0.00) ou les fractions f_+ et f_- (avec f_0 fixée à 0.70). Le nombre d'événements est normalisé au nombre d'événements dans les données multiplié par un facteur 10000 afin de rendre l'erreur statistique négligeable. Il a été vérifié que l'erreur statistique ainsi déterminée correspond à celle attendue et varie correctement avec l'augmentation de la statistique. Au total, 1000 lots de données factices sont construits pour chaque valeur de fraction générée (avec un pas de 0.05 pour les fractions testées). Les lots simulés utilisés pour créer ces données factices sont décorrélés de ceux auxquels ils sont comparés pour effectuer la mesure (les distributions de références). La distribution des 1000 fractions mesurées pour une fraction générée est ajustée par une gaussienne, dont la moyenne est extraite. Ces moyennes des fractions mesurées sont présentées en fonction des fractions générées sur la figure de gauche des figures 5.6 et 5.7. Cette courbe, correspondant à une courbe de calibration, peut être ajustée par une droite. Cette calibration révèle l'existence d'un biais pour chaque canal et chacune des fractions mesurées, soit une pente différente de l'unité et/ou un décalage à l'origine. Le tableau 5.1 récapitule les paramètres de la calibration pour les deux canaux d'études, pour chacune des fractions. Toutes les mesures réalisées ultérieurement ont été faites en corrigeant la mesure brute suivant ces courbes de calibration.

	f_0		f_+	
	Pente	Décalage à l'origine	Pente	Décalage à l'origine
e+jets	1.00	-0.021	0.97	-0.004
mu+jets	1.00	-0.031	0.95	-0.014

TAB. 5.1: Paramètres des courbes de calibration pour les mesures de f_0 et f_+.

Les figures de droite des figures 5.6 et 5.7 représentent la distribution de la fraction d'hélicité mesurée pour une fraction générée correspondant à la valeur du Mo-

[1]Par abus de langage et afin de simplifier, la contribution QCD, déterminée à partir des données, sera considérée comme faisant partie de la simulation, par analogie avec les composantes W+jets et $t\bar{t}$ et par opposition aux données réelles utilisées pour la mesure.

dèle Standard. La statistique utilisée pour obtenir ces distributions est normalisée au nombre d'événements dans les données, et 50000 Ensemble Tests ont été produits. Ces distributions illustrent la dispersion des mesures autour de la vraie valeur. Leur largeur correspond à l'erreur statistique attendue pour notre mesure : $\sigma_{stat}^{e+jets}(f_0) = 0.219$, $\sigma_{stat}^{mu+jets}(f_0) = 0.267$, $\sigma_{stat}^{e+jets}(f_+) = 0.114$ et $\sigma_{stat}^{mu+jets}(f_+) = 0.135$.

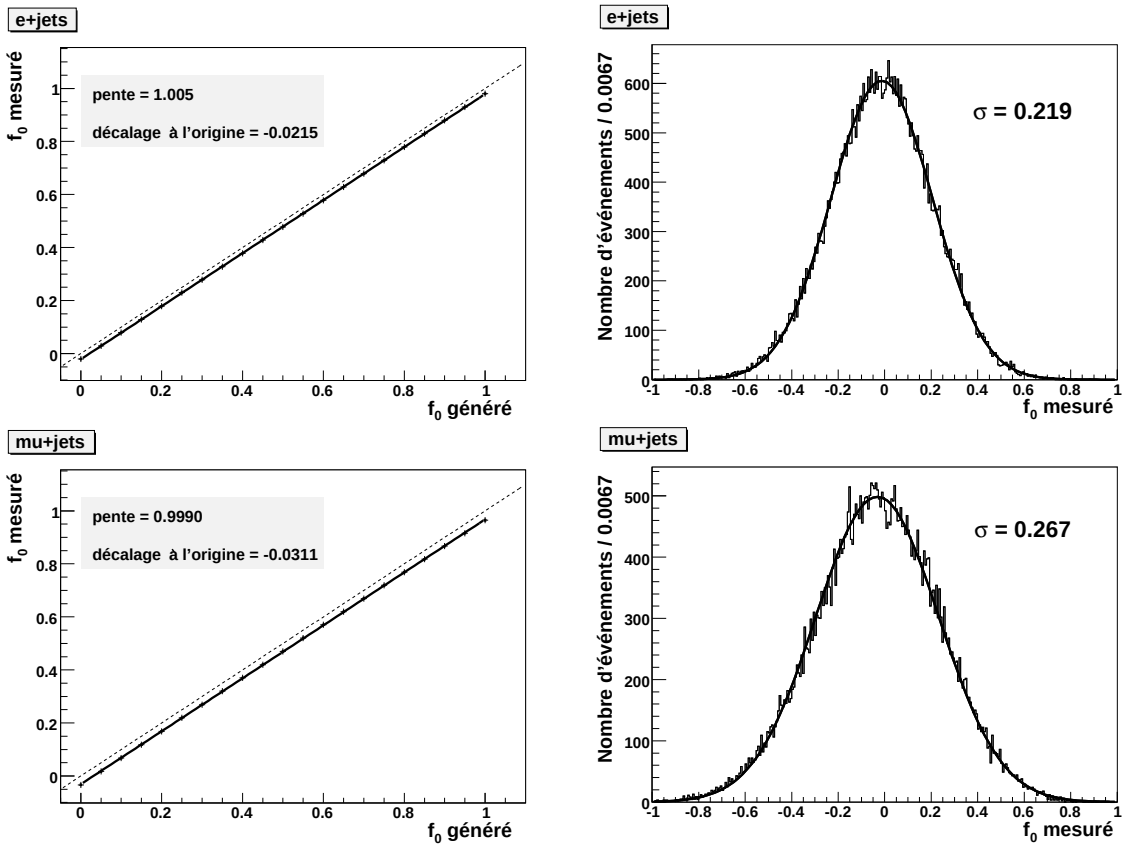

FIG. 5.6: Figures de gauche : fraction f_0 de W longitudinaux mesurée en fonction de la fraction f_0 générée (pas de 0.05 et statistique correspondant à la statistique des données multipliée par 10000). Tous les points de cette courbe sont corrélés. En pointillé est représentée la bissectrice correspondant au cas idéal où la mesure de la fraction correspondrait à la fraction générée. Un ajustement est réalisé afin de déterminer le biais induit par notre mesure.
Figures de droite : distribution de la fraction d'hélicité mesurée pour une fraction générée correspondant à la valeur du Modèle Standard. La largeur de la distribution correspond à l'erreur statistique attendue sur la mesure.
Figures du haut : canal e+jets. Figures du bas : canal mu+jets.

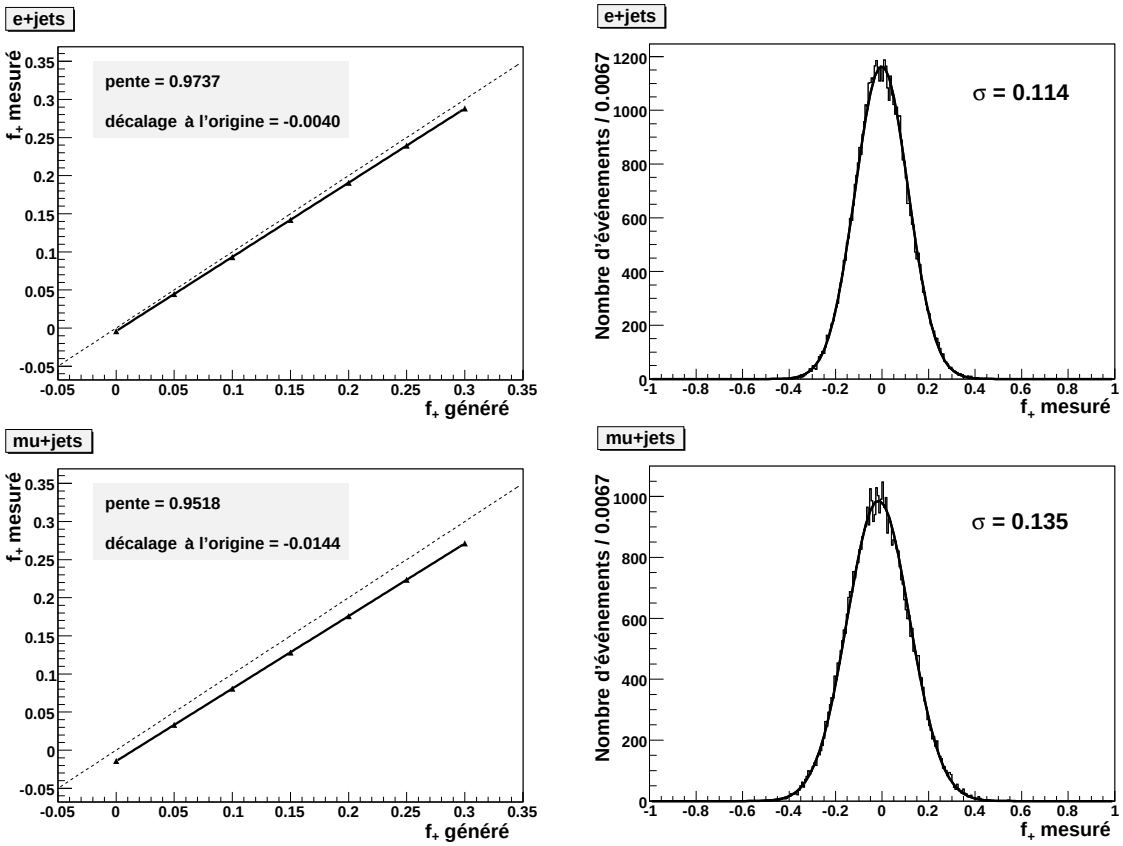

FIG. 5.7: Figures de gauche : fraction f_+ de W transverses droits mesurée en fonction de la fraction f_+ générée (pas de 0.05 et statistique correspondant à la statistique des données multipliée par 10000). Tous les points de cette courbe sont corrélés. En pointillé est représentée la bissectrice correspondant au cas idéal où la mesure de la fraction correspondrait à la fraction générée. Un ajustement est réalisé afin de déterminer le biais induit par notre mesure.
Figures de droite : distribution de la fraction d'hélicité mesurée pour une fraction générée correspondant à la valeur du Modèle Standard. La largeur de la distribution correspond à l'erreur statistique attendue sur la mesure.
Figures du haut : canal e+jets. Figures du bas : canal mu+jets.

5.2.3 Mesure dans les données

Les fonctions de vraisemblance sont déterminées dans chaque canal séparément en les corrigeant du biais déterminé précédemment, puis combinées. Le minimum de la fonction combinée correspond à la mesure de la fraction d'hélicité. L'erreur statistique est directement extraite de cette fonction.

5.2.3.1 Fraction d'hélicité longitudinale f_0

Nous mesurons f_0 suivant la méthode que nous avons présentée. La figure 5.8 présente le résultat de la minimisation de la vraisemblance dans les canaux e+jets et mu+jets, ainsi que le résultat combiné. La fraction d'hélicité longitudinale mesurée est $f_0 = 0.493^{+0.170}_{-0.171}$ (stat.). L'erreur statistique mesurée dans chaque canal correspond à l'erreur statistique attendue déterminée à la section 5.2.2. Le tableau 5.2 récapitule les mesures effectuées dans les deux canaux séparement et leur combinaison.

e+jets	mu+jets	Combinaison
$0.327^{+0.216}_{-0.213}$	$0.778^{+0.270}_{-0.282}$	$0.493^{+0.170}_{-0.171}$

TAB. 5.2: Mesure de la fraction f_0 de bosons W longitudinaux, dans les canaux e+jets et mu+jets, et leur combinaison. L'erreur indiquée est l'erreur statistique.

La figure 5.9 compare la distribution de l'observable $\cos\theta$ observée dans les données à celle simulée en combinant les contributions $t\bar{t}$, W+jets et QCD, dans les états finals e+jets et mu+jets. La contribution du signal $t\bar{t}$ est un mélange des trois lots d'états d'hélicité du W transverse droit, transverse gauche et longitudinal. Les proportions du mélange correspondent à celles mesurées dans les données pour le canal correspondant en fixant f_+ à sa valeur standard. Cette figure présente également le rapport simulation sur données des distributions de $\cos\theta$.

5.2.3.2 Fraction d'hélicité droite f_+

Pour cette mesure, la fraction de bosons W longitudinaux est fixée à sa valeur standard $f_0 = 0.70$, et la fonction de vraisemblance discrète est minimisée de manière à extraire la fraction f_+ de bosons W d'hélicité transverse droite. La fraction prédite par le Modèle Standard est $f_+ = 0.00$. La figure 5.10 présente le résultat de la minimisation de la vraisemblance dans les canaux e+jets et mu+jets, ainsi que le résultat combiné. La fraction d'hélicité longitudinale mesurée est $f_+ = -0.109^{+0.099}_{-0.074}$ (stat.). L'erreur statistique mesurée dans chaque canal correspond à l'erreur statistique attendue déterminée à la section 5.2.2. Le tableau 5.3 récapitule les mesures effectuées dans les deux canaux séparement et leur combinaison.

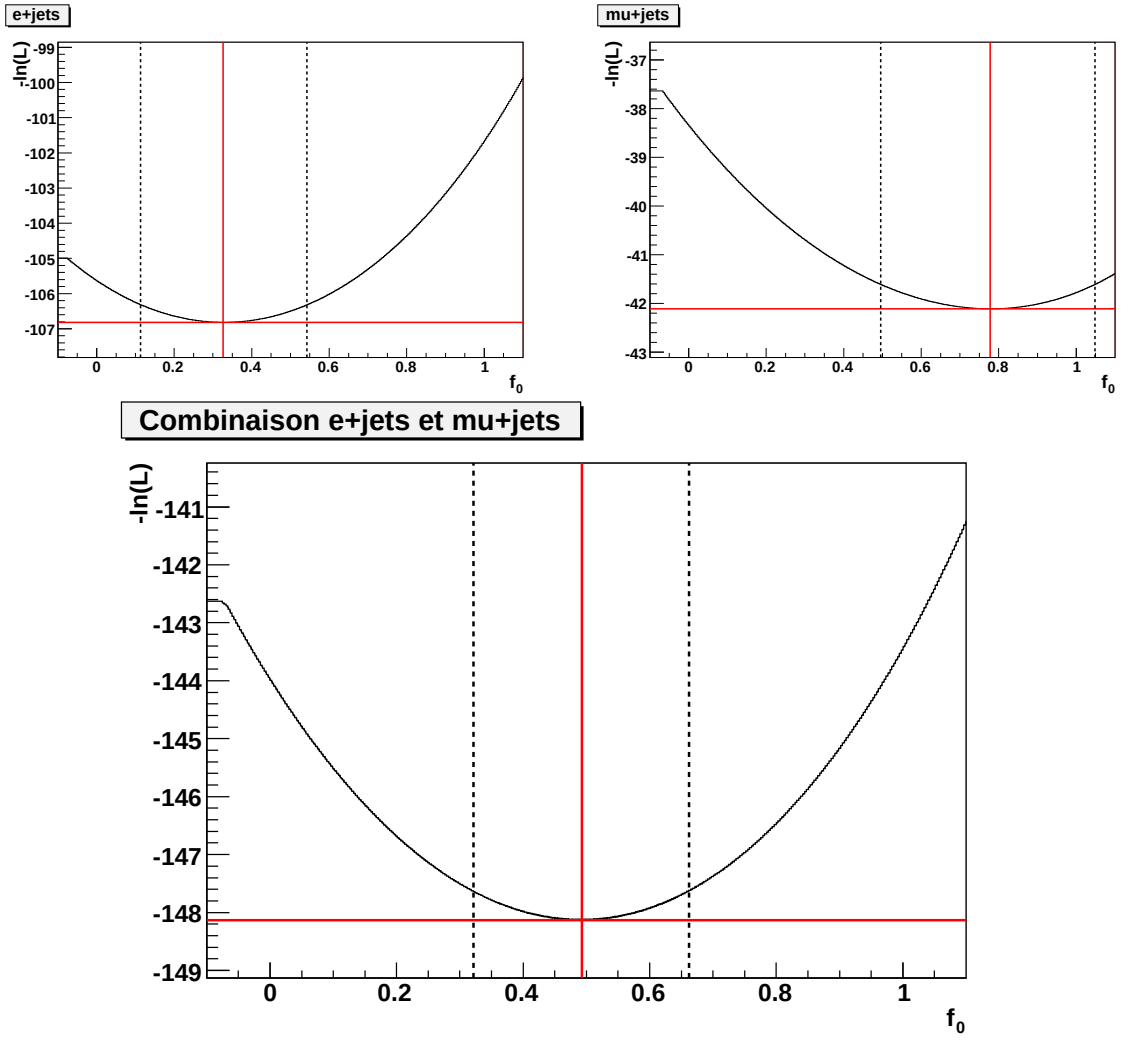

FIG. 5.8: Mesure de la fraction f_0 de W longitudinaux par minimisation de la fonction de vraisemblance, après correction du biais. Figure haut gauche : mesure effectuée dans le canal e+jets, $f_0^{e+jets} = 0.327^{+0.216}_{-0.213}$. Figure haut droite : mesure effectuée dans le canal mu+jets, $f_0^{mu+jets} = 0.778^{+0.270}_{-0.282}$. Figure du bas : mesure combinée, $f_0 = 0.493^{+0.170}_{-0.171}$.

e+jets	mu+jets	Combinaison
$-0.162^{+0.109}_{-0.099}$	$0.007^{+0.151}_{-0.143}$	$-0.109^{+0.099}_{-0.074}$

TAB. 5.3: Mesure de la fraction f_+ de bosons W transverses droits, dans les canaux e+jets et mu+jets, et leur combinaison. L'erreur indiquée est l'erreur statistique.

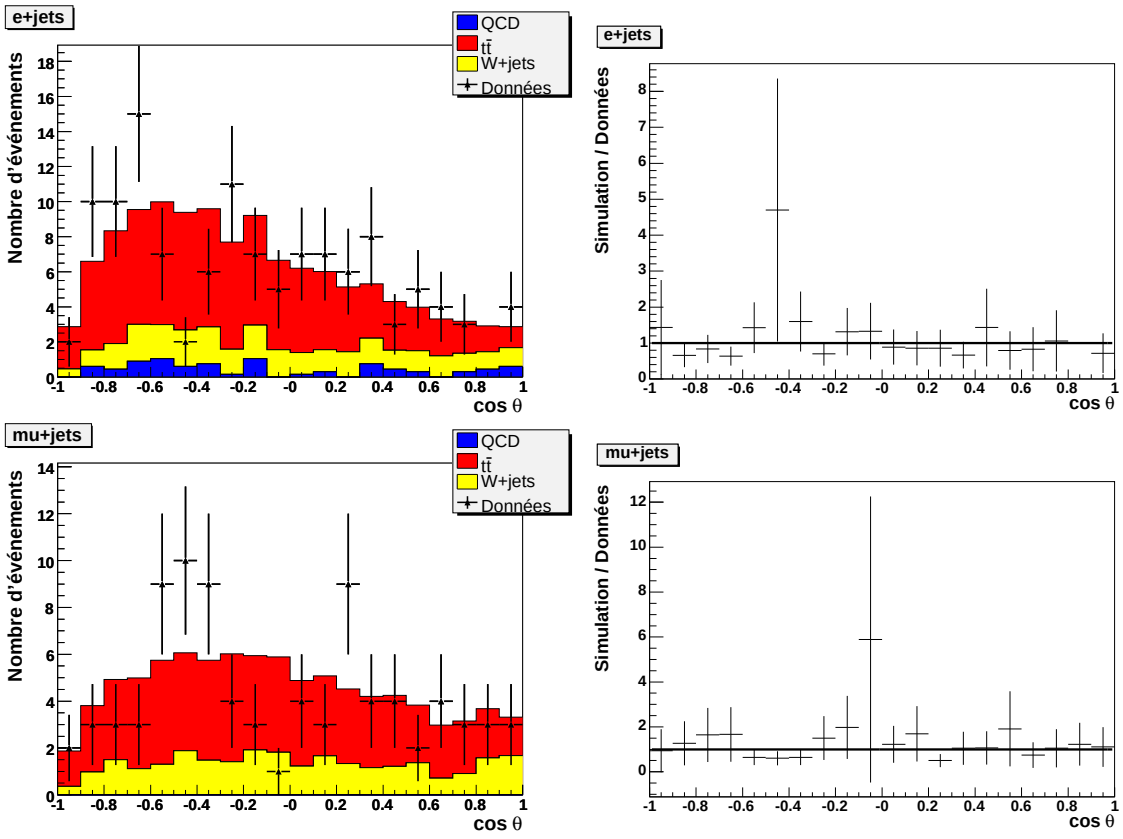

FIG. 5.9: Accord entre les données et la simulation de la distribution de l'observable $\cos\theta$. Les figures de gauche comparent la distribution issue des données à celle créée par la combinaison des distributions des trois composantes $t\bar{t}$, W+jets (simulées) et QCD (issue des données). La composante $t\bar{t}$ utilisée ici correspond à la combinaison des trois lots $t\bar{t}$ simulés avec des fractions d'hélicité purement f_0, f_+ et f_-, respectivement. La contribution de chacun des trois lots correspond à celle mesurée pour le canal correspondant, lorsque l'on a fixé $f_+ = 0.00$ et mesuré f_0. Les figures de droite présentent le rapport de la distribution simulée sur la distribution issue des données.
Les figures du haut correspondent aux distributions obtenues dans le canal e+jets, celles du bas aux distributions dans le canal mu+jets.

La figure 5.11 est identique à la figure 5.9, sauf que cette fois-ci les proportions du mélange des états d'hélicité correspondent à ceux obtenus en mesurant f_+ dans nos données.

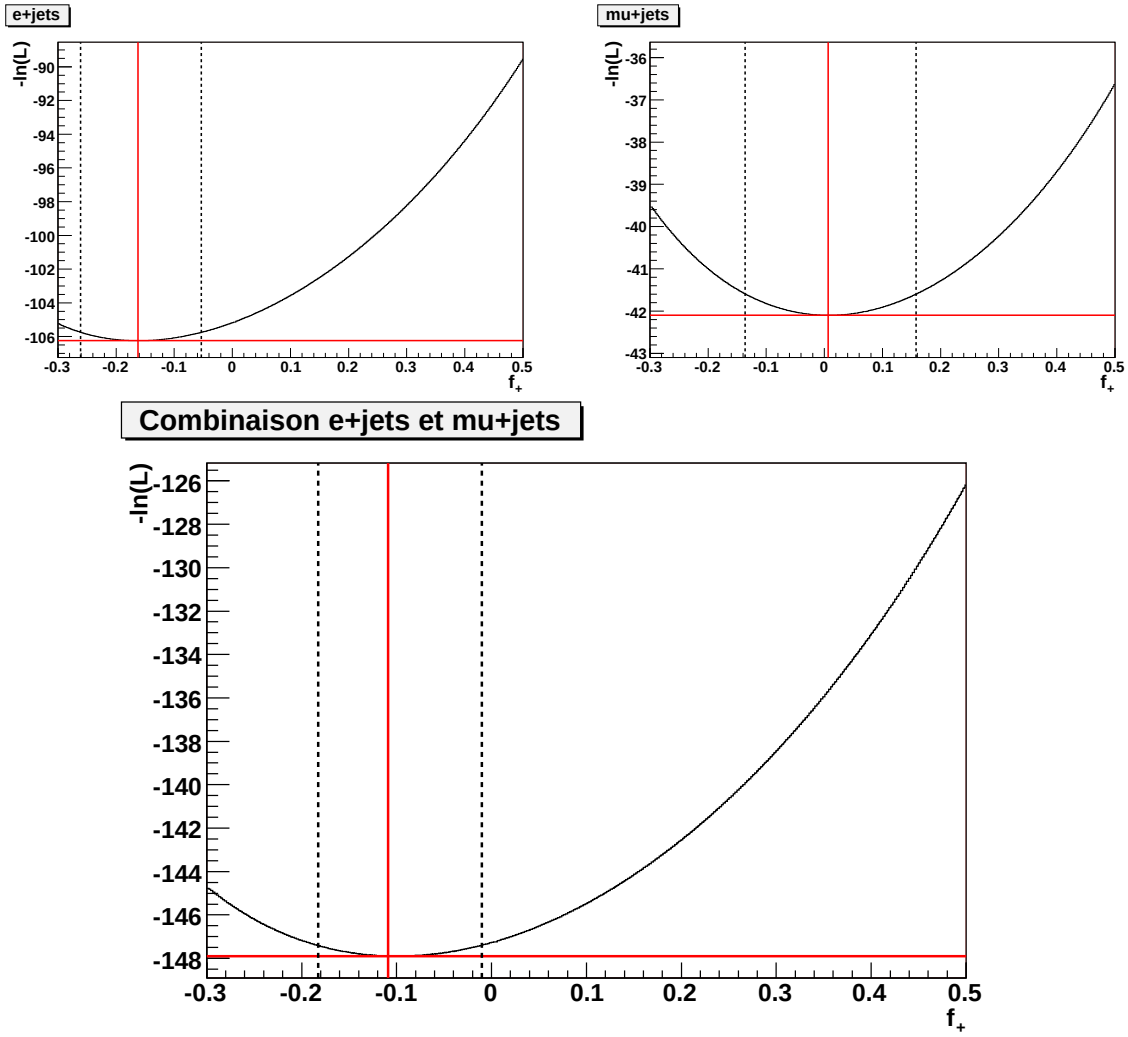

FIG. 5.10: Mesure de la fraction f_+ de W transverses droits par minimisation de la fonction de vraisemblance, après correction du biais. Figure haut gauche : mesure effectuée dans le canal e+jets, $f_+^{e+jets} = -0.162^{+0.109}_{-0.099}$. Figure haut droite : mesure effectuée dans le canal mu+jets, $f_+^{mu+jets} = 0.007^{+0.151}_{-0.143}$. Figure du bas : mesure combinée, $f_+ = -0.109^{+0.099}_{-0.074}$.

5.3 Etude des erreurs systématiques

5.3.1 Principe de la mesure des erreurs

Notre mesure est basée sur la comparaison des données à des distributions de référence. Ces références sont obtenues à partir d'événements simulés. Cette simulation utilise des modèles (pour la fragmentation par exemple) et des paramètres fondamentaux dont les valeurs sont connues avec une erreur expérimentale (comme la masse du

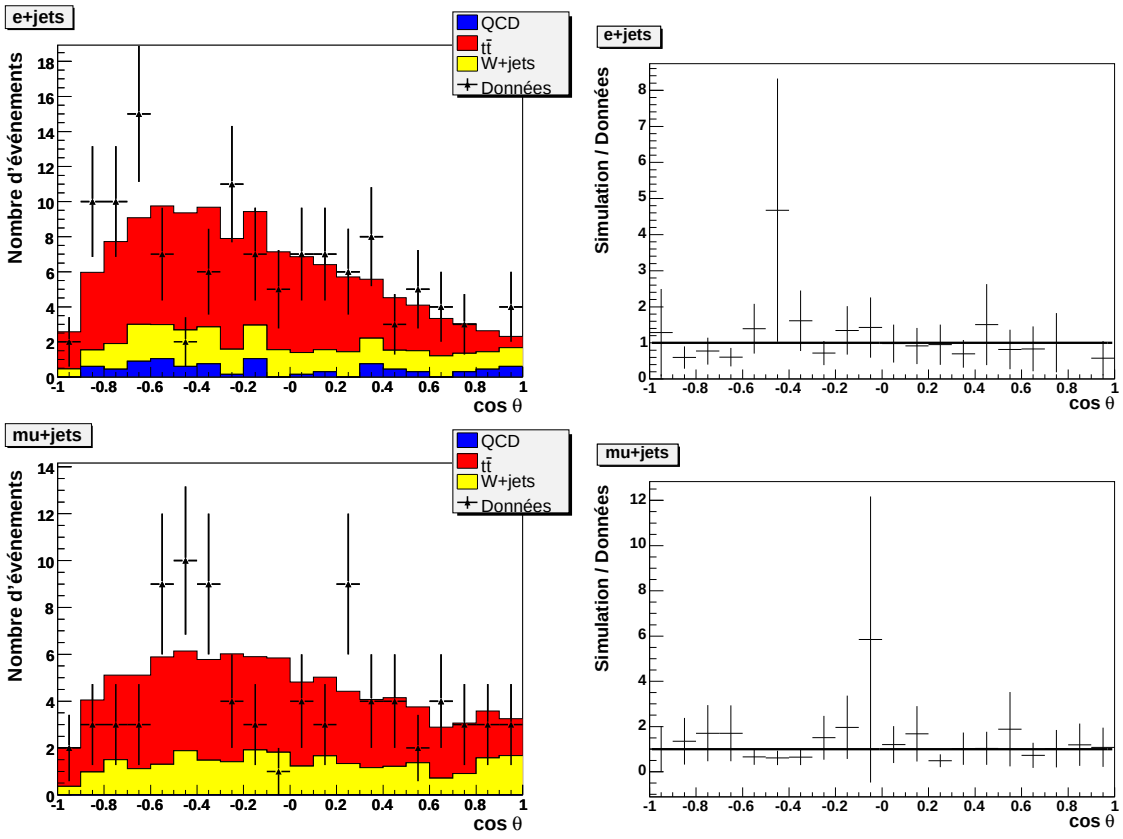

FIG. 5.11: Accord entre les données et la simulation de la distribution de l'observable $\cos\theta$. Les figures de gauche comparent la distribution issue des données à celle créée par la combinaison des distributions des trois composantes $t\bar{t}$, W+jets (simulées) et QCD (issue des données). La composante $t\bar{t}$ utilisée ici correspond à la combinaison des trois lots $t\bar{t}$ simulés avec des fractions d'hélicité purement f_0, f_+ et f_-, respectivement. La contribution de chacun des trois lots correspond à celle mesurée pour le canal correspondant, lorsque l'on a fixé $f_0 = 0.70$ et mesuré f_+. Les figures de droite présentent le rapport de la distribution simulée sur la distribution issue des données.
Les figures du haut correspondent aux distributions obtenues dans le canal e+jets, celles du bas aux distributions dans le canal mu+jets.

quark top par exemple).

Pour estimer l'impact de ces hypothèses sur notre mesure, nous avons utilisé la méthode (déjà présentée) des Ensembles Tests. Nous générons de nombreuses données factices avec les mêmes caractéristiques que les événements servant à notre mesure (proportions de signal et de bruit de fond) en procédant à des tirages aléatoires dans des distributions simulées. Pour simuler ces distributions nous changeons uniquement le pa-

ramètre dont nous voulons estimer l'effet systématique sur la mesure. Les distributions de référence auxquelles nous comparons les données factices, elles, sont inchangées, i.e. exactement les mêmes que celles utilisées pour la mesure dans les données réelles. La modification d'un paramètre change la cinématique de l'événement, ce qui peut induire un effet sur les distributions de $\cos\theta$, sur les distributions de la fonction discriminante (définie au chapitre 4 et permettant d'estimer les proportions de signal et de bruit de fond) et sur les événements sélectionnés. Les effets touchent donc à la fois la forme de la distribution de $\cos\theta$ et les normalisations. Ainsi les effets sur la fonction discriminante sont aussi pris en compte et toute la procédure menant à la mesure est reproduite. Nous générons 800 lots d'événements factices, et procédons ainsi pour chaque paramètre modifié à 800 mesures. La distribution de ces 800 mesures est gaussienne et la moyenne est comparée à la moyenne de la distribution obtenue avec la même méthode en n'ayant modifié aucun paramètre. La différence entre ces deux moyennes est utilisée comme estimation de l'erreur systématique induite par l'incertitude attachée au paramètre modifié.

5.3.2 Estimation des erreurs systématiques

5.3.2.1 Sources d'erreurs systématiques

Dans cette section nous présentons les paramètres que nous avons modifiés pour étudier leur impact sur notre mesure. Les sources d'erreurs systématiques sont les mêmes pour la mesure de f_0 et la mesure de f_+.

• **Statistique finie des distributions de référence :**
Les lots d'événements servant à construire les distributions de référence de l'observable $\cos\theta$ ont une statistique finie, qui contribue à l'erreur systématique sur la mesure des fractions d'hélicité. L'étude de cet effet est différent de celui des autres sources d'erreur systématique. Dans ce cas uniquement, les distributions de référence sont modifiées. Un tirage aléatoire poissonnien est réalisé dans chacune des distributions de référence correspondant aux trois composantes, $t\bar{t}$, W+jets et QCD. La statistique des nouvelles distributions de référence est identique à la statistique des distributions initiales. 1000 nouvelles distributions de référence sont ainsi produites et servent à mesurer les fractions d'hélicité f_0 et f_+ dans les données réelles. La distribution des 1000 mesures est une gaussienne, dont la largeur correspond à l'erreur systématique induite par la statistique limitée des distributions de référence utilisées pour la mesure.

• **Courbe de calibration :**
Le biais observé entre la valeur de fraction d'hélicité du W que nous mesurons et la valeur réelle, implique l'utilisation d'une courbe de calibration pour corriger la mesure brute. L'erreur induite par cette correction est estimée en ne corrigeant que le canal e+jets du biais, ce qui induit une mesure combinée (e+jets + mu+jets) modifiée de Δf_e, puis en ne corrigeant que le canal mu+jets, induisant une modification Δf_μ. L'er-

reur systématique est alors prise comme $\sqrt{(\Delta f_e)^2 + (\Delta f_\mu)^2}$.

• **La correction de l'énergie des jets :**
La *Jet Energy Scale* (cf. section 2.6.3.3) est la correction appliquée à l'énergie des jets, à la fois dans les données et dans la simulation. C'est une des sources principales d'incertitude pour notre mesure, car elle modifie à la fois la forme de la distribution de l'observable $\cos\theta$ mais également les événements sélectionnés. Une modification de $\pm 1\sigma$ est effectuée, et l'énergie transverse manquante de l'événement est recalculée en en tenant compte.

• **L'identification des jets :**
L'efficacité d'identification et de reconstruction des jets dans la simulation est diminuée de 1σ, conduisant au rejet d'un plus grand nombre de jets dans la simulation. Cette efficacité d'identification ne peut pas être augmentée de 1σ car elle dépasse alors 1. Nous en déduisons ainsi l'erreur systématique associée à l'identification des jets dans la simulation.

• **La résolution sur l'énergie des jets :**
Les paramètres utilisés pour corriger la résolution en énergie des jets dans la simulation (cf. section 2.6.3.4) sont modifiés de $\pm 1\sigma$ afin d'estimer l'influence de cette correction sur la mesure finale. Une modification de $+1\sigma$ correspond à une dégradation de la résolution en énergie des jets, tandis qu'une modification de -1σ correspond à une meilleure résolution.

• **La fragmentation des hadrons beaux :**
La fragmentation des hadrons beaux est effectuée par PYTHIA. Elle utilise la fonction de Bowler, dont trois paramètres sont libres. Les simulations utilisées ont été générées avec les valeurs de ces trois paramètres définies par défaut dans PYTHIA. Pour l'étude de cette systématique, la fonction de Bowler a été renormalisée à celle définie en utilisant les paramètres déterminés par les expériences du LEP. La méthode employée est décrite dans la référence [93].

• **L'identification des jets b :**
Comme nous l'avons vu dans la section 3.1.7, l'identification des jets b dans la simulation se fait au moyen d'une pondération des événements par leur probabilité de contenir au moins 1 jet de saveur belle. Ces paramétrisations de l'efficacité d'identification des b par le réseau de neurones (TRF) peuvent être variées de $\pm 1\sigma$, afin d'estimer l'erreur systématique induite par cette incertitude.

• **La modélisation du signal $t\bar{t}$:**
Les événements modélisant le signal $t\bar{t}$ ont été générés en utilisant ALPGEN. Afin d'étudier l'impact du modèle choisi pour la production des états finals des paires $t\bar{t}$, un lot d'événements générés par PYTHIA a été utilisé à la place. D'autre part un lot d'événement généré par PYTHIA, mais pour lequel la paramétrisation de l'*underlying event* a été modifiée (ce lot est appelé $t\bar{t}$ DW), a aussi été utilisé. Deux erreurs systématiques sont estimées ainsi, l'une en comparant PYTHIA avec ALPGEN, l'autre en comparant PYTHIA et DW. L'erreur systématique associée à la modélisation des événements $t\bar{t}$ est prise comme la somme quadratique de ces deux erreurs.

• **La composition en saveurs des processus W+jets simulés :**
La composition en saveurs du bruit de fond W+jets n'est pas connue parfaitement. Afin de déterminer l'influence d'une mauvaise estimation des contributions relatives, nous avons comparé les mesures obtenues en ayant produit des données factices comportant comme bruit de fond W+jets uniquement des processus $Wb\bar{b}jj$ ou uniquement des processus W+jets avec des jets légers. Cette incertitude correspond à l'entrée W_{bb}^{jj} dans les tableaux 5.5 et 5.6.

• **L'échelle de factorisation des processus W+jets :**
L'échelle de factorisation utilisée pour simuler les événements des processus W+jets est $Q^2 = M_W^2 + (P_T^W)^2$. Une erreur systématique induite est estimée en utilisant des lots alternatifs d'événements W+jets générés avec l'échelle de factorisation $Q^2 = M_W^2 + \sum(P_T^{jet})^2$. Cependant, seuls les lots de processus W + saveurs légères sont disponibles avec cette autre échelle de factorisation. Nous comparons donc les distributions obtenues pour chaque hypothèse de factorisation mais en nous restreignant aux processus de saveurs légères. En effet, l'incertitude associée au choix de l'échelle de factorisation est supposée décorrélée de la saveur des processus. L'erreur estimée ainsi correspond à l'entrée $W_P_T^{jet}$ dans les tableaux 5.5 et 5.6.

• **Erreur sur les proportions de bruit de fond et de signal :**
Les normalisations relatives des processus W+jets et QCD déterminées par la *Matrix Method* et par l'ajustement de la fonction discriminante ont été modifiées indépendamment de $\pm 1\sigma$ afin d'étudier l'influence sur la mesure. Pour l'étude de l'influence de la fraction d'événements W+jets, les fractions d'événements $t\bar{t}$ et QCD ont été redeterminées en prenant en compte cette modification. Et inversement lorsque l'influence de la fraction de QCD a été étudiée. Il s'avère néanmoins finalement que les incertitudes engendrées sont totalement négligeables comparées aux autres sources d'erreurs systématiques, et il n'y a pas d'entrées correspondantes dans les tableaux 5.5 et 5.6.

• **La masse du quark top :**
Les distributions de référence sont simulées avec une masse du quark top de 175 GeV/c^2. La courbe de calibration a été construite en comparant des données factices, générées elles aussi avec M_{top} = 175 GeV/c^2, à ces distributions de référence. Dans les données réelles, le quark top a une masse mesurée de 172.6±1.4 GeV/c^2. Pour estimer l'effet induit par cette différence de masse du top, nous avons utilisé des donnes factices simulées avec M_{top} = 170 GeV/c^2. Les différences observées sont :

$$\Delta f_0(175) - \Delta f_0(170) = 0.060$$

$$\Delta f_+(175) - \Delta f_+(170) = 0.034$$

où $\Delta f = f_{mesure} - f_{genere}$. Avec ces deux points nous pouvons corriger l'effet de la masse du top sur notre mesure effectuée dans les données. Les valeurs de f_0 et de f_+ données respectivement dans les sections 5.2.3.1 et 5.2.3.2 deviennent, pour une masse du top de 172.6 GeV/c^2 :

$$f_0(172.6 \text{ GeV/c}^2) = 0.521$$

$$f_+(172.6 \text{ GeV/c}^2) = -0.093$$

Pour estimer l'erreur systématique induite par cette correction de la masse du top, nous propageons aux points 172.6+1.4 GeV/c^2 et 172.6-1.4 GeV/c^2 (correspondant à l'incertitude expérimentale sur la connaissance de la masse du top) l'erreur de la mesure faite à 170 GeV/c^2 avec les données factices (due au nombre fini de lots de données factices). Le point à 175 GeV/c^2 étant par construction affecté d'une erreur nulle (ce point est la référence).

5.3.2.2 Bilan des erreurs systématiques

Le tableau 5.4 récapitule les simulations qui ont dû être modifiées pour étudier une source donnée d'erreur systématique.

Les tableaux 5.5 et 5.6 présentent les résultats de l'étude des erreurs systématiques, respectivement pour la mesure de f_0 et f_+. Ces résultats sont les suivants :

$$\sigma_{syst}(f_0) = ^{+0.083}_{-0.067} \quad \text{et} \quad \sigma_{syst}(f_+) = ^{+0.040}_{-0.033}$$

L'erreur sur les mesures des fractions d'hélicité du W avec 1 fb^{-1} de données du Tevatron sont encore dominées par le manque de statistique. L'erreur systématique est essentiellement due aux incertitudes sur l'échelle d'énergie des jets, le modèle de la fragmentation et la composition en saveur du bruit de fond W+jets.

Source d'erreur	$t\bar{t}$	W+jets	QCD
Jet Energy Scale	X	X	X
JetId	X	X	-
Jet Energy Resolution	X	X	-
TRF	X	X	-
b fragmentation	X	X	-
Modèles du top	X	-	-
Processus w+jets	-	X	-

TAB. 5.4: Les symboles X indiquent les simulations modifiées pour estimer une erreur systématique.

5.4 Estimation d'une limite supérieure sur la fraction d'hélicité droite

Notre méthode de mesure des fractions d'hélicité du W est non bayesienne, c'est à dire que nous n'avons pas cantonné ces fractions dans le domaine physique [0,1], pour ne pas biaiser la mesure. La mesure observée de f_+ se révèle non physique, étant négative. Nous avons alors aussi déterminé une limite supérieure à 95 % de niveau de confiance sur cette fraction, en utilisant une méthode fréquentiste. Pour déterminer cette limite supérieure nous employons la méthode suivante : nous représentons la valeur mesurée (obtenue à partir de lots de données factices) en fonction de la valeur générée, pour chaque valeur générée (cf. figure 5.12). Nous connaissons l'erreur totale $\sigma_{tot} = \sqrt{\sigma_{stat}^2 + \sigma_{syst}^2}$ à partir des erreurs statistique et systématique que nous avons déterminées. A chaque point sur la courbe représentant f_+^{mesure} en fonction de f_+^{genere} nous affectons une erreur de $\pm$ 1.64 σ_{tot} (correspondant à une mesure à 90 % de niveau de confiance, ou 90 % CL). Nous définissons ainsi un domaine suivant la valeur de f_+ correspondant à 90 % de niveau de confiance. La limite supérieure sur notre mesure de f_+ est lue comme l'abscisse de l'intersection de la droite $y = f_+^{mesure}$ et la limite basse du domaine à 95 % CL (i.e. 5 % sont au-dessus et 5 % sont en dessous du domaine défini à 90 % CL).
Nous obtenons ainsi la limite supérieure suivante sur la fraction d'hélicité droite :

$$f_+ < 0.040 \text{ (95 \% CL)}$$

5.5 Conclusion

Nous avons mesuré la fraction f_+ de bosons W transverses droits et la fraction f_0 de bosons W longitudinaux dans des données constituées de paires de quarks top se désintégrant dans les modes semi-leptoniques e+jets et mu+jets. La mesure procède

Source	Δf_0
Courbe de calibration	$\pm$ 0.021
Jet Energy Scale $+1\sigma$	0.046
Jet Energy Scale -1σ	-0.018
JetId	$\pm$ 0.001
Jet Energy Resolusion $+1\sigma$	0.014
Jet Energy Resolution -1σ	0.011
TRF $+1\sigma$	0.024
TRF -1σ	-0.011
Stat. finie des distributions de référence	$\pm$ 0.017
Fragmentation des hadrons beaux	$\pm$ 0.041
Masse du top	+0.017 -0.019
Modélisation des processus $t\bar{t}$	+0.011 -0.010
W_{bb}^{jj}	$\pm$ 0.031
$W_P_T^{jet}$	$\pm$ 0.013
Total	+0.083 -0.067

TAB. 5.5: Erreurs systématiques associées à la mesure de la fraction d'hélicité f_0 de bosons W longitudinaux.

de la minimisation d'une fonction de vraisemblance discrète comparant la distribution de l'observable $\cos\theta$ observée dans les données à celle constituée par la combinaison des processus $t\bar{t}$ et W+jets (simulés) et QCD, où le processus $t\bar{t}$ est composé d'un mélange d'événements comportant des bosons W d'états d'hélicités longitudinale, droite et gauche en proportions variables. Les diverses sources d'erreurs systématiques sont étudiées par la méthode dite des Ensembles Tests. Les événements $t\bar{t}$ utilisés pour la mesure ayant été générés avec une hypothèse de masse du quark top de 175 GeV/c^2, cette mesure a été corrigée par un facteur d'échelle afin d'être ramenée à une hypothèse de $M_{top} = 172.6$ GeV/c^2 correspondant à la valeur mesurée officielle.

Nous mesurons ainsi une fraction d'hélicité longitudinale :

$$f_0 = 0.521^{+0.170}_{-0.171}(\text{stat.})^{+0.083}_{-0.067}(\text{syst.})$$

pour f_+ fixée à sa valeur standard $f_+ = 0.00$, et une fraction d'hélicité tranverse droite :

$$f_+ = -0.093^{+0.099}_{-0.074}(\text{stat.})^{+0.040}_{-0.033}(\text{syst.})$$

pour f_0 fixée à $f_0 = 0.70$. Ces résultats sont en accord avec les prédictions du Modèle Standard.

Source	Δf_+
Courbe de calibration	± 0.012
Jet Energy Scale $+1\sigma$	0.024
Jet Energy Scale -1σ	-0.014
JetId	± 0.004
Jet Energy Resolusion $+1\sigma$	0.005
Jet Energy Resolution -1σ	0.004
TRF $+1\sigma$	0.012
TRF -1σ	-0.010
Stat. finie des distributions de référence	± 0.010
Fragmentation des hadrons beaux	± 0.017
Masse du top	+0.010 -0.011
Modélisation des processus $t\bar{t}$	+0.009 -0.003
W_{bb}^{jj}	± 0.011
$W_P_T^{jet}$	± 0.003
Total	+0.040 -0.033

TAB. 5.6: Erreurs systématiques associées à la mesure de la fraction d'hélicité f_+ de bosons W transverses droits.

Toutefois, la valeur mesurée de f_+ se situe hors de la région physique de la fraction d'hélicité, car notre méthode de mesure est non bayesienne. Nous avons donc déterminé une limite supérieure à 95 % de niveau de confiance sur la valeur de f_+ en utilisant une méthode fréquentiste :

$$f_+ < 0.040 \text{ (95 \% CL)}$$

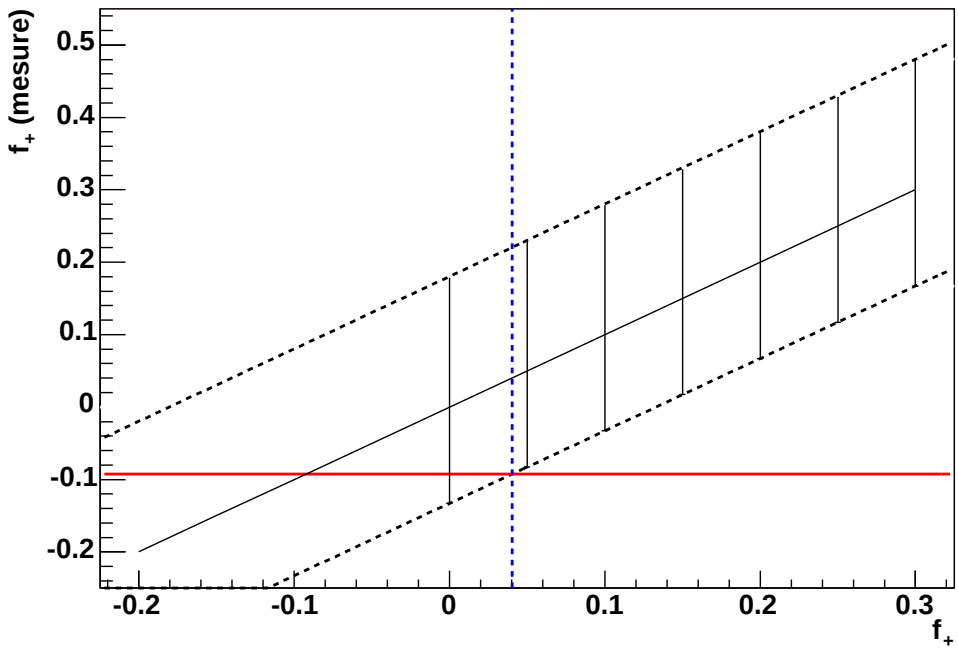

FIG. 5.12: Détermination de la limite supérieure à 95 % CL sur la fraction de bosons W transverses droits.

Conclusion

Ce mémoire de thèse a présenté l'étude du couplage top-Wb par la mesure de l'hélicité du boson W que nous avons effectuée en analysant 1 fb^{-1} de données enregistrées par l'expérience DØ auprès du collisionneur proton-antiproton du Tevatron. Le Tevatron est jusqu'à présent le collisionneur le plus puissant jamais mis en service. Il délivre des collisions $p\bar{p}$ avec une énergie dans le centre de masse de $\sqrt{s} = 1.96$ TeV. Nous avons analysé les données enregistrées entre juillet 2002 et février 2006 au *Run IIa* du Tevatron pour mesurer l'hélicité du W. Cette mesure a été réalisée dans les états finals semi-leptoniques e+jets et mu+jets d'événements $t\bar{t}$.

Pour mesurer les fractions d'hélicité du W nous avons donc sélectionné et reconstruit des paires de quarks top. La topologie des événements correspondants consiste en un lepton, de l'énergie manquante traduisant la présence d'un neutrino échappant à la détection et quatre jets dont deux issus de quarks beaux. Une première étape de sélection consiste en des critères de déclenchement de l'acquisition des données, des coupures sur l'énergie transverse manquante, l'impulsion transverse et la pseudo-rapidité des jets, ainsi que des critères de qualité des leptons.

Les candidats top ainsi préselectionnés peuvent être encore efficacement enrichis en signal en utilisant un algorithme d'identification des jets de quarks beaux basé sur un réseau de neurones. En effet, le quark top se désintègre par interaction faible chargée en un boson W réel et un quark beau, qui s'hadronise en un jet par la suite. Le réseau de neurones combine les informations de divers algorithmes d'identification dont JLIP, développé et maintenu par l'équipe DØ de l'IPHC Strasbourg et qui nécessite une calibration effectuée durant cette thèse. Cette méthode d'identification des jets de quarks beaux est basée sur les paramètres d'impact des traces, qui sont plus élevés dans le cas de traces d'un jet b. Appliquer une coupure sur le réseau de neurones permet ainsi de sélectionner les jets issus de l'hadronisation de quarks b, avec un compromis entre l'efficacité de sélection du signal et l'efficacité de réjection du bruit de fond. Le point de fonctionnement du réseau de neurones a été choisi de manière à maximiser le rapport signal sur bruit de la coupure.

Après cette phase de sélection, le lot d'événements est constitué du signal $t\bar{t}$ mais également de bruits de fond, dont deux types ne sont pas négligeables. Ces bruits de fond

consistent en des événements multijets QCD qui simulent l'état final de notre signal par l'identification erronée d'un faux lepton, ou des processus W+jets présentant le même état final. Nous avons combiné deux méthodes afin de déterminer les contributions relatives de chaque processus, la *Matrix Method* largement employée par la collaboration, et l'ajustement d'une fonction discriminante utilisant un certain nombre de variables topologiques et cinématiques au fort pouvoir discriminant entre le signal $t\bar{t}$ et le bruit de fond W+jets. L'application d'une coupure sur cette fonction discriminante a permis une légère réjection supplémentaire des bruits de fond.

Une fois la sélection des événements achevée et leur contenu en signal et bruit de fond estimé, l'observable $\cos\theta$, où θ est l'angle entre le lepton chargé dans le référentiel de repos du W et le W dans le référentiel de repos du top, est obtenue dans les données par un ajustement cinématique contraint des quadri-impulsions des jets et des leptons. La distribution de $\cos\theta$ est comparée par une méthode de vraisemblance discrète aux distributions théoriques de cette même observable obtenues par des simulations Monte-Carlo pour les contributions du signal $t\bar{t}$ et du bruit de fond W+jets, et déterminée dans les données pour le QCD. Nous mesurons ainsi une fraction d'hélicité longitudinale :

$$f_0 = 0.521^{+0.170}_{-0.171}(\text{stat.})^{+0.083}_{-0.067}(\text{syst.})$$

pour f_+ fixée à sa valeur standard $f_+ = 0.00$, et une fraction d'hélicité tranverse droite :

$$f_+ = -0.093^{+0.099}_{-0.074}(\text{stat.})^{+0.040}_{-0.033}(\text{syst.})$$

pour f_0 fixée à $f_0 = 0.70$. Ces résultats sont en accord avec les prédictions du Modèle Standard. De plus, nous avons déterminé une limite supérieure à 95 % de niveau de confiance sur la valeur de f_+ en utilisant une méthode fréquentiste :

$$f_+ < 0.040 \text{ (95 \% CL)}$$

L'augmentation de la luminosité intégrée delivrée par le collisionneur Tevatron (3.5 fb^{-1} disponibles au printemps 2008) va permettre d'améliorer la sensibilité de cette mesure, ainsi que l'utilisation de plusieurs états finals de la paire de quarks top (semi-leptonique et di-leptonique) et la combinaison des mesures des expériences DØ et CDF. Il sera également possible alors de procéder à une mesure simultanée des trois fractions d'hélicité. Malgré tout il est peu probable qu'une telle mesure au Tevatron permette de mettre en défaut le Modèle Standard de la physique des particules. D'ici fin 2008 le collisionneur LHC devrait entrer en fonction avec une énergie dans le centre de masse 7 fois plus élevée qu'au Tevatron et une luminosité bien plus importante. Cette machine sera une usine à quarks top et elle permettra rapidement une mesure des fractions d'hélicité du W avec une précision atteignant le pourcent qui, si elle est en désaccord avec le Modèle Standard, pourrait dévoiler des processus physiques nouveaux.

Bibliographie

[1] A. Salam *et al*, Phys.Lett. **13**, 168-171, 1964.

[2] S.L. Glashow *et al*, Nucl.Phys. **22** ,579-588, 1961.

[3] S. Weinberg, Phys.Rev.Lett. **19**, 1264,1266, 1967.

[4] M. Gell-Mann, Phys.Rev. **125**, 1067-1084, 1962.

[5] Y. Ne'eman, Nucl.Phys. **26**, 222-229, 1961.

[6] G. Zweig, CERN-TH-412, 74pp, 1964.

[7] D.J. Gross *et al*, Phys.Rev.D **8**, 3633-3652, 1973.

[8] H.D. Politzer, Phys.Rev.Lett. **30**, 1346-1349, 1973.

[9] C.S. Wu *et al.*, Phys. Rev. **105**, 1413-1414, 1957.

[10] M. Goldhaber *et al*, Phys.Rev. **109**, 1015-1017, 1958.

[11] Super-Kamiokande Collaboration (Y. Fukuda *et al.*), Phys.Rev.Lett. **81**, 1562-1567, 1998.

[12] P.W. Higgs, Phys. Lett. **12**, 132, 1964.

[13] F. Englert et R. Brout, Phys. Rev. Lett. **13**, 321, 1964.

[14] LEP Electroweak Working Group, http ://lepewwg.web.cern.ch/LEPEWWG/.

[15] Supernova Cosmology Project Collaboration (S. Perlmutter et al.), Astrophys.J.**517**, 565-586, 1999.

[16] J. Wess et B. Zumino, Phys.Lett.B, **37**, 95, 1971.

[17] Cleo Collaboration M.S. Alam *et al.*, *First measurement of the rate for the inclusive radiative penguin decay* $b \to s\gamma$. Phys.Rev.Lett. **74**, 2885-2889, 1995.

[18] CDF Collaboration (F. Abe *et al.*), Phys.Rev.Lett. **74**, 2626, 1995.

[19] J.L. Hewett V.D. Barger et T.G. Rizzo, Phys.Rev.Lett. **65**, 1313-1316, 1990.

[20] Tevatron Electroweak Working Group pour les Collaborations CDF et DØ, arXivx :0803.1683, mars 2008.

[21] DØ Collaboration (V.M. Abazov *et al.*), *Experimental Discrimination between Charge 2e=3 Top Quark and Charge 4e=3 Exotic Quark Production Scenario*. Phys. Rev. Lett. **98**, 041801, 2007.

[22] M. Cacciari *et al.*, JHEP 0404, 068, 2004.

[23] N. Kidonakis et R. Vogt, Eur.Phys.J.C **33**, 466, 2004.

[24] DØ Collaboration (V.M. Abazov *et al.*), FERMILAB-PUB-04/08-010-E.

[25] Z. Sullivan, Phys.Rev.D **70**, 114012, 2004.

[26] T.M.P. Tait, Phys.Rev.D **61**, 034001, 1999.

[27] DØ Collaboration (V.M. Abazov *et al.*), Phys.Rev.Lett. **98**, 181802, 2007.

[28] CDF Collaboration (T. Aaltonen *et al.*), *Résultats préliminaires pour les conférences d'hiver*. CDF note 9251, mars 2008.

[29] DØ Collaboration (V.M. Abazov *et al.*), FERMILAB-PUB-08/056-E and arXiv :0803.0739.

[30] http ://www.fnal.gov.

[31] http ://www-cdf.fnal.gov.

[32] http ://www-d0.fnal.gov.

[33] DØ Collaboration,
http ://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html.

[34] S. Abachi *et al* (DØ Collaboration), *The DØ detector*. Nucl. Instrum. Methods in Phys. Res. A 338, 185, 1994.

[35] V.M. Abazov *et al* (DØ Collaboration), *The Upgraded DØ detector*. Nucl. Instrum. Methods in Phys. Res. A 565, 463, 2006.

[36] DØCollaboration, *DØ Silicon Tracker Technical Design Report*. http ://d0server1.fnal.gov/projects/silicon/www/tdr_final.ps, juillet 1994.

[37] M. Adams *et al*, *Design Report of the Central Preshower Detector for the DØ Upgrade*. DØ Note 3014, 15 janvier 1996.

[38] V.M. Abazov *et al*, *The muon system of the Run II DØ detector*. Nucl. Instrum. Methods in Phys. Res. A 552 372-398, 2005.

[39] R. Jayanti *et al*, *The DØ Muon System Upgrade*. DØ Note 2780, janvier 1996.

[40] B. Baldin *et al*, *Technical Design of the Central Muon System*. DØ Note 3365, 12 mars 1998.

[41] G. Alexeev *et al*, *Technical Design Report for the DØ Forward Muon Tracking Detector Based on Mini-drift Tubes*. DØ Note 3366, décembre 1997.

[42] V. Abramov *et al*, *Technical Design Report for the DØ forward trigger scintillator counters*. DØ Note 3237, novembre 1997.

[43] F. Deliot, *Reconstruction et identification des muons dans l'expérience DØ. Etude de la production résonnante de sleptons*. Thèse de doctorat, Université Denis Diderot, Paris 7, avril 2002.

[44] B. Gomez, *Time-to-Distance Relationship for the DØ Proportional Drift Tubes PDTs using garfield.* DØ Note 3892, 2001.

[45] C. Luo, *Muon Identification and B Physics Studies at the Tevatron Collider Experiment DØ.* Thèse de doctorat, Indiana University, septembre 2003.

[46] A. Lo *et al*, *Luminosity Monitor Technical Data Report.* DØ Note 3320, 29 septembre 1997.

[47] S. Protopopescu *et al*, *ThumbNail : a compact data format.* DØ Note 3979, 2002.

[48] Sjöstrand *et al*, *PYTHIA 6.2 Physics and Manual.* Comp. Phys. Comm. 135 (2001) 238,, 2001.

[49] M.L. Mangano *et al*, *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions.* JHEP 07 (2003) 001,, 2003.

[50] Y. Fisyak et J. Womersley, *DØ GEANT Simulation of the Total Apparatus Response.* DØ Note 3191, 1997.

[51] S. Agostinelli *et al* (GEANT4 Collaboration), *GEANT4 : A simulation toolkit.* Nucl. Instrum. Meth. A 506, 250, 2003.

[52] G. Hesketh, *Central Track Extrapolation Through the DØ Detector.* DØ Note 4079, 1 octobre 2003.

[53] A. Khanov, *HTF : histogramming method for finding tracks. The algorithm description.* DØ Note 3778, 23 septembre 2000.

[54] H. Greenlee, *The DØ Kalman Track Fit.* DØ Note 4303, 6 janvier 2004.

[55] G. Borissov, *Ordering a Chaos or... Technical Details of AA Tracking.* http ://www-d0.fnal.gov/global_tracking/talks/20030228/talk-adm-030228.ps, 28 février 2003.

[56] A. García-Bellido *et al*, *Primary Vertex certification in p14.* DØ Note 4320, 7 janvier 2004.

[57] A. Schwartzman *et al*, *Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting.* DØ Note 4918, 9 septembre 2005.

[58] Y. Peters *et al*, *Certification of the adaptive primary vertex in p17.* DØ Note 5192, 19 novembre 2006.

[59] E. Busato *et al*, *Jet Algorithms in the DØ Run II Software : Description and User's Guide.* DØ Note 4457, 13 août 2004.

[60] A. Harel, *Jet ID Optimization.* DØ Note 4919.

[61] *Jet Energy Scale.*
http ://www-d0.fnal.gov/phys_id/jes/public/plots_v7.1/index.html.

[62] N. Makovec *et al*, *The Relative Data-Monte Carlo Jet Energy Scale.* DØ Note 4807, 18 juin 2005.

[63] N. Makovec *et al*, *Shifting, Smearing and Removing Simulated Jets.* DØ Note 4914, 9 novembre 2005.

[64] N. Makovec, *Recherche de nouvelle physique dans la topologie à jets et énergie transverse manquante avec le détecteur DØ au Tevatron.* Thèse de doctorat, Université Paris XI Orsay, mai 2006.

[65] F. Beaudette *et al*, *The Road Method (an algorithm for the identification of electrons in jets).* DØ Note 3976, 11 avril 2002.

[66] P. Calfayan *et al*, *Muon Identification Certification for p17 data.* DØ Note 5157, février 2007.

[67] P. Calfayan *et al*, *Muon Id. Group.* http ://www-d0.fnal.gov/computing/algorithms/muon/p17/muonId_quality_type.html.

[68] S. Trincaz-Duvoid *et al*, *Missing ET Reconstruction in p17.* DØ Note 4474, 21 juin 2004.

[69] L. Sawyer *et al*, *Missing ET Reconstruction : Variables and Methods.* DØ Note 3957, 1 juillet 2003.

[70] S. Calvet *et al*, *Towards MissingET Certification and Unclustered Energy Studies.* DØ Note 4927, 27 septembre 2005.

[71] H.A. Bethe, Phys. Rev. 89, 1256, 1953.

[72] W.-M. Yao *et al*, Journal of Physics, G 33, 1, 2006.

[73] V. Siccardi, *Etiquetage des jets de quarks beaux et application à la mesure de la section efficace de production de paires de quarks top dans l'expérience DØ au Tevatron.* Université Louis Pasteur - Strasbourg, septembre 2004.

[74] B. Clément *et al*, *SystemD or how to get signal, backgrounds and their efficiencies with real data.* DØ note 4159, juin 2003.

[75] B. Clement, *Production électrofaible du quark top au Run II de l'expérience DØ.* Thèse de doctorat, Université Louis Pasteur - Strasbourg I, avril 2006. IPHC 06-004 N d'ordre ULP 5086.

[76] D. Bloch *et al*, *Jet Lifetime b-tagging.* DØ note 4069, décembre 2002.

[77] D. Bloch *et al*, *Performance of the Jet LIfetime Probabilty method for b-tagging in DØ data.* DØ note 4158, juin 2003.

[78] D. Bloch *et al*, *Performance of the JLIP b-tagger in p14.* DØ note 4348, mars 2004.

[79] D. Bloch *et al*, *Update of the JLIP b-tagger Performance in p14/pass2 with Jes 5.3.* DØ note 4824, juin 2005.

[80] S. Greder, *Étiquetage des quarks beaux et mesure de la section efficace de production de paires de quarks Top à* $\sqrt{s} = 1.96$ *TeV dans l'expérience DØ.* Thèse de doctorat, Université Louis Pasteur - Strasbourg I, octobre 2004. IRES 05-006 N d'ordre ULP 4652.

[81] T. Scanlon, *A Neural Network b-tagging Tool.* DØ Note 4889, 23 août 2005.

[82] M. Anastasoaie *et al*, *Performance of the NN b-tagging Tool on p17 Data.* DØ Note 5213, 25 janvier 2007.

[83] T. Gadfort *et al*, *Performance of the NN b-tagging Tool on p20 Data.* DØ Note 5554, 2R février 2008.

[84] *Combined CDF and DØ Upper Limits on Standard Model Higgs-Boson Production.* http ://arxiv.org/abs/0712.2383.

[85] J. Kenneth *et al*, *Measurement of the W Boson Helicity in Top Quark Decays Using 1 fb^{-1} of RunII Data.* DØ Note 5313, 2007.

[86] S. Hoche *et al*, *Matching Parton Showers and Matrix Elements.* arXiv :hep-ph/0602031, 2006.

[87] A.F. Barfuss, *Compte-rendu des journées jeunes chercheurs à la Rochelle.* 2006.

[88] S. Park *et al*, *Eefficiency of the Data Quality Calorimeter Flags.* DØ Note 5324, 18 janvier 2007.

[89] S. Snyder, Thèse de doctorat, State University of New York at Stony Brook, Mai 1995.

[90] E. Barberis *et al*, *The Matrix Method and its Error Calculation.* DØ note 4564, 8 août 2004.

[91] M.A. Pleier, *Measurement of the Electron and Muon Fake Rates in Lepton+Jets Datasets.* DØ Note 5469, 2007.

[92] J. Brunelle *et al*, *Measurement of the $t\bar{t}$ Production Cross-Section at $\sqrt{s}$ = 1.96 TeV in the Lepton+Jets Final State using a Topological Method on 1 fb^{-1} of DØ Data.* DØ note 5249, 30 octobre 2006.

[93] Y. Peters *et al*, *Reweighting of the fragmentation function for the DØ Monte Carlo.* DØ Note 5325, 2007.

Institut Pluridisciplinaire Hubert Curien

Service de Reprographie
Michel Schlée

23, rue du Loess
67037 Strasbourg Cedex 2