

# THÈSE

présentée pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ LOUIS PASTEUR STRASBOURG I

*Spécialité : Physique des Particules*

présentée par :

Benoit CLÉMENT

---

### Production électrofaible du quark top au Run II de l'expérience DØ

---

Soutenue publiquement le 28 avril 2006 devant la commission d'examen :

M. Denis GELÉ	Directeur de Thèse
M. Daniel HUSS	Rapporteur interne
M. Eric KAJFASZ	Examineur
M. Claude VALLÉE	Rapporteur externe
M. John WOMERSLEY	Rapporteur externe



# Remerciements

Je tiens à remercier Daniel Huss de m'avoir accueilli à l'IReS (aujourd'hui IPHC) et d'avoir accepté de présider ce jury de thèse

Je remercie les membres du jury Claude Vallée, John Womersley et Eric Kajfasz pour avoir relu ce document et pour leurs commentaires et corrections.

Un grand merci à mon directeur de thèse Denis Gelé pour sa présence, ses conseils, les discussions infinies sur le calcul d'une systématique ou la construction d'un likelihood et les nombreux fous rires.

Je remercie également Arnaud Lucotte (promis j'essaierais de venir faire du single top dans ATLAS) qui a patiemment relu et commenté une partie de cette thèse.

Merci encore à mes collègues de Strasbourg : Isabelle Ripp-Baudot pour les discussions à refaire la physique et le monde autour d'une tasse de thé, Daniel Bloch pour avoir soutenu la plupart des idées farfelues que j'ai pu proposer durant cette thèse (et si on mesurait les fractions de saveurs lourdes dans les données avec la méthode de H1?), François Charles, Abdenour Lounis et Walter Geist ainsi que les étudiants : Sebastien le maître du code, Anne-Catherine qui danse (paraît-il) sur les tables, Jean-Laurent et Vincent le p'tit dernier (bon courage pour la suite)!

Un immense merci à tout ceux dont j'ai croisé la route en France ou à Fermilab et qui ont rendu ces trois années de travail plus supportables : Alexis, Anne-Marie, Aran, Arnaud (D. et G.), Aurélien, Bernard, Brigitte, Boris, Damien, Eric (déjà cité), Florent, Gerald, Gerard, Gordon, Ioanis, Jovan, Jessica, Marine, Manu et Mathieu qui m'ont précédés dans la quête du top célibataire, Miruna, Nikola, Philip, Shabnam, Slava, Steeve, Supriya, Thomas (G. et M.), Samuel, Sabine, et tous les autres.

Je remercie finalement mes parents pour leur soutien constant et leur courageuse relecture de ce document, à la chasse aux fôtes d'ortografe.



# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Panorama du Modèle Standard : une vue du sommet</b>	<b>3</b>
1.1 Introduction au Modèle Standard . . . . .	3
1.2 L'électrodynamique quantique comme exemple de théorie de jauge . . . . .	4
1.2.1 Champs et symétries . . . . .	4
1.2.2 Symétries de jauge . . . . .	5
1.2.3 Développement perturbatif et renormalisation . . . . .	7
1.3 Description du Modèle Standard et physique du quark top . . . . .	8
1.3.1 Propriétés du quark top . . . . .	8
1.3.2 L'interaction forte : chromodynamique quantique . . . . .	8
1.3.2.1 Introduction de la charge de couleur . . . . .	8
1.3.2.2 Le lagrangien de $SU(3)_{couleur}$ . . . . .	9
1.3.2.3 Phénoménologie de QCD . . . . .	10
1.3.2.4 Physique des collisionneurs hadroniques . . . . .	11
1.3.2.5 Production du quark top par interaction forte . . . . .	13
1.3.3 L'interaction électrofaible . . . . .	13
1.3.3.1 Le lagrangien de $SU(2)_L \times U(1)_Y$ . . . . .	14
1.3.3.2 Production et désintégration du quark top par interaction électrofaible . . . . .	15
1.3.3.3 Mesure de la section efficace de production de paires de top . . . . .	18
1.3.4 La masse et le mécanisme de Higgs . . . . .	19
1.3.4.1 Brisure spontanée de la symétrie électrofaible . . . . .	19
1.3.4.2 La masse du quark top et la masse du boson de Higgs . . . . .	21
1.3.4.3 Couplage électrofaible du quark top . . . . .	22
1.3.4.4 Matrice de mélange . . . . .	25
1.4 Conclusion : Le Modèle Standard et au-delà . . . . .	26
<b>2 Le Tevatron et l'expérience DØ</b>	<b>31</b>
2.1 L'accélérateur Tevatron . . . . .	31
2.1.1 Introduction . . . . .	31

2.1.2	Le faisceau de protons . . . . .	34
2.1.3	Production et recyclage du faisceau d'antiprotons . . . . .	34
2.2	L'expérience DØ . . . . .	36
2.2.1	Description générale . . . . .	36
2.2.2	Le détecteur de traces . . . . .	38
2.2.2.1	Le détecteur de vertex ( <i>SMT</i> ) . . . . .	38
2.2.2.2	Le détecteur à fibres scintillantes ( <i>CFT</i> ) . . . . .	40
2.2.2.3	La reconstruction des traces . . . . .	40
2.2.2.4	La reconstruction du vertex primaire . . . . .	42
2.2.3	La calorimétrie . . . . .	45
2.2.3.1	Les détecteurs de pieds de gerbe . . . . .	45
2.2.3.2	Le calorimètre . . . . .	46
2.2.3.3	La région inter-cryostat . . . . .	47
2.2.3.4	La reconstruction des électrons et des photons . . . . .	47
2.2.3.5	La reconstruction des jets . . . . .	49
2.2.3.6	La correction en énergie des jets . . . . .	50
2.2.3.7	L'énergie transverse manquante . . . . .	51
2.2.4	Le système à muons . . . . .	52
2.2.4.1	Le spectomètre à muons . . . . .	52
2.2.4.2	La reconstruction des muons . . . . .	53
2.2.5	La mesure de la luminosité . . . . .	53
2.2.6	Le système de déclenchement . . . . .	55
2.2.7	Le programme de reconstruction et la simulation du détecteur . . . . .	56
<b>3</b>	<b>Identification des jets de quarks beaux</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.2	L'algorithme <i>JLIP</i> . . . . .	61
3.2.1	Significance en paramètre d'impact . . . . .	61
3.2.2	Probabilité de temps de vol par trace et par jet . . . . .	62
3.3	Lots de données pour la calibration et la certification . . . . .	64
3.4	Calibration de l'algorithme <i>JLIP</i> . . . . .	66
3.4.1	Qualité des jets . . . . .	66
3.4.2	Critères de sélection . . . . .	67
3.4.3	Réjection des $V^0$ . . . . .	68
3.4.4	Correction de l'erreur sur le paramètre d'impact . . . . .	69
3.4.5	Fonctions de résolution . . . . .	71
3.5	Certification de l'algorithme <i>JLIP</i> . . . . .	74
3.5.1	La méthode SystemD . . . . .	76
3.5.1.1	Un peu de mathématiques . . . . .	76
3.5.1.2	Application à l'étiquetage des jets de quark $b$ . . . . .	78
3.5.1.3	Test de <i>System8</i> . . . . .	82
3.5.1.4	Estimation des erreurs systématiques . . . . .	82

3.5.2	Efficacité d'identification des saveurs lourdes . . . . .	87
3.5.2.1	Efficacité d'identification dans la simulation . . . . .	87
3.5.2.2	Efficacité d'identification dans les données réelles . . . . .	87
3.5.3	Mesure du taux de mauvaise identification . . . . .	90
3.5.3.1	Efficacité d'étiquetage négatif . . . . .	94
3.5.3.2	Différences entre les lots QCD et EM1TRK . . . . .	94
3.5.3.3	Estimation des erreurs systématiques . . . . .	98
3.5.4	Paramétrisation des efficacités . . . . .	98
3.5.4.1	Factorisation des efficacités . . . . .	98
3.5.4.2	Erreurs systématiques . . . . .	100
3.6	Les autres algorithmes d'identification des jets $b$ à $D\emptyset$ . . . . .	105
3.7	Conclusion . . . . .	105
<b>4</b>	<b>Production du quark top par voie électrofaible à <math>D\emptyset</math></b>	<b>109</b>
4.1	Introduction . . . . .	109
4.2	Stratégie de l'analyse . . . . .	110
4.2.1	Signature des événements <i>single top</i> . . . . .	110
4.2.2	Principaux bruits de fond . . . . .	111
4.2.3	Structure de l'analyse . . . . .	114
4.3	Données et événements simulés utilisés pour l'analyse . . . . .	115
4.4	Sélection initiale . . . . .	116
4.4.1	Critères de déclenchement . . . . .	116
4.4.2	Vertex primaire . . . . .	118
4.4.3	Leptons isolés . . . . .	119
4.4.3.1	Electrons . . . . .	119
4.4.3.2	Muons . . . . .	120
4.4.4	Jets . . . . .	122
4.4.5	Energie transverse manquante . . . . .	122
4.4.6	Coupures triangulaires . . . . .	122
4.5	Etiquetage des jets de quarks beaux . . . . .	124
4.6	Pondération de la simulation . . . . .	127
4.6.1	Facteurs correctifs entre données réelles et simulation . . . . .	127
4.6.1.1	Vertex primaire . . . . .	127
4.6.1.2	Electrons . . . . .	128
4.6.1.3	Muons . . . . .	128
4.6.2	Efficacité de déclenchement . . . . .	128
4.6.2.1	Description générale . . . . .	128
4.6.2.2	Efficacité de déclenchement des jets . . . . .	130
4.6.3	Utilisation des $TRF$ pour l'identification des jets $b$ . . . . .	132
4.6.3.1	Probabilité d'étiquetage par événement . . . . .	132
4.6.3.2	Efficacité de <i>taggabilité</i> . . . . .	135
4.6.3.3	Permutations et identification des jets $b$ étiquetés . . . . .	136

4.7	Estimation du signal et des bruits de fond attendus . . . . .	139
4.7.1	Acceptances des lots simulés . . . . .	139
4.7.2	Normalisation sur les sections efficaces théoriques . . . . .	140
4.7.3	Normalisation à l'aide des données réelles . . . . .	140
4.7.3.1	Mesure du nombre de vrais leptons isolés . . . . .	140
4.7.3.2	Le bruit de fond $W$ +jets . . . . .	144
4.7.3.3	Le bruit de fond multijet . . . . .	149
4.7.4	Nombres d'événements attendus après sélection . . . . .	151
4.8	Analyse multivariable . . . . .	151
4.8.1	Variables discriminantes . . . . .	151
4.8.2	Fonctions de vraisemblance . . . . .	155
4.9	Erreurs systématiques . . . . .	164
4.10	Calcul des limites . . . . .	170
4.10.1	Limites bayésiennes sur la production <i>single top</i> dans les voie $s$ et $t$ . . . . .	170
4.10.2	Limites à deux dimensions et physique au-delà du MS . . . . .	178
4.10.3	Comparaison avec l'analyse par réseaux de neurones de $D\emptyset$ . . . . .	181
4.11	Conclusion . . . . .	182
	<b>Conclusion</b>	<b>185</b>
	<b>A Combinaison des systématiques des fonds <math>W</math>+jets et multijet</b>	<b>187</b>
A.1	L'erreur systématique sur la méthode de la matrice . . . . .	187
A.2	Combinaison des fonds $Wjj$ et $Wbb$ . . . . .	189
A.3	Combinaison des fonds $W$ +jets et multijet . . . . .	190
	<b>B Reconstruction du neutrino</b>	<b>191</b>
	<b>C Histogrammes de contrôle</b>	<b>195</b>
	<b>D Densités de probabilité</b>	<b>219</b>
	<b>E Limite bayésienne</b>	<b>229</b>
E.1	La méthode bayésienne . . . . .	229
E.2	Application à la recherche du <i>single top</i> . . . . .	231
	<b>Liste des figures</b>	<b>233</b>
	<b>Liste des tableaux</b>	<b>237</b>
	<b>Bibliographie</b>	<b>240</b>

*First Witch* —

*When shall we three meet again  
In thunder, lightning, or in rain?*

William Shakespeare, “*Macbeth*”, Act I, Scene 1.

# Introduction

Cette thèse présente les résultats de mon travail au sein du groupe DØ de l’IPHC de Strasbourg. Le détecteur DØ est installé auprès du collisionneur proton-antiproton Tevatron au Fermilab près de Chicago. Il est utilisé par une collaboration internationale de plus de 600 physiciens. Les sujets de recherche abordés dans l’analyse des données de DØ couvrent un large spectre de la physique des hautes énergies, allant de la spectroscopie hadronique (production et propriétés des hadrons beaux) à la recherche de physique au-delà du Modèle Standard (Supersymétrie, dimensions supplémentaires,...) en passant par l’étude du Modèle Standard et de ses limites (production de jets, étude du quark top, recherche du boson de Higgs,...).

Avec une énergie de collision de  $\sqrt{s} = 1.8$  TeV au Run I (1992-1996) puis de  $\sqrt{s} = 1.96$  TeV au Run II (depuis 2001), Tevatron est le seul accélérateur au monde capable de produire le quark top, en attendant le démarrage du *Large Hadron Collider* au Cern en 2007. Ce sixième quark a été mis en évidence au Run I en 1995 avec une masse  $m_{top} = 178.1 \pm 4.1$  GeV/c<sup>2</sup>. Le Run II a confirmé l’existence de cette particule et les mesures précises de ses propriétés se poursuivent. Les prédictions du Modèle Standard sont en accord avec les mesures actuelles.

A ce jour, seule la production par interaction forte de paires de quarks top a été observée. Le Modèle Standard prévoit également la production de quarks top célibataires par interaction électrofaible. Ce mode de production est plus rare que celui par interaction forte et souffre de bruits de fond plus importants. L’observation de ce processus est l’un des objectifs de DØ au Run II. La mesure de sa section efficace de production peut également permettre de sonder les limites du Modèle Standard où le quark top n’intervient dans l’interaction électrofaible par courant chargé que via le vertex  $Wtb$ . Inversement, ce vertex apparaît dans le quasi-unique mode de désintégration du quark top,  $t \rightarrow Wb$ . L’état final contient ainsi un jet issu de la fragmentation du quark  $b$ . L’identification de tels jets est un outil important en physique du top. Par ailleurs les jets de quark  $b$  interviennent dans de nombreux états finals, notamment dans le mode dominant de désintégration d’un boson de Higgs léger,  $H \rightarrow b\bar{b}$ . Les algorithmes d’indentification des jets de quark  $b$  sont donc des outils importants pour l’ensemble de la collaboration DØ.

Le premier chapitre de cette thèse introduit les principaux éléments du Modèle Standard de la physique des particules. A partir de chaque composante du lagrangien, nous tenterons de mettre en évidence les propriétés du quark top étudiable au Tevatron.

La seconde section est consacrée à la description du dispositif expérimental : le complexe d'accélération des protons et des antiprotons ainsi que les différents sous-détecteurs constituant l'expérience DØ .

Le troisième chapitre décrit l'algorithme d'identification des jets de quarks  $b$  développé par le groupe de Strasbourg ainsi que la procédure de validation de l'algorithme pour son utilisation par l'ensemble de la collaboration.

Le dernier chapitre présente l'analyse réalisée à l'été 2005 et visant à améliorer les limites existantes sur les sections efficaces de production électrofaible du quark top. Cette analyse repose, entre autres, sur l'identification des jets de quarks  $b$  décrite précédemment.

Enfin, certains détails techniques ainsi que de nombreux histogrammes de contrôle relatifs à l'analyse sont fournis en annexe.

*Hamlet — O God, I could be bounded in a nut shell and count myself a king of infinite space, were it not that I have bad dreams.*

William Shakespeare, “*Hamlet*”, Act II, Scene 2.

# 1

## Panorama du Modèle Standard : une vue du sommet

### 1.1 Introduction au Modèle Standard

L'ensemble des théories décrivant les interactions entre fermions élémentaires (voir tableau 1.1) des basses énergies à l'échelle du  $\text{TeV}^1$  est communément appelé Modèle Standard de la physique des particules (MS) [1, 2]. Résultat de plusieurs dizaines d'années d'efforts tant théoriques qu'expérimentaux, le MS repose sur le formalisme de la théorie quantique des champs [3]. Aujourd'hui le MS regroupe deux grandes théories : la théorie de Glashow, Salam et Weinberg [4] qui unifie les interactions électromagnétique et faible et la théorie de la chromodynamique quantique décrivant l'interaction forte. Si le MS a obtenu de nombreux succès expérimentaux, il demeure incomplet. Seules les interactions électromagnétique, nucléaire faible et nucléaire forte sont décrites alors que la gravitation de la relativité générale n'est pas formalisable en théorie quantique des champs. L'interaction gravitationnelle reste néanmoins entièrement négligeable en dessous de l'échelle de Planck ( $10^{18}$  GeV). En outre, le MS n'est pas satisfaisant en soi : il ne permet de prédire ni le nombre de fermions ni leur masse, qui restent des paramètres arbitraires, et n'offre pas une unification des interactions électrofaible et forte. D'autres modèles plus complexes mais souvent moins aboutis tels la Supersymétrie, les modèles de dimensions supplémentaires, les modèles de grande unification et les théories de Supercordes sont avancés pour résoudre certains de ces problèmes. Le MS ne serait alors qu'une approximation valable aux basses énergies d'une théorie plus fondamentale, les

---

<sup>1</sup>Tera-electronVolt.  $1 \text{ TeV} = 10^{12} \text{ eV}$  et  $1 \text{ eV} = 1.6 \cdot 10^{-19} \text{ J}$

effets d'une nouvelle physique ne se faisant pas ou peu sentir aux énergies accessibles par les expérimentateurs. L'un des enjeux des expériences auprès des grands collisionneurs (LEP, Tevatron et le futur LHC) est donc une étude détaillée des processus du MS, afin de mettre en évidence d'éventuelles déviations par rapport aux prédictions théoriques, ainsi que la recherche de nouveaux processus mettant en jeu cette nouvelle physique. La physique du quark top, avec sa masse voisine de l'échelle d'énergie électrofaible, est un domaine clé de ce programme de recherche.

Après l'introduction des principes théoriques du MS à travers l'exemple de l'électrodynamique quantique, nous nous attacherons à présenter la physique du quark top dans le MS : modes de production et de désintégration, observables physiques accessibles expérimentalement, en introduisant au fur et à mesure les éléments théoriques nécessaires. Les processus au delà du MS seront évoqués mais sans entrer dans les détails.

Particule		Masse (MeV/c <sup>2</sup> )
Leptons		
Electron	$e$	0.511
Neutrino électron	$\nu_e$	$< 3.10^{-6}$
Muon	$\mu$	105
Neutrino mu	$\nu_\mu$	$< 0.19$
Tau	$\tau$	1777
Neutrino tau	$\nu_\tau$	$< 18.2$
Quarks		
Down	$d$	5 – 8.5
Up	$u$	1.5 – 4
Strange (étrange)	$s$	80 – 155
Charm (charmé)	$c$	$\sim 1300$
Bottom (beau)	$b$	$\sim 4200$
Top	$t$	$172.7 \pm 2.9 \text{ GeV}/c^2$ [6]

TAB. 1.1: Les fermions élémentaires du Modèle Standard [5].

## 1.2 L'électrodynamique quantique comme exemple de théorie de jauge

### 1.2.1 Champs et symétries

Les modèles théoriques en physique sont généralement structurés autour de la conservation, observée expérimentalement, de grandeurs physiques. Le théorème de Noether démontre l'équivalence de cette conservation avec l'invariance du système physique considéré sous certaines symétries. Comme à toute observable conservée dans un processus physique est associée une symétrie, il est possible de construire une théorie à partir

des symétries qui laissent invariantes les équations du mouvement (équations d'Euler-Lagrange). Ceci revient à construire, pour le système, un lagrangien invariant par symétrie. C'est cette idée qui sert de base à la théorie quantique des champs. Les symétries peuvent être séparées en :

- symétries d'espace-temps : rotations d'espace, translations d'espace et de temps, transformations spéciales de Lorentz.
- symétries internes :
  - continues : symétries de jauge.
  - discrètes : parité, conjugaison de charge, renversement du temps.

Mathématiquement, une symétrie continue est associée aux représentations d'un groupe de Lie. Les théories composant le Modèle Standard sont des théories relativistes. Elle doivent donc être, au minimum, invariantes sous le groupe de Lorentz-Poincaré. Les représentations de ce groupe sont caractérisées par le spin du champ, qui est relié à son nombre de composantes. Pour le spin  $\frac{1}{2}$ , il existe deux représentations de dimension 2, qui correspondent aux champs de chiralité gauche et droite, que l'on regroupe en un spineur de Lorentz  $(\chi_L, \chi_R)$ . Le tableau 1.2 présente les lagrangiens invariants de Lorentz pour des champs de différents spins.

Champ	Spin	Particule	Lagrangien
scalaire $\varphi(\mathbf{x})$	0	boson de Higgs	Lagrangien de Klein-Gordon $\mathcal{L}_{KG} = \frac{1}{2}(\partial_\mu \varphi \partial^\mu \varphi - m^2 \varphi^2)$
spineur $\psi(\mathbf{x}) = \begin{pmatrix} \chi_L(\mathbf{x}) \\ \chi_R(\mathbf{x}) \end{pmatrix}$	$\frac{1}{2}$	quarks et leptons	Lagrangien de Dirac $\mathcal{L}_{Dirac} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$
vecteur $A^\mu(\mathbf{x})$	1	bosons de jauge $\gamma, Z, W^\pm$ , gluons	Lagrangien de Maxwell $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ (cas abélien) $\mathcal{L}_{Maxwell} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$

TAB. 1.2: Lagrangiens invariants de Lorentz pour les champs fermioniques et bosoniques usuels.

### 1.2.2 Symétries de jauge

Considérons un champ fermionique (spineur de Lorentz), par exemple celui du quark top  $\psi_t(\mathbf{x})$ . Le champ libre doit satisfaire au lagrangien de Dirac. De plus, puisqu'il s'agit d'une théorie quantique, on voudrait pouvoir effectuer un changement de phase globale  $\theta$  du champ (i.e. indépendant des coordonnées d'espace-temps  $\mathbf{x}$ ) sans en modifier les propriétés physiques. L'ensemble de ces transformations forment le groupe  $U(1)$ .

$$\psi_t(\mathbf{x}) \xrightarrow{U(1)_{global}} \psi'_t = e^{+i\theta} \psi_t(\mathbf{x}) \tag{1.1}$$

$$\bar{\psi}_t(\mathbf{x}) \xrightarrow{U(1)_{global}} \bar{\psi}'_t = \bar{\psi}_t(\mathbf{x}) e^{-i\theta} \tag{1.2}$$

$$\text{avec } \bar{\psi}_t(\mathbf{x}) = \psi_t^\dagger(\mathbf{x}) \gamma_0 \tag{1.3}$$

La phase ne dépend pas des coordonnées d'espace-temps  $x^\mu$ . Ainsi  $e^{+i\theta}$  commute avec l'opérateur de dérivation  $\partial_\mu$ . Cette transformation laisse ainsi le lagrangien invariant. Il semble naturel de demander que l'invariance sous la symétrie  $U(1)$  soit locale (invariance de jauge) et donc que la phase dépende de  $x_\mu$ .

$$\psi_t(\mathbf{x}) \xrightarrow{U(1)_{local}} e^{+i\theta(\mathbf{x})}\psi_t(\mathbf{x}) \quad (1.4)$$

$$\begin{aligned} \partial_\mu \psi_t \xrightarrow{U(1)_{local}} (\partial_\mu \psi_t)' &= (\partial_\mu e^{+i\theta})\psi_t + e^{+i\theta}(\partial_\mu \psi_t) \\ &= e^{+i\theta}(\partial_\mu + i\partial_\mu \theta)\psi_t \end{aligned} \quad (1.5)$$

Dès lors, l'opérateur de symétrie ne commute plus avec l'opérateur de dérivation et il est nécessaire d'introduire un nouveau champ vectoriel  $A_\mu$  pour annuler le commutateur. La dérivée covariante  $D_\mu$  est ainsi définie par :

$$D_\mu = \partial_\mu + iqA_\mu \quad , \quad \text{telle que } (D_\mu \psi_t)' = e^{+i\theta}(D_\mu \psi_t) \quad (1.6)$$

où  $q$ , la charge de l'électron, est la constante de couplage électromagnétique et

$$A_\mu \xrightarrow{U(1)_{local}} (A_\mu)' = A_\mu - \frac{1}{q}\partial_\mu \theta \quad (1.7)$$

En ajoutant un terme cinétique pour le nouveau champ  $A_\mu$  libre, le lagrangien devient finalement :

$$\begin{aligned} \mathcal{L}_{Dirac}^{QED} &= \bar{\psi}_t(i\gamma^\mu D_\mu - m)\psi_t - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \underbrace{\bar{\psi}_t(i\gamma^\mu \partial_\mu - m)\psi_t}_{\text{champ } \psi_t \text{ libre}} + \underbrace{q\bar{\psi}_t\gamma^\mu\psi_t A_\mu}_{\text{couplage entre } \psi_t \text{ et } A_\mu} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{champ } A_\mu \text{ libre}} \end{aligned} \quad (1.8)$$

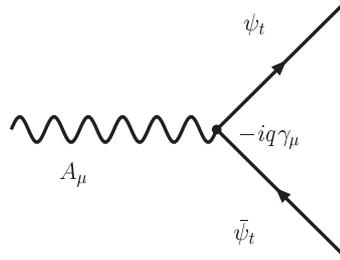


FIG. 1.1: Vertex QED couplant le quark top au photon.

Le champ vectoriel  $A_\mu$ , de masse nulle, est associé au photon. Imposer l'invariance de jauge sous  $U(1)_{local}$  fait naturellement apparaître le boson de jauge de l'électromagnétisme. Le seul couplage possible est déterminé par le terme d'interaction. Le diagramme de Feynman correspondant est donné par la figure 1.1. Le seul mode de production du quark top par couplage électromagnétique est ainsi une production par paire.

### 1.2.3 Développement perturbatif et renormalisation

La résolution des équations d'Euler-Lagrange se fait généralement dans le cadre d'un développement perturbatif en puissance de la constante de couplage. L'ordre dans le développement perturbatif d'un diagramme de Feynman correspond au nombre de vertex. Le vertex de la figure 1.1 n'est valable qu'à l'ordre 0. Pour comparer des calculs théoriques aux mesures de précisions expérimentales, il est nécessaire de prendre en compte des corrections d'ordre supérieur pour les vertex et les propagateurs. Ces corrections au premier ordre sont représentées par la figure 1.2. Ces corrections font apparaître des

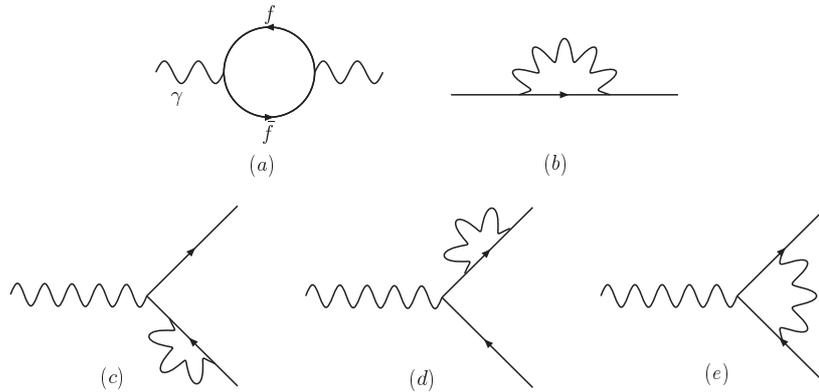


FIG. 1.2: Corrections radiatives au premier ordre pour : (a) le propagateur du photon, (b) le propagateur d'un fermion chargé, (c-e) le vertex photon-fermions.

boucles dont le calcul fait intervenir des intégrales divergentes car la quadri-impulsion des particules dans la boucle n'est pas contrainte. Une procédure analytique complexe, la renormalisation, permet d'éliminer ces divergences. Pour un processus à une échelle d'énergie donnée, les contributions de toutes les boucles aux ordres supérieurs sont absorbées dans la constante de couplage. La constante de couplage initiale est ainsi remplacée par un couplage effectif qui dépend de l'échelle d'énergie du processus. L'évolution avec l'échelle d'énergie des constantes de couplage du Modèle Standard est illustrée sur la figure 1.3. Les boucles fermioniques ont un effet d'écrantage qui diminue l'intensité de la force quand l'énergie décroît. Dans le cadre plus général d'une théorie de jauge non abélienne, il existe aussi des boucles bosoniques, qui ont une contribution de signe opposé aux boucles de fermions. C'est le cas pour les interactions forte et faible qui seront discutées aux § 1.3.2 et 1.3.3.

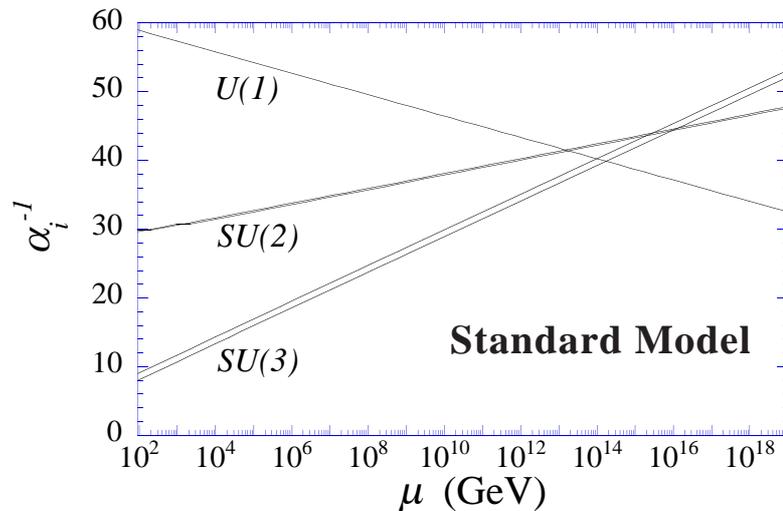


FIG. 1.3: Evolution de l'inverse des constantes de couplage électrofaibles ( $U(1)$  et  $SU(2)$ ) et forte ( $SU(3)$ ) avec l'énergie.

## 1.3 Description du Modèle Standard et physique du quark top

### 1.3.1 Propriétés du quark top

L'intérêt principal du quark top provient de sa masse élevée, voisine de l'énergie typique de la brisure de symétrie électrofaible. Ceci lui confère des propriétés uniques, mais rend également sa production expérimentalement difficile. Découvert en 1995 [8], il demeure la particule du Modèle Standard dont les propriétés sont les plus mal connues. La physique du quark top est donc un domaine unique pour tester les limites du modèle et sonder la nouvelle physique, si elle existe. Cette section vise à présenter les différentes théories composant le Modèle Standard et d'en déduire les modes de production et les propriétés du quark top ainsi les observables pertinentes de la physique du top [7].

### 1.3.2 L'interaction forte : chromodynamique quantique

#### 1.3.2.1 Introduction de la charge de couleur

Le modèle des quarks fut introduit en 1963 par Gell-Mann [9] pour expliquer la grande multiplicité des états hadroniques observés. Un méson est un état lié quark-antiquark et un baryon se compose de trois quarks. Le modèle élaboré à partir de 3 quarks (up, down et étrange) suffisait à reconstituer tous les spectres hadroniques et à prédire de nouvelles particules. La découverte en 1964 du baryon  $\Omega$  [10] confirma la validité de ce modèle. Or la particule  $\Omega$  est un état lié  $sss$  de spin  $\frac{3}{2}$ . De ce fait les

trois quarks étranges portent les mêmes nombres quantiques, ce qui est interdit par le principe de Pauli. Il faut donc que chacun des ces quarks possède un nombre quantique supplémentaire différent, conservé par l'interaction forte : c'est la charge de couleur (par convention : rouge, vert ou bleu). La charge de couleur augmente également le nombre de degrés de liberté disponibles lors de la production de hadrons. Elle peut ainsi être mise en évidence en comparant les taux de hadrons et de leptons produits dans des collisions  $e^+e^-$  ou lors de la désintégration du lepton  $\tau$  par exemple.

### 1.3.2.2 Le lagrangien de $SU(3)_{couleur}$

La Chromodynamique Quantique (QCD) est la théorie de jauge décrivant la conservation de la couleur dans l'interaction forte. Le groupe de symétrie correspondant le plus simple est  $SU(3)$ . Contrairement à  $U(1)$  qui intervient en QED, ce groupe n'est pas commutatif. Il possède huit générateurs  $\lambda_{a=1..8}$  (matrices de Gell-Mann) et son algèbre de Lie est donnée par  $[\lambda_a, \lambda_b] = f_{abc}\lambda_c$  où les  $f_{abc}$  sont les constantes de structure du groupe. Pour rappeler que cette symétrie conserve la couleur, le groupe est généralement noté  $SU(3)_c$ . Chaque quark (ici le quark top) peut alors être décrit par un triplet complexe de  $SU(3)_c$  :

$$\Psi_t(\mathbf{x}) = \begin{pmatrix} \psi_t^{rouge} \\ \psi_t^{vert} \\ \psi_t^{bleu} \end{pmatrix} \quad (1.9)$$

La dérivée covariante est introduite de manière analogue au chapitre 1.2.2. Pour conserver l'invariance locale sous  $SU(3)_c$  huit bosons de jauge, autant que de générateurs du groupe, sont introduits. Ce sont les gluons  $G_\mu^a$  :

$$D_\mu = \partial_\mu + ig_s \lambda_a G_\mu^a \quad , \text{ où } g_s \text{ est la constante de couplage fort.} \quad (1.10)$$

De même le terme cinétique pour les gluons s'obtient en généralisant le cas de QED. La chromodynamique est une théorie non abélienne, les gluons étant eux-mêmes colorés. Ceci introduit des termes supplémentaires dans le lagrangien du champ des gluons libres et donc de nouveaux couplages. Finalement le lagrangien total pour un seul quark s'écrit :

$$\begin{aligned} \mathcal{L}_{Dirac}^{QCD} &= \bar{\Psi}_t(i\gamma^\mu D_\mu - m)\Psi_t + \mathcal{L}_{gluon} \\ &= \underbrace{\bar{\Psi}_t(i\gamma^\mu \partial_\mu - m)\Psi_t}_{\text{quarks libres}} + \underbrace{g_s \bar{\Psi}_t \gamma^\mu \lambda_a \Psi_t G_\mu^a}_{\text{couplage entre quarks et gluon}} - \underbrace{\frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}}_{\text{gluons libres}} \quad (1.11) \end{aligned}$$

$$\text{avec } G_{\mu\nu}^a = \underbrace{\partial_\mu G_\nu^a - \partial_\nu G_\mu^a}_{\text{terme analogue à QED}} - \underbrace{f_{bc}^a G_\mu^b G_\nu^c}_{\text{terme non-abelien}} \quad (1.12)$$

La partie non abélienne du lagrangien fait apparaître des termes de la forme  $\partial GGG$  et  $GGGG$ , soit des vertex à trois ou quatre gluons, présentés dans la figure 1.4. Contrairement à QED, les bosons de jauge portent également une charge de couleur (couleur et anticouleur) et sont eux-même sujets à l'interaction forte.

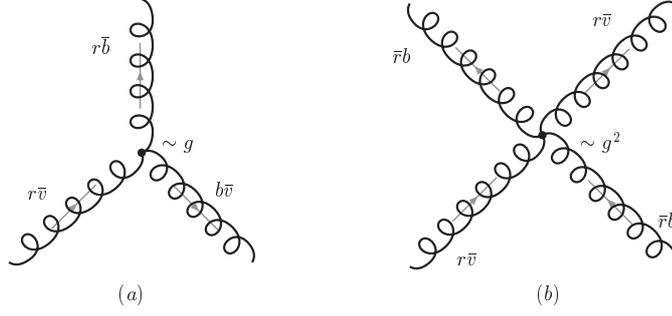


FIG. 1.4: Exemples de vertex d'auto-couplage à trois (a) et quatre (b) gluons.

### 1.3.2.3 Phénoménologie de QCD

Dans le domaine d'énergie où la théorie QCD est perturbative, l'évolution de la constante de couplage fort en fonction de l'impulsion transférée  $Q^2$  peut s'écrire en fonction de la constante de couplage à une échelle d'énergie fixée  $\mu^2$  :

$$\alpha_s = \frac{g_s}{4\pi} \quad (1.13)$$

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{4\pi} b_0 \ln\left(\frac{Q^2}{\mu^2}\right)}, \text{ avec } b_0 = -\frac{2}{3}N_f + \frac{11}{3}N_c [2] \quad (1.14)$$

où  $N_f$  est le nombre de saveurs et  $N_c$  le nombre de couleurs. Dans le Modèle Standard  $N_f = 6$  et  $N_c = 3$ , et donc  $b_0 > 0$ . A grand moment transféré, c'est-à-dire à courte distance la constante de couplage fort tend vers 0. Ce phénomène est connu sous le nom de liberté asymptotique. Un quark isolé polarise le vide autour de lui, créant des paires virtuelles de quarks et de gluons. Mais les boucles de gluons, porteurs d'une charge de couleur, créent un anti-écranage si bien que la charge de couleur effective diminue avec la distance. Inversement à petit  $Q^2$  ( $Q^2 \approx m_{quarks}^2$ ) la constante de couplage augmente et la théorie n'est plus perturbative. Elle ne peut plus se développer en terme de diagrammes de Feynman. L'intensité de l'interaction entre quarks augmente avec la distance qui les sépare. En conséquence un quark ou un gluon libre ne peut exister, car il a tendance à attirer vers lui toute autre particule colorée jusqu'à annulation de la charge. Les seuls états liés stables doivent être "blancs" : c'est le confinement. Les systèmes les plus simples sont constitués d'une paire quark-antiquark (mésone :  $M = \frac{1}{\sqrt{3}}\delta^{\alpha\beta}|q_\alpha\bar{q}_\beta\rangle$  où  $\alpha, \beta = r, v, b$ ) ou de trois quarks (baryon :  $B = \frac{1}{\sqrt{6}}\epsilon^{\alpha\beta\gamma}|q_\alpha q_\beta q_\gamma\rangle$ ). Le confinement est un effet non-perturbatif et il est très complexe de décrire exactement un état lié de plusieurs quarks.

### 1.3.2.4 Physique des collisionneurs hadroniques

La compréhension de QCD est fondamentale pour décrire aussi bien l'état initial que l'état final des processus étudiés auprès d'un collisionneur hadronique. Le Tevatron est un collisionneur proton-antiproton fonctionnant à une énergie dans le centre de masse  $\sqrt{s} = 1.96$  TeV. Cette énergie correspond à une distance caractéristique de l'ordre de  $10^{-18}$  m soit inférieure de trois ordres de grandeur au rayon classique du proton (environ  $10^{-15}$  m). A cette échelle, QCD est une théorie perturbative. Une interaction à cette échelle d'énergie est donc sensible à la structure interne du proton. Ce sont les constituants du proton et de l'antiproton, appelés partons qui interagissent. Ces partons peuvent être :

- un quark de valence : l'un des trois quarks constituant le proton,
- un quark de la mer : l'un des quarks virtuels du vide QCD polarisé,
- un gluon.

La section efficace totale d'un tel processus est donnée par [11] :

$$\sigma(p\bar{p} \rightarrow X) = \sum_{i,j} \int_0^1 \int_0^1 f_i^p(x_1) f_j^{\bar{p}}(x_2) \hat{\sigma}(ij \rightarrow X) dx_1 dx_2 \quad (1.15)$$

Cette expression fait intervenir :

- $\hat{\sigma}(ij \rightarrow X)$  : la section efficace du processus faisant intervenir les partons  $i$  et  $j$ . Cette section efficace est calculable analytiquement en QCD perturbative.
- $f_i^{p(\bar{p})}(x)$  : la densité de parton de type  $i$  dans le proton (antiproton). Les densités de parton (*Parton Density Functions, PDF*) dépendent de la fraction  $x$  de l'énergie totale du proton portée par le parton. Elles sont déterminées à partir de différentes mesures de diffusion profondément inélastique  $e + p \rightarrow X$ . Plusieurs collaborations de phénoménologues telles CTEQ [12] ou MRST [13] fournissent des ajustements de ces *PDF* aux données de multiples expériences sur cibles fixes et sur collisionneurs. La figure 1.5 montre les *PDF* déterminée par la collaboration CTEQ [12]. L'erreur sur la mesure des PDFs est l'une des principales incertitudes dans le calcul des sections efficaces théoriques des processus physiques étudiés. Ces incertitudes sont particulièrement grandes pour les gluons à grand  $x$ . Ceci est illustré sur la figure 1.6. La contribution des saveurs lourdes (quarks  $b$  et  $c$ ) aux *PDF* est très faible. Les rares mesures existantes [14] sont en accord avec les modèles théoriques mais souffrent de large incertitudes.

La QCD intervient de nouveau pour expliquer les états finals observés dans le détecteur. Les quarks et les gluons de haute énergie produits au cours de l'interaction ne peuvent pas demeurer libres à cause du confinement. Ils polarisent le vide autour d'eux jusqu'à fournir suffisamment d'énergie pour créer une nouvelle paire quark-antiquark. De proche en proche, le parton initial cède son énergie pour produire une multitude de hadrons instables associés dont la désintégration produit un jet. Au niveau du détecteur ces jets sont caractérisés par plusieurs traces chargées collimées et un ou plusieurs amas d'énergie dans le calorimètre. Ce processus de fragmentation et d'hadronisation est extrêmement

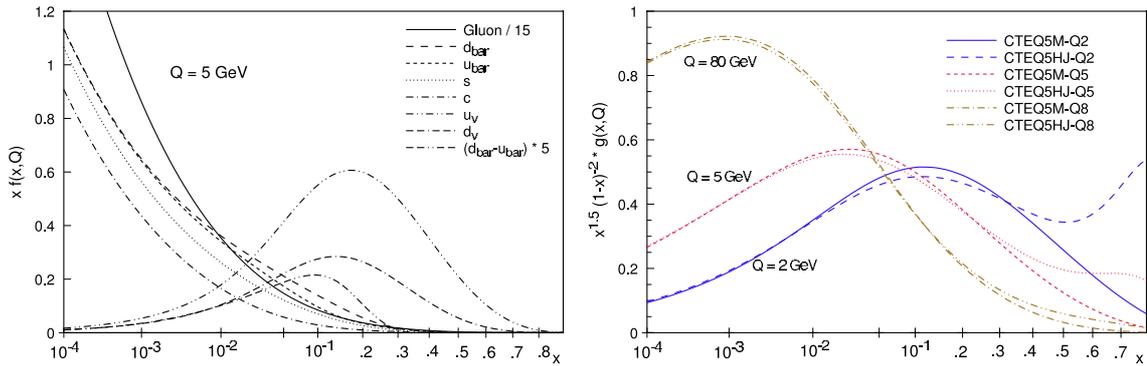


FIG. 1.5: Fonctions de distribution de parton de la collaboration CTEQ.

A gauche : *PDF* des différents partons (CTEQ5M) pour une impulsion transférée  $Q = 5 \text{ GeV}/c$ . A droite : *PDF* du gluon (CTEQ5M en traits pleins et CTEQ5HQ en pointillés) pour trois impulsions transférées  $Q = 2 \text{ GeV}/c$  (bleu),  $5 \text{ GeV}/c$  (rouge) et  $80 \text{ GeV}/c$  (marron).

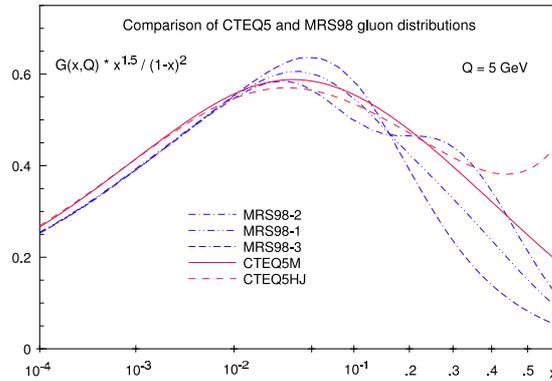


FIG. 1.6: Fonction de densité du gluon pour différentes paramétrisations des collaborations CTEQ (5M et 5HQ) et MRST (98-1, 98-2, 98-3). L'important désaccord observé à grand  $x$  traduit la large incertitude sur la *PDF* du gluon dans cette région.

complexe et ne peut être calculé par un développement perturbatif. Il existe cependant différents modèles phénoménologiques décrivant l'hadronisation. Ce type de modèle est utilisé dans des programmes de simulation Monte-Carlo comme PYTHIA [15] (modèle de Lund) ou HERWIG [16]. Un quark reste libre environ  $10^{-23} \text{ s}$  avant de s'hadroniser. Du fait de sa masse très élevée, la durée de vie du quark top ( $\sim 10^{-24} \text{ s}$ ) est plus courte que ce temps d'hadronisation. Le quark top est donc le seul quark se désintégrant avant d'avoir pu s'hadroniser. Ainsi, il n'existe pas de hadron contenant un quark top.

### 1.3.2.5 Production du quark top par interaction forte

Il nous est désormais possible d'introduire les différents diagrammes de Feynman correspondant à la production d'une paire de quarks top par interaction forte illustrés par la figure 1.7. Il s'agit du mode dominant de production du quark top au Tevatron dont la section efficace théorique est de  $6.77 \pm 1.22$  pb [17, 18] à  $\sqrt{s} = 1.96$  TeV. Au Tevatron, les processus de fusion de gluons ne représentent que 15% de la section efficace totale. La mesure de cette section efficace sera présentée au § 1.3.3.3.

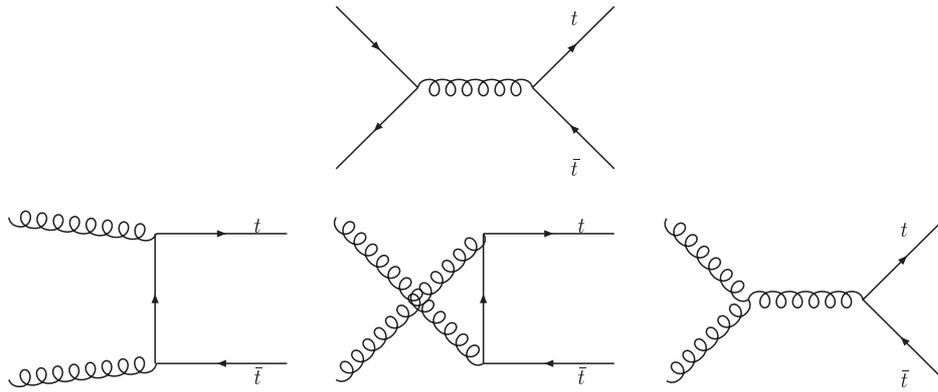


FIG. 1.7: Production de paire de quarks top par interaction forte : annihilation de quarks (en haut) et fusion de gluons (en bas)

### 1.3.3 L'interaction électrofaible

L'électrodynamique quantique a été introduite au § 1.2 à titre illustratif. Dans le Modèle Standard, électromagnétisme et interaction nucléaire faible sont décrites dans un même formalisme. La découverte en 1956 par C.S. Wu [19] de la violation de parité dans les désintégrations  $\beta$  montra que l'interaction faible se comporte différemment vis à vis des deux projections de chiralité des champs fermioniques. Pour un champ spinoriel  $\psi(x)$ , les deux projections de chiralité gauche et droite sont définies par :

$$\psi(x) = \psi_L(x) + \psi_R(x) , \text{ avec } \psi_L(x) = \frac{1 - \gamma^5}{2}\psi(x) \text{ et } \psi_R(x) = \frac{1 + \gamma^5}{2}\psi(x) \quad (1.16)$$

L'expérience montre que cette violation de parité est maximale : les courants chargés de l'interaction faible ne se couplent qu'aux fermions gauches (anti-fermions droits). Les champs gauches d'une même famille peuvent être groupés en un doublet. Le système ainsi formé est analogue à celui d'un spin  $\frac{1}{2}$ . Par analogie, on introduit un nouveau nombre quantique conservé, l'isospin faible, associé à une nouvelle symétrie. Chaque doublet possède un isospin  $T = \frac{1}{2}$  et chaque champ correspond à une projection de l'isospin,

$T_3 = \pm\frac{1}{2}$ . Le groupe de symétrie associé à la conservation de l'isospin est  $SU(2)$ , noté généralement  $SU(2)_L$  pour rappeler que cette symétrie ne concerne que les fermions gauches. Les champs droits sont des singlets pour cette symétrie et possèdent donc un isospin nul. Dans le Modèle Standard, les neutrinos, de masse nulle, n'interagissent que par interaction faible. Comme cette dernière ne se couple qu'aux champs gauches il ne peut exister de neutrino droit. Les interactions faible et électromagnétique peuvent être unifiées en introduisant une autre grandeur scalaire conservée, dépendant de la charge électrique  $Q$  avec une valeur unique pour chaque doublet. C'est l'hypercharge  $Y$ , donnée par la relation de Gell-Mann et Nishijima.

$$Y = 2(Q - T_3) \quad (1.17)$$

Les valeurs de  $T_3$  et  $Y$  pour les fermions du Modèle Standard sont résumées dans le tableau 1.3.

	Leptons	$T_3$	$Y$	Quarks	$T_3$	$Y$
doublets	$(\nu_e)_L, (\nu_\mu)_L, (\nu_\tau)_L$ $(e)_L, (\mu)_L, (\tau)_L$	$(+\frac{1}{2})$ $(-\frac{1}{2})$	-1	$(u)_L, (c)_L, (t)_L$ $(d)_L, (s)_L, (b)_L$	$(+\frac{1}{2})$ $(-\frac{1}{2})$	$\frac{1}{3}$ $\frac{1}{3}$
singlets	$e_R, \mu_R, \tau_R$	0	-2	$u_R, c_R, t_R$ $d_R, s_R, b_R$	0 0	$+\frac{4}{3}$ $-\frac{2}{3}$

TAB. 1.3: Isospin faible et hypercharge des fermions du Modèle Standard.

Finalement, la théorie de jauge des interactions électrofaibles est construite en imposant l'invariance sous les transformations locales du groupe  $SU(2)_L \times U(1)_Y$ .

### 1.3.3.1 Le lagrangien de $SU(2)_L \times U(1)_Y$

Le lagrangien du modèle de Glashow, Salam et Weinberg [4], se construit comme pour QED et QCD.  $SU(2)_L$  possède trois générateurs  $T_a$ , ce qui introduit trois bosons de jauge  $I_\mu^a$ . Un quatrième boson  $B_\mu$  provient de la symétrie  $U(1)_Y$ . Pour une seule famille de quarks, il existe un doublet gauche  $Q_L = \begin{pmatrix} u \\ d \end{pmatrix}_L$  et deux singlets droits  $b_R$  et  $t_R$ . Le lagrangien s'écrit alors :

$$\mathcal{L}_{Dirac}^{EW} = \bar{Q}_L(i\gamma^\mu D_\mu)Q_L + \bar{t}_R(i\gamma^\mu D_\mu)t_R + \bar{b}_R(i\gamma^\mu D_\mu)b_R - \frac{1}{4}I_{\mu\nu}^a I^{\mu\nu a} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad (1.18)$$

avec

$$D_\mu = \partial_\mu + igT_a I_\mu^a + i\frac{g'}{2}Y B_\mu \quad (1.19)$$

où  $g$  et  $g'$  sont les constantes de couplage de  $SU(2)_L$  et  $U(1)_Y$ .

Les tenseurs  $I_{\mu\nu}^a$  et  $B_{\mu\nu}$  sont définis de manière analogue à QCD et QED. La théorie doit également décrire l'électromagnétisme, il faut donc faire apparaître le champ du

photon. Comme aucun des champs  $I_\mu^a$  et  $B_\mu$  ne correspond au photon, ce ne sont pas des champs "physiques". Le photon ne modifiant pas la saveur des quarks, il ne peut être qu'une combinaison linéaire de  $I_\mu^3$  et  $B_\mu$ . De même le couplage up/down est une combinaison linéaire de  $I_\mu^1$  et  $I_\mu^2$ . Les champs physiques sont finalement définis par :

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W I_\mu^3 \quad (1.20)$$

$$Z_\mu = -\sin \theta_W B_\mu + \cos \theta_W I_\mu^3 \quad (1.21)$$

$$W_\mu^+ = \frac{I_\mu^1 - iI_\mu^2}{\sqrt{2}} \quad (1.22)$$

$$W_\mu^- = \frac{I_\mu^1 + iI_\mu^2}{\sqrt{2}} \quad (1.23)$$

où  $\theta_W$  est l'angle de mélange de Weinberg. Le champ  $A_\mu$  décrivant le photon, sa constante de couplage est celle de QED. On a alors :

$$\tan \theta_W = \frac{g'}{g} \text{ et } g' \cos \theta_W = g \quad (1.24)$$

Par construction, les bosons de jauge sont de masse nulle, or l'expérience montre que les bosons  $Z$  et  $W$  sont massifs, ce qui explique la courte portée de l'interaction faible. Ces bosons ont été mis en évidence expérimentalement en 1983 au CERN par les expériences  $UA1$  et  $UA2$  [20, 21]. Leurs masses [5] ont été mesurées avec précision au  $LEP$  :

$$m_Z = 91.1876 \pm 0.0021 \text{ GeV}/c^2 \quad (1.25)$$

$$m_W = 80.425 \pm 0.038 \text{ GeV}/c^2 \quad (1.26)$$

De plus, contrairement au cas de QED et QCD, le terme de masse  $-m\bar{\psi}\psi$  du lagrangien de Dirac viole à la fois  $SU(2)_L$  et  $U(1)_Y$  car il couple fermions gauche et droit. Ainsi le lagrangien électrofaible ne décrit que des fermions de masse nulle. Puisque les masses des fermions et des bosons de jauge  $W$  et  $Z$  sont non nulles, il faut introduire un terme dans le lagrangien électrofaible qui brise la symétrie  $SU(2)_L \times U(1)_Y$  et fait apparaître les termes de masse. C'est la brisure spontanée de la symétrie électrofaible décrite au § 1.3.4 qui permet d'introduire les masses des bosons de jauge et des fermions.

### 1.3.3.2 Production et désintégration du quark top par interaction électrofaible

Bien qu'incomplet, le lagrangien 1.19 permet déjà d'introduire de nouveaux vertex mettant en jeu le quark top (Fig. 1.8). Le couplage au photon et au  $Z$ , fait de nouveau apparaître deux quarks top. Ces vertex pourraient être considérés pour la production par paires du quark top, mais dans une collision hadronique, ces diagrammes sont négligeables devant la production de paires par interaction forte (voir § 1.3.2.5).

Le dernier vertex est beaucoup plus intéressant. Il couple le quark top au boson  $W$  et au quark  $b$ . C'est le seul mode de désintégration du quark top possible dans le Modèle

Standard. Avec sa masse de  $172.7 \text{ GeV}/c^2$ , le quark top se désintègre par interaction faible :  $t \rightarrow W^+b$ , avec la caractéristique notable que le boson  $W$  est réel car  $m_t > m_w + m_b$  est sur sa couche de masse.

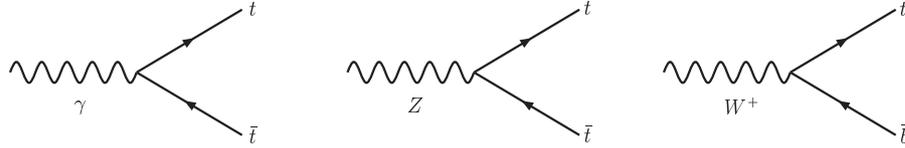


FIG. 1.8: Couplages du quark top aux bosons électrofaibles.

Ce vertex intervient également dans la production électrofaible du quark top, dont la recherche à  $D\bar{O}$  sera détaillée au chapitre 4. Contrairement à la production par couplage fort qui se fait par paire  $t\bar{t}$ , il n'y a qu'un seul quark top dans l'état final. Pour cette raison on parle généralement de production de *single top*. Il existe trois diagrammes à l'ordre de Born pour la production de *single top* :

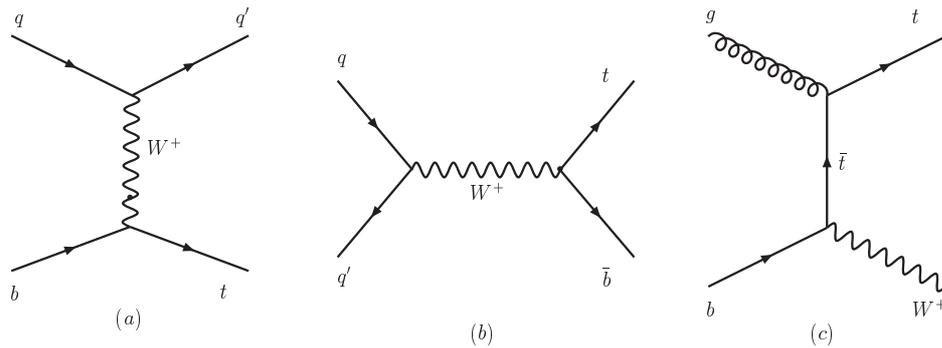


FIG. 1.9: Diagrammes de production de *single top* à l'ordre de Born dans le Modèle Standard. (a) voie  $t$ , (b) voie  $s$ , (c) production associée  $Wt$ .

- La voie  $t$  (figure 1.9a) correspond à l'échange d'un boson  $W$  virtuel du genre temps ( $Q_W^2 < 0$ ). Ce diagramme nécessite un quark  $b$  en entrée. Dans les collisions  $p\bar{p}$  la  $PDF$  du quark  $b$  est très inférieure à celle gluon et ce quark  $b$  provient principalement de la conversion d'un gluon en paire de quarks  $b$  ( $g \rightarrow b\bar{b}$ ). Le diagramme dominant pour un collisionneur hadronique est donc celui de la figure 1.10 (à gauche). C'est pourquoi ce mode de production est parfois appelé fusion de  $W$  et de gluon. Pour calculer ce diagramme il faut aussi prendre en compte le diagramme où le gluon se convertit en une paire  $t\bar{t}$ , qui donne une interférence négative, comme l'illustre la figure 1.10. La détermination théorique de la section efficace est encore compliquée par des problèmes de double comptage entre les processus  $2 \rightarrow 2$  et

$2 \rightarrow 3$  (figures 1.9a et 1.10) [22]. Les incertitudes théoriques sur ces sections efficaces proviennent principalement de la masse du quark top et des *PDF* du gluon et du quark  $b$ .

- La voie  $s$  (figure 1.9b) correspond à l'échange d'un boson  $W$  virtuel du genre espace ( $Q_W^2 > 0$ ). Ce processus est analogue au Drell-Yann.
- La production associée  $Wt$  (figure 1.9c) correspond à la production d'un boson  $W$  réel ( $Q_W^2 = m_W^2$ ) accompagné d'un quark top. Ce processus nécessite un quark  $b$  et un gluon dans la voie d'entrée et, pour les mêmes arguments que pour la voie  $t$ , deux gluons. Ce processus n'est donc pas favorisé au Tevatron où les interactions quark-quark et quark-gluon dominent.

Les sections efficaces théoriques de ces trois processus au Tevatron, calculées au deuxième ordre (***NLO***, *N*ext-to-*L*eading *O*rdre) pour une énergie dans le centre de masse  $\sqrt{s} = 1.96$  TeV sont données dans le tableau 1.4.

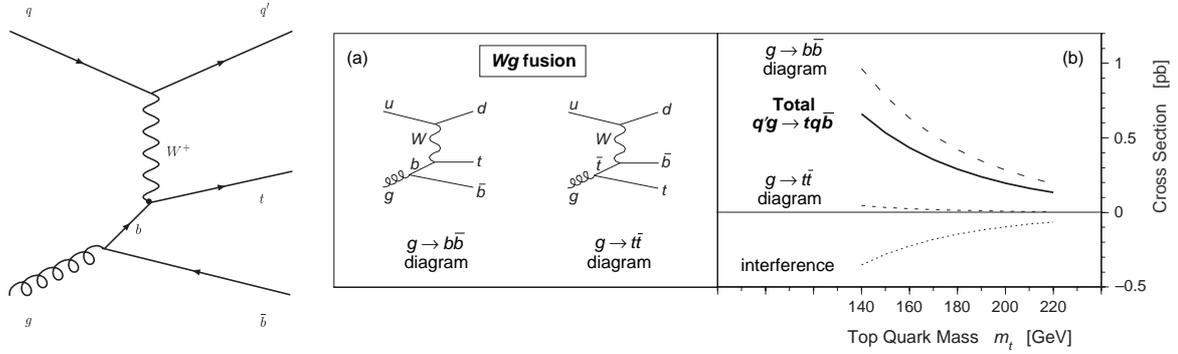


FIG. 1.10: A gauche : Diagramme dominant pour la production de *single top* dans la voie  $t$  par fusion de  $W$  et de gluon.

A droite :

(a) Diagrammes avec les conversions de gluon  $g \rightarrow b\bar{b}$  et  $g \rightarrow t\bar{t}$ .

(b) Section efficace de production de *single top* à  $\sqrt{s} = 1.8$  TeV pour chacun de ces diagrammes et l'interférence négative entre les deux processus en fonction de la masse du quark top [23].

Mode de production	Sections efficaces (pb)
voie $s$	$0.88 \pm 0.14$
voie $t$	$1.98 \pm 0.30$
production associée $Wt$	$0.093 \pm 0.024$

TAB. 1.4: Sections efficaces théoriques *NLO* de production *single top* à  $\sqrt{s} = 1.96$  TeV pour  $m_t = 175$  GeV [22] calculées avec les *PDF* CTEQ5M. Pour les voies  $s$  et  $t$ , l'erreur systématique inclut l'incertitude sur la masse du quark top [24].

### 1.3.3.3 Mesure de la section efficace de production de paires de top

La mesure de la section efficace de production de paires de quarks top est l'analyse la plus simple à réaliser. La valeur théorique de cette section efficace a été donnée au §1.3.2.5. Les états finals contiennent nécessairement deux jets de quark  $b$  et peuvent être caractérisés par les modes de décroissance des bosons  $W$ . L'identification des jets de quark  $b$  sera discutée au chapitre 3. Par abus de langage, on parle de lepton pour désigner un électron ou un muon dans l'état final. Les modes de désintégration leptonique du  $\tau$  ( $\tau \rightarrow e\nu_e\nu_\tau$  et  $\tau \rightarrow \mu\nu_\mu\nu_\tau$ ) sont implicitement inclus dans les états finals *dileptons* et *lepton + jets*. Jusqu'à présent, les modes avec des désintégrations hadroniques des leptons  $\tau$  ne sont pas pris en considération dans les analyses de  $D\bar{O}$ , du fait de la difficulté d'identification des  $\tau$  dans leurs modes hadroniques.

Les principaux canaux d'analyse sont :

- les modes *dileptons* : chaque  $W$  décroît en un muon ou un électron accompagné d'un neutrino. Ce canal est très propre car il nécessite l'identification de deux leptons isolés associés à deux jets et de l'énergie manquante, mais souffre de faibles rapports d'embranchement ( $ee, \mu\mu : 1.2\%$ ,  $e\mu : 2.4\%$ ). Les principaux bruits de fond sont la production associée de bosons  $Z$  et  $W$  et de jets, en particulier à des jets de saveurs lourdes ( $b$  et  $c$ ).
- les modes *lepton + jets* : Un  $W$  décroît en deux quarks formant des jets tandis que l'autre  $W$  se désintègre leptoniquement. La signature consiste donc en un lepton isolé, de l'énergie manquante (neutrino) et quatre jets, dont deux jets de quark  $b$ . Les bruits de fond principaux proviennent de la production associée  $W$ +jets et d'événements multijet mal reconstruits. Ce canal possède un fort taux d'embranchement (30%) et différentes techniques peuvent être mises en œuvre pour séparer le signal du bruit de fond : soit des méthodes d'analyse multivariable (fonction de vraisemblance) soit l'identification de jets  $b$  dans l'événement.
- le mode *hadronique* : Les deux bosons  $W$  se désintègrent en hadrons. L'état final possède six jets mais aucun lepton, et souffre ainsi d'un très important bruit de fond multijet malgré un grand rapport d'embranchement (44%). La mesure à  $D\bar{O}$  de la section efficace dans ce canal nécessite la combinaison d'une analyse multivariable (réseau de neurones) et de l'identification des jets de quark  $b$ .

Dans tous ces canaux, la mesure de la section efficace est réalisée par comptage d'événements, après l'application d'une sélection topologique. Le nombre d'événements de bruit de fond estimé (principalement à partir de simulation Monte Carlo) est confronté au nombre d'événements observés et à l'acceptance estimée pour le signal. La figure 1.11 résume l'ensemble des mesures de section efficace  $t\bar{t}$  à  $D\bar{O}$ . Les événements sélectionnés pour la mesure de la section efficace forment des lots de quarks top relativement purs et dont les quantités de bruit de fond résiduel sont connues. Ces lots sont aussi utilisés pour mesurer d'autres propriétés du quark top.

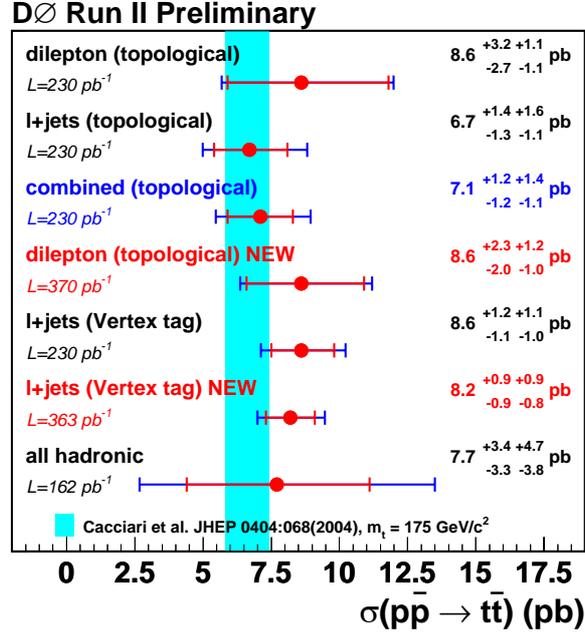


FIG. 1.11: Mesures de la section efficace de production de paires de quarks top à DØ dans différents canaux. Résultats préliminaires de l'été 2005 [25].

## 1.3.4 La masse et le mécanisme de Higgs

### 1.3.4.1 Brisure spontanée de la symétrie électrofaible

Les bosons de jauge sont, par construction, de masse nulle. De plus, nous avons vu au § 1.3.3.1 que la conservation de  $SU(2)_L \times U(1)_Y$  n'est possible que si les fermions sont également de masse nulle. Or l'expérience montre que tous les fermions ainsi que les bosons  $W^\pm$  et  $Z$  sont massifs. En 1964, un mécanisme fut proposé séparément par Higgs [26] et par Brout et Englert [27]. Ce mécanisme consiste en l'introduction dans la théorie d'un nouveau champ dont le lagrangien satisfait la symétrie  $SU(2)_L \times U(1)_Y$  mais introduit des termes qui, individuellement, brisent la symétrie. Ceci est réalisé pour un potentiel qui possède plusieurs minima équivalents. Le choix d'un minimum du potentiel pour la valeur dans le vide du champ brise la symétrie alors que le lagrangien total la conserve. De tels mécanismes de brisure spontanée de symétrie se rencontrent également en physique du solide pour décrire le ferromagnétisme ou encore l'effet Debye. Ainsi, on introduit arbitrairement dans le Modèle Standard un champ complexe scalaire sous forme d'un doublet d'isospin décrit par le lagrangien  $\mathcal{L}_{Higgs}$  :

$$\Phi(\mathbf{x}) = \begin{pmatrix} \varphi^{(+)} \\ \varphi^{(0)} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_3 + i\varphi_4 \end{pmatrix} \text{ et } \mathcal{L}_{Higgs} = (D_\mu \Phi^\dagger)(D^\mu \Phi) - V(\Phi) \quad (1.27)$$

Avec un potentiel de Higgs de la forme :

$$V(\Phi) = \mu^2 \Phi^+ \Phi + \lambda (\Phi^+ \Phi)^2 \quad (1.28)$$

Pour  $\mu^2 < 0$  et  $\lambda > 0$ , le potentiel est minimum pour  $\Phi^+ \Phi = \frac{-\mu^2}{2\lambda} = v^2/2$ . Le champ de Higgs prend une valeur non nulle dans le vide. La forme d'un tel potentiel est donnée par la figure 1.12. En effectuant un changement de repère, on peut toujours choisir comme

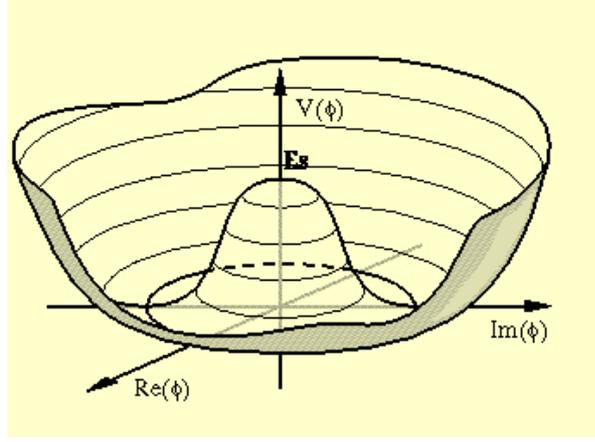


FIG. 1.12: Forme du potentiel de Higgs pour un champ scalaire complexe,  $V(\Phi) = \mu^2 \Phi^+ \Phi + \lambda (\Phi^+ \Phi)^2$  avec  $\mu^2 < 0$  et  $\lambda > 0$ .

valeur dans le vide :  $\Phi(0) = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$  et écrire le champ  $\Phi(\mathbf{x})$  comme un développement en série de Taylor autour de ce minimum :

$$\Phi(\mathbf{x}) = \frac{e^{\theta_a(\mathbf{x})T^a}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(\mathbf{x}) \end{pmatrix}, \text{ où } T^{a=1,2,3} \text{ sont les 3 générateurs de } SU(2). \quad (1.29)$$

Les trois champs  $\theta_a(\mathbf{x})$  sont des bosons de Goldstone. L'invariance du lagrangien sous  $SU(2)_L$  permet de faire disparaître le terme  $e^{\theta_a(\mathbf{x})T^a}$ . Il ne reste qu'un seul champ scalaire, le boson de Higgs,  $h(\mathbf{x})$  et la valeur constante du champ de Higgs dans le vide  $v$ . Les termes couplant le Higgs aux bosons de jauge dans le lagrangien 1.27 sont de la forme  $W_\mu^+ W^{-\mu} (v+h(\mathbf{x}))^2$  et  $Z_\mu Z^\mu (v+h(\mathbf{x}))^2$ . Le couplage proportionnel à  $v$  (valeur dans le vide du champ de Higgs) produit des termes analogues à un terme de masse ( $m_W^2 W_\mu^+ W^{-\mu}$  et  $m_Z^2 Z_\mu Z^\mu$ ). On vérifie également que la composante inférieure du doublet de Higgs ne se couple pas au photon : le boson de Higgs est électriquement neutre et le photon n'acquiert pas de masse. Les masses des bosons  $W$ ,  $Z$  et du boson de Higgs sont données par :

$$M_W = \frac{gv}{2}, \quad m_Z = \frac{m_W}{\cos \theta_W} \text{ et } m_h = \mu\sqrt{2} \quad (1.30)$$

Tout terme du lagrangien qui n'est pas interdit par le modèle, doit a priori être pris en compte. Il faut ainsi ajouter l'interaction entre le boson de Higgs et les fermions. Ce sont les couplages de Yukawa. Ils couplent les composantes gauche et droite des fermions et font ainsi apparaître des termes de masse pour les fermions. Le neutrino conserve une masse nulle car il n'existe pas de neutrino droit dans le Modèle Standard.

$$\mathcal{L}_{Yukawa} = -\lambda_t \bar{t}_R \Phi^\dagger t_L - \lambda_b \bar{b}_R \Phi^\dagger b_L + \text{conjugué hermitien} \quad (1.31)$$

La masse des fermions est reliée au couplage de Yukawa. Par exemple :

$$m_t = \lambda_t \frac{v}{\sqrt{2}} \quad (1.32)$$

Dans le Modèle Standard, les couplages de Yukawa  $\lambda_i$  sont arbitraires. Les masses des fermions sont donc des paramètres libres du modèle, au même titre que les constantes de couplage et la valeur dans le vide du champ de Higgs. Cette dernière est connue expérimentalement par la mesure de la constante de Fermi :

$$G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2} \text{c}^{-4}, \quad v = \sqrt{\frac{1}{\sqrt{2}G_F}} = 246 \text{ GeV}/\text{c}^2, \quad (1.33)$$

et fixe l'échelle d'énergie de la brisure spontanée de la symétrie électrofaible.

#### 1.3.4.2 La masse du quark top et la masse du boson de Higgs

Le boson de Higgs est à ce jour la seule particule du Modèle Standard qui n'ait pas été observée. Les expériences du LEP ont permis de mettre une limite directe sur sa masse,  $m_h > 114.3 \text{ GeV}/\text{c}^2$  [5]. Il est néanmoins possible d'obtenir des contraintes indirectes sur sa masse à partir des mesures de précision des masses du quark top et du boson  $W$ . La formule donnant la masse du boson  $W$  (équation 1.30) n'est valable qu'à l'ordre de Born). Comme pour les couplages, il faut prendre en compte les corrections radiatives aux ordres supérieurs. Du fait de la masse très élevée du quark top par rapport aux autres fermions, ce sont les boucles du top qui vont dominer.

La masse du  $W$  peut se réécrire en fonction de la constante de couplage électromagnétique  $\alpha$  (constante de structure fine) et de la constante de Fermi  $G_F$  sous la forme :

$$m_W = \frac{\pi\alpha}{\sqrt{2}G_F} \frac{1}{\sin^2 \theta_W} \quad (1.34)$$

$$\text{avec } \alpha = \frac{1}{4\pi} \frac{g^2 g'^2}{g^2 + g'^2} \text{ et } G_F = \frac{1}{\sqrt{2}v^2} \text{ à l'ordre de Born.} \quad (1.35)$$

En prenant en compte les corrections radiatives au premier ordre, la masse devient :

$$m_W = \frac{\pi\alpha}{\sqrt{2}G_F} \frac{1}{\sin^2 \theta_W (1 + \Delta r)} \quad (1.36)$$

Les contributions dominantes au terme correctif  $\Delta r$  proviennent des boucles de quarks top (figure 1.13) et, dans une moindre mesure du boson de Higgs (figure 1.14). Le calcul donne :

$$\Delta r_{top} \approx -\frac{3G_F m_t^2}{8\sqrt{2}\pi^2 \tan^2 \theta_W} \quad (1.37)$$

$$\Delta r_{Higgs} \approx \frac{11G_F m_Z^2 \cos^2 \theta_W}{24\sqrt{2}\pi^2} \ln \frac{m_h^2}{m_Z^2} \quad (1.38)$$

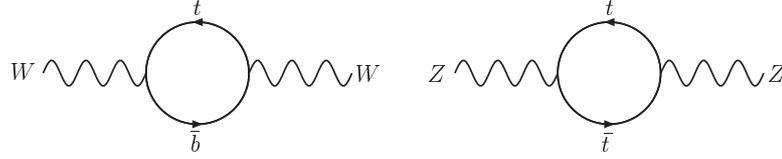


FIG. 1.13: Contribution du quark top aux corrections radiatives à une boucle aux masses du  $W$  et du  $Z$ .

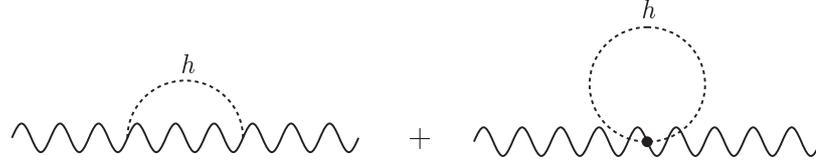


FIG. 1.14: Contribution du boson de Higgs aux corrections radiatives à une boucle aux masses du  $W$  et du  $Z$ .

La masse du boson  $W$  dépend maintenant de 5 paramètres. Les paramètres  $G_F$ ,  $\alpha$  et  $M_Z$  ont été mesurés avec une grande précision. En inversant le problème, connaître les masses du boson  $W$  et du quark top permet de contraindre la masse du boson de Higgs. La masse du  $W$  a été mesurée au LEP avec une précision de 0.05%. La masse du top provient des mesures combinées du Tevatron (DØ et CDF). Ces mesures sont présentées sur la figure 1.15. La figure 1.16 donne les contraintes actuelles sur la masse du boson de Higgs, avec :

$$m_W = 80.425 \pm 0.038 \text{ GeV}/c^2 \text{ et } m_t = 172.7 \pm 2.9 \text{ GeV}/c^2. \quad (1.39)$$

Les mesures tendent à favoriser un boson de Higgs léger. L'un des objectifs du Tevatron est de ramener l'incertitude sur la masse du  $W$  à 25 MeV/ $c^2$  et celle du top à 1.5 GeV/ $c^2$ .

### 1.3.4.3 Couplage électrofaible du quark top

Le couplage de Yukawa du quark top ( $\sqrt{2}\frac{m_t}{v}$ ) est proche de l'unité et la désintégration du top peut être sensible à l'existence d'une nouvelle physique à l'échelle d'énergie de

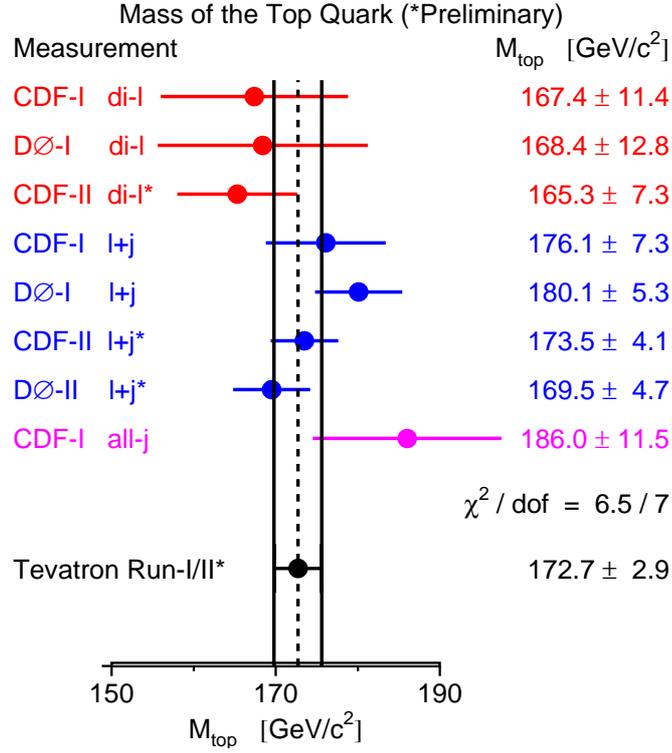


FIG. 1.15: Mesures de la masse du quark top à DØ et à CDF au Run I et II, dans différents canaux d'analyse. La moyenne mondiale actuelle est  $m_t = 172.7 \pm 2.9 \text{ GeV}/c^2$  [6].

la brisure spontanée de la symétrie électrofaible. Dans le Modèle Standard, le courant chargé a une structure purement  $V - A$ , c'est-à-dire que le boson  $W$  ne se couple qu'aux fermions de chiralité gauche. Il est possible de sonder la nature de ce couplage en étudiant la distribution angulaire entre le quark top et les produits de désintégration du boson  $W$  dans le référentiel au repos du  $W$ . Pour le fermion d'isospin  $-\frac{1}{2}$  (leptons chargés, quarks  $d$  et  $s$ ) cette distribution angulaire s'écrit en fonction de l'angle de désintégration  $\theta$  (voir figure 1.17) :

$$\omega(\theta) = \frac{4}{3} (m_t^2 \omega_0(\theta) + 2M_W^2 \omega_-(\theta) + m_b^2 \omega_+(\theta)), \quad (1.40)$$

avec :

$$\omega_0(\theta) = \frac{3}{4} \sin^2 \theta \quad (1.41)$$

$$\omega_-(\theta) = \frac{3}{8} (1 - \cos \theta)^2 \quad (1.42)$$

$$\omega_+(\theta) = \frac{3}{8} (1 + \cos \theta)^2 \quad (1.43)$$

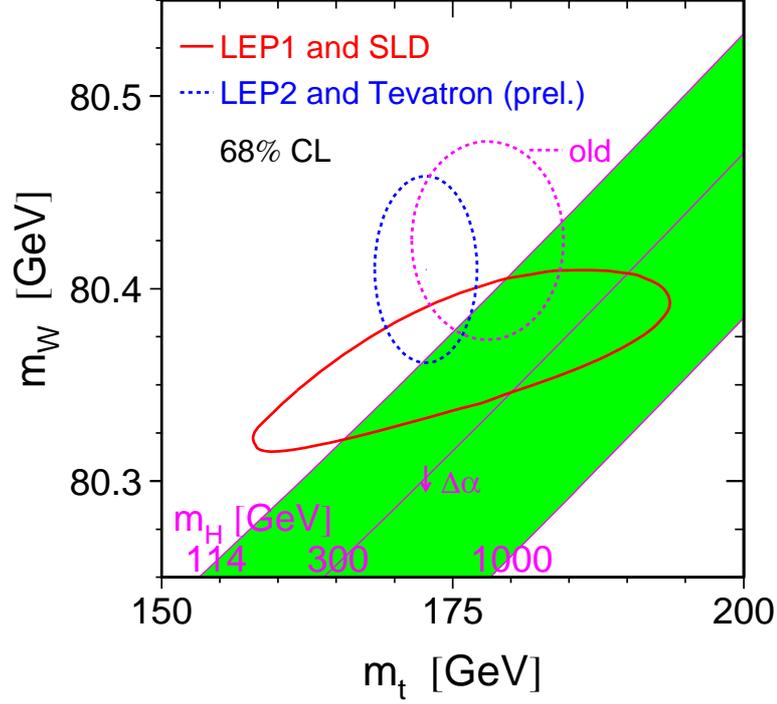


FIG. 1.16: Contraintes des masses du boson  $W$  et du quark top sur la masse du boson Higgs. Les contours en pointillé délimitent la région correspondant à un niveau de confiance de 68% sur les mesures directes de  $m_W$  (LEP2) et  $m_t$  (Tevatron). Le contour violet utilise la masse du top publiée au Run I ( $m_t = 178.1 \pm 4.3 \text{ GeV}/c^2$ ), le contour bleu inclut les résultats préliminaires du Tevatron de l'été 2005. Le contour rouge correspond aux mesures électrofaibles indirectes combinant les résultats de LEP1 et SLD. Les droites indiquent les isomasses pour le boson de Higgs, la bande verte correspondant à la région de masse comprise entre 114  $\text{GeV}/c^2$  (limite inférieure expérimentale) et 1000  $\text{GeV}/c^2$  (limite supérieure théorique) [29].

Cette distribution peut se réécrire en introduisant les fractions  $f_i$  de bosons  $W$  d'hélicité longitudinale ( $i = 0$ ), gauche ( $i = -1$ ) et droite ( $i = +1$ ).

$$\omega'(\theta) = \frac{\omega(\theta)}{\frac{4}{3}(m_t^2 + 2M_W^2 + m_b^2)} = f_0\omega_0(\theta) + f_{-}\omega_{-}(\theta) + f_{+}\omega_{+}(\theta) \quad (1.44)$$

Ces fractions correspondent aux rapports d'embranchement de  $t \rightarrow Wb$  pour chaque projection d'hélicité. La fraction  $f_0$  se calcule aisément par :

$$f_0 = \frac{m_t^2}{m_t^2 + 2M_W^2 + m_b^2} = (69.7 \pm 0.7)\%, \quad (1.45)$$

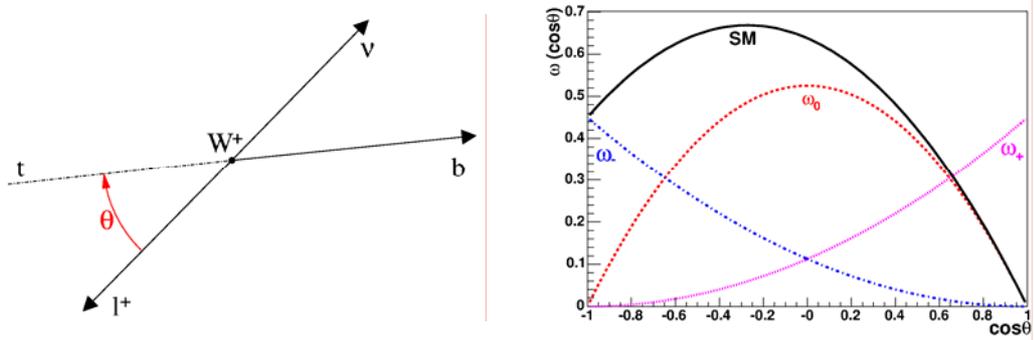


FIG. 1.17: A gauche : définition de l'angle  $\theta$  dans le référentiel du boson  $W$ .

A droite : distributions angulaires des produits de désintégration du boson  $W$  pour chaque polarisation. La ligne noire indique la distribution totale attendue dans le Modèle Standard.

pour  $m_t = 172.7 \pm 2.9 \text{ GeV}/c^2$ ,  $M_W = 80.425 \pm 0.038 \text{ GeV}/c^2$  et  $m_b = 4.5 \pm 0.4 \text{ GeV}/c^2$  [5]. Le quark  $b$  dans la désintégration du top est fortement *boosté* ( $m_b \ll m_W < m_t$ ) si bien que son hélicité et sa chiralité sont pratiquement équivalentes. Ainsi, la fraction  $f_+$  est fortement supprimée d'un facteur  $\frac{m_b^2}{m_t^2}$  car le boson  $W$  ne se couple qu'aux champs de chiralité gauche. Les contributions des différentes polarisations du boson  $W$  sont présentées sur la figure 1.17. Si le quark top possédait en plus un couplage de type  $V + A$ , cela augmenterait  $f_+$  en laissant  $f_0$  invariante. Une mesure de  $f_+$  permet donc de sonder la contribution d'un couplage  $V + A$  dans la désintégration du quark top. Une mesure de  $f_0$  déviant par rapport au Modèle Standard pourrait également mettre en évidence un couplage scalaire ou tensoriel. Expérimentalement, ces mesures sont réalisées en reconstruisant soit directement l'angle  $\cos \theta$  soit une variable corrélée (masse invariante du système lepton-quark  $b$ , impulsion transverse du lepton) et en utilisant des événements  $t\bar{t}$  dans les modes dilepton et lepton+jets. La principale difficulté consiste à reconstruire complètement l'événement en associant correctement le lepton chargé et le jet de quark  $b$  d'un même top. On utilise pour cela un ajustement contraint par les masses du  $W$  et du top. Au RunII, CDF a mesuré  $f_0 = 0.73^{+0.22}_{-0.34}$  et  $f_+ < 0.27$  à 95% de niveau de confiance [30]. DØ n'a mesuré que  $f_+ = 0.04 \pm 0.13$  soit  $f_+ < 0.25$  à 95% de niveau de confiance [31]. Aucune déviation significative par rapport au Modèle Standard n'a encore été observée.

#### 1.3.4.4 Matrice de mélange

L'interaction faible n'a été décrite jusqu'ici que pour une seule famille de quarks. Le passage aux trois familles connues du Modèle Standard introduit de nouveaux paramètres libres dans le lagrangien. Les observations montrent qu'un quark  $c$ , par exemple, peut se désintégrer en  $c \rightarrow Ws$  et  $c \rightarrow Wd$ . Ces deux désintégrations conservent l'isos-

pin faible mais pas la saveur des quarks. Ceci signifie que les états propres de masse ou de propagation des quarks (intervenant dans le lagrangien des fermions libres) diffèrent des états propres d'interaction faible. Les états propres d'interaction faible pour une projection d'isospin donnée ( $T_3 = \pm\frac{1}{2}$ ) sont donc une combinaison linéaire des états propres de masse. Du fait de l'invariance de jauge du lagrangien, il est toujours possible d'effectuer une rotation telle que les états propres de masse et d'interaction des quarks *up* ( $u, c, t; T_3 = +\frac{1}{2}$ ) coïncident. Une matrice de mélange doit être introduite pour relier états propres de masse ( $d, s, b$ ) et d'interaction ( $d', s', b'$ ) des quarks *down* ( $T_3 = -\frac{1}{2}$ ). Cette matrice complexe unitaire est appelée matrice de Cabbibo-Kobayashi-Maskawa (*CKM*) [32].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (1.46)$$

Le vertex électrofaible  $Wq\bar{q}'$ , couplant le boson  $W$  aux quarks est proportionnel au carré de l'amplitude de l'élément de la matrice *CKM* correspondant. La norme des éléments diagonaux est proche de l'unité ce qui favorise les couplages au sein d'une même famille. Les termes non diagonaux non nuls sont responsables de la majorité des désintégrations des hadrons étranges, charmés et beaux. Tous les termes de la matrice *CKM* ont été mesurés directement par l'étude de diverses désintégrations électrofaibles, à l'exception de  $|V_{tb}|$ , qui n'est connu que par les contraintes d'unitarité de la matrice. La valeur de  $|V_{tb}|$  est très voisine de l'unité ( $0.9990 < |V_{tb}| < 0.9992$  [5]). Par conséquent, le quark top se couple exclusivement au quark  $b$  ce qui justifie les modes de production et de désintégration du quark top déduits d'un modèle à une seule famille au § 1.3.3.2 et 1.3.3.3. La production électrofaible du quark top fait directement intervenir le vertex  $Wtb$  et son étude permet une mesure directe de  $|V_{tb}|$ . Dans le secteur leptonique, une matrice analogue est nécessaire si on considère des neutrinos massifs. Ce n'est pas dans la formulation habituelle du Modèle Standard où le nombre leptonique est conservé. Depuis 1998 on dispose cependant d'évidences expérimentales d'oscillation des neutrinos. Le Modèle Standard peut être élargi pour des neutrinos de faible masse et une nouvelle matrice de mélange est introduite [33] (matrice de Pontecorvo-Maki-Nakagawa-Sakata). La très faible masse des neutrinos pourrait indiquer que cette masse serait issue d'un mécanisme différent du mécanisme de Higgs [34].

## 1.4 Conclusion : Le Modèle Standard et au-delà

A ce jour, le Modèle Standard décrit dans ce chapitre n'a pas été mis en défaut, si ce n'est par l'observation des oscillations de neutrinos. Le secteur électrofaible a été étudié avec une grande précision au pôle du  $Z$  à LEP1 et dans la production de paires  $W^+W^-$  sans révéler de déviation significative avec le Modèle Standard (voir figure 1.18). En chromodynamique quantique les résultats de DESY et du Tevatron (la mesure de

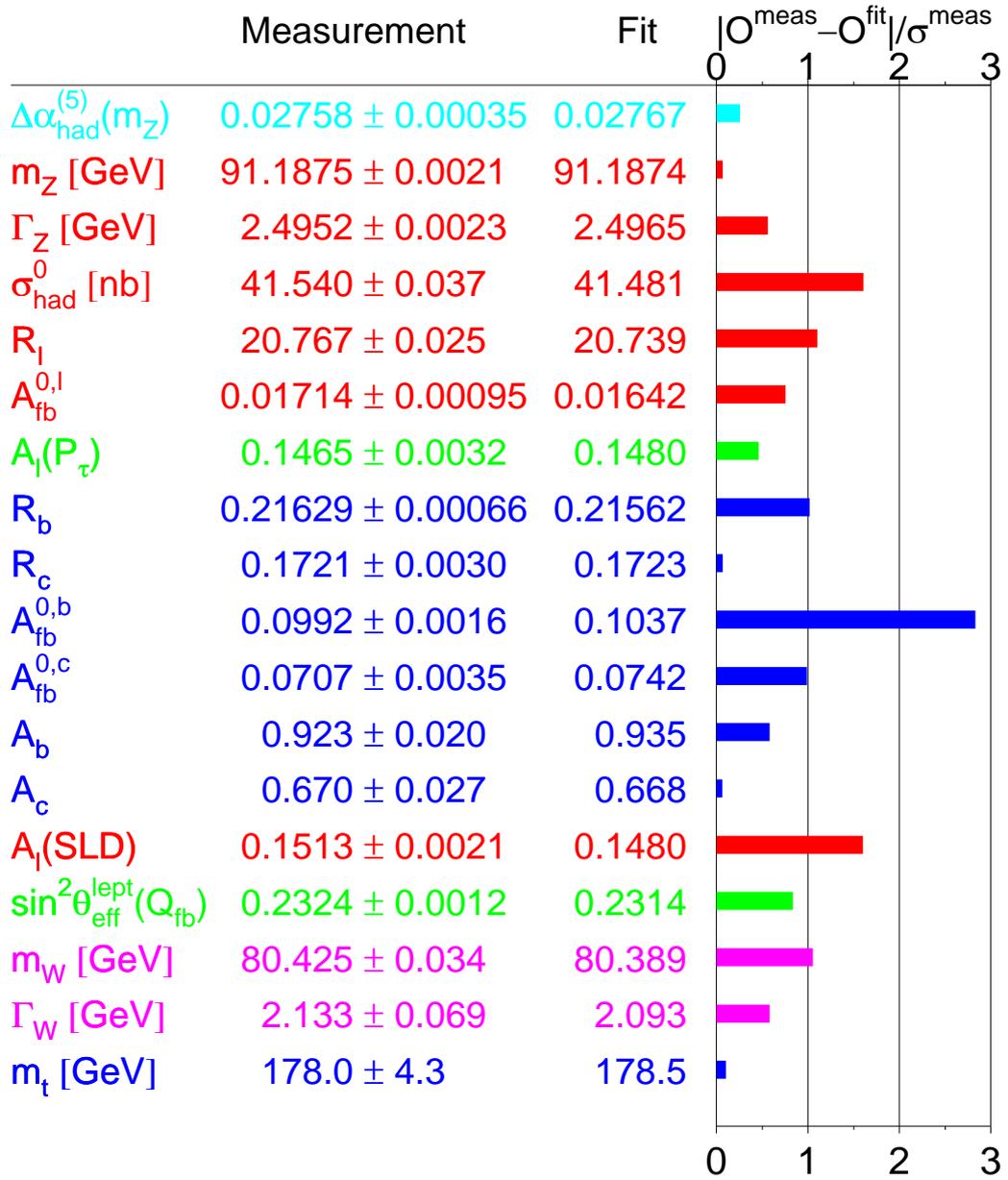


FIG. 1.18: Déviations relatives (*pull*) par rapport aux calculs théoriques des principales observables électrofaibles étudiées au pôle du  $Z$  à SLD et à LEP1 (constantes de couplages, masse et largeur du  $Z$ , asymétries avant/arrière, rapports d'embranchement), à LEP2 (masse et largeur du  $W$ ) et au Tevatron (masse du top) [29].

la section efficace de production de jets, les corrélations angulaires entre jets,...) sont également en accord avec la théorie dans le domaine perturbatif. Toutes les particules, à l'exception du boson de Higgs ont été observées. Cependant ce modèle n'est pas entièrement satisfaisant :

- Le Modèle Standard possède un total de 18 paramètres libres (9 masses de fermions, 3 constantes de couplages, 4 paramètres de la matrice  $CKM$ , 2 paramètres du potentiel de Higgs), voire 24 si on considère une extension avec neutrinos massifs. Il n'existe aucun argument théorique permettant de prédire leur valeur.
- Les corrections radiatives à la masse du Higgs divergent quadratiquement et peuvent dépasser de plusieurs ordres de grandeurs la masse au premier ordre. Seul un ajustement fin des paramètres peut faire disparaître ces divergences mais ce n'est pas une procédure naturelle. Ce problème, dit "de hiérarchie", est général pour toute particule scalaire fondamentale.
- Il n'existe pas d'unification des interactions électrofaible et forte à très haute énergie, les constantes de couplages ne convergent pas comme l'illustre la figure 1.3.
- L'interaction gravitationnelle, c'est-à-dire la relativité générale n'est pas incluse dans le Modèle Standard, qui repose sur la relativité restreinte (invariance de Lorentz). De plus la gravitation ne peut être formalisée dans le cadre d'une théorie de jauge car elle nécessite un boson de spin 2, le graviton.

Ces différents points semblent indiquer l'existence de symétries plus générales brisées à basse énergie. De nouveaux processus physiques sont donc attendus à haute énergie. Les modèles les plus séduisants sont probablement les théories supersymétriques qui introduisent une symétrie entre fermions et bosons. Dans le Modèle Supersymétrique Standard Minimal (MSSM), chaque fermion se voit ainsi associé un super-partenaire bosonique et réciproquement. Ce type de modèle introduit un grand nombre de paramètres libres supplémentaires mais présente cependant un certains avantages :

- Dans les corrections radiatives (à la masse du Higgs notamment) les boucles de bosons sont exactement compensées par les boucles des super-partenaires fermioniques, ce qui résout naturellement le problème de hiérarchie.
- Le potentiel de Higgs découle naturellement du potentiel supersymétrique. En revanche, deux doublets de Higgs sont nécessaires pour donner une masse aux particules d'isospin  $T^3 = -\frac{1}{2}$  et  $T^3 = +\frac{1}{2}$  (un seul est nécessaire dans le modèle standard, mais le nombre de doublets n'est pas limité). Après brisure de la symétrie électrofaible, il demeure cinq degrés de liberté, soit cinq bosons de Higgs :  $h$ ,  $H$ ,  $H^\pm$  et  $A$ .
- Les constantes de couplages mobiles semblent converger vers une valeur unique à haute énergie ( 1.19) ce qui accrédite la possibilité d'unifier les interactions électrofaible et forte au sein d'une symétrie plus large, brisée à basse énergie.
- La supersymétrie apparaît naturellement dans les théories plus générales de Supercordes.

A ce jour, aucune particule supersymétrique n'a été observée. Les contraintes existantes sur les paramètres des modèles supersymétriques indiquent que les masses des

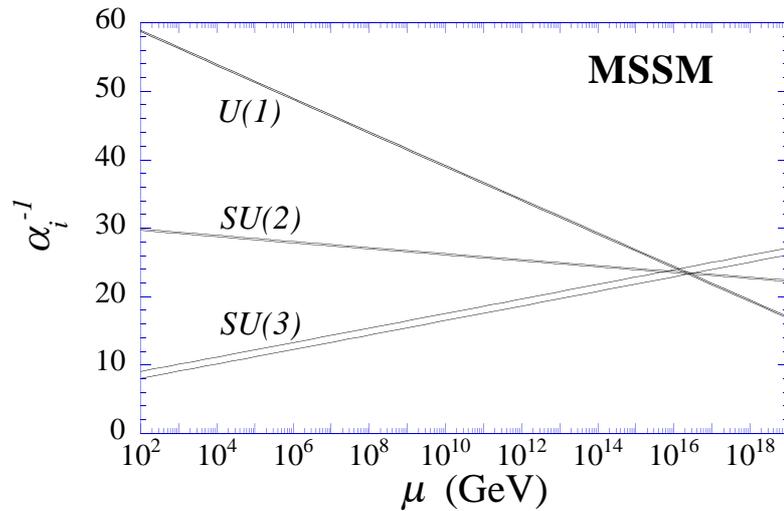


FIG. 1.19: Evolution de l'inverse des constantes de couplage électrofaibles ( $U(1)$  et  $SU(2)$ ) et forte ( $SU(3)$ ) dans le modèle supersymétrique minimal.

super-partenaires devraient se situer entre quelques centaines de GeV et quelques dizaines de TeV. D'autres modèles (modèles à dimensions supplémentaires, modèles technicouleur, modèles théoriques de grande unification incluant le MS dans un groupe de symétrie plus large :  $SO(10)$ ,  $E_6$ , ...) permettent aussi de résoudre certains problèmes du MS. Ces modèles prédisent également de nouvelles particules. D'une manière générale, ces différents modèles ne sont pas exclusifs. Si de nouvelles particules hors MS devaient être découvertes au Tevatron ou au LHC, l'interprétation en terme de modèles de nouvelle physique s'avérerait très complexe.

La physique du top est un secteur où cette nouvelle physique pourrait se manifester. De nouvelles particules pourraient modifier les sections efficaces de production et les modes de désintégration du top. Par exemple un boson de Higgs chargé ou un boson de jauge lié à une symétrie plus large  $W'$  peut se substituer au boson  $W$  dans les diagrammes de production électrofaible et de désintégration du quark top. Dans ce dernier cas si les rapports d'embranchement leptoniques et hadroniques du nouveau boson diffèrent de ceux du boson  $W$ , les sections efficaces mesurées pour la production  $p\bar{p} \rightarrow t\bar{t}$  dans les canaux dilepton, lepton+jets et hadronique seraient affectées. Un autre exemple serait celui d'une particule neutre, lourde se désintégrant en paire de quarks. Cette particule modifierait la section efficace de production  $t\bar{t}$  et pourrait être mise en évidence en reconstruisant le spectre en impulsion transverse des quarks top ou la masse invariante  $t\bar{t}$ . Le quark top du Modèle Standard devient également un bruit de fond majeur dans la recherche de ces nouvelles particules et sa compréhension en est d'autant plus importante.



*Two households, both alike in dignity,  
In fair Verona, where we lay our scene,  
From ancient grudge break to new mutiny,  
Where civil blood makes civil hands unclean.*

William Shakespeare, “*Romeo and Juliet*”, Prologue.

# 2

## Le Tevatron et l’expérience DØ

### 2.1 L’accélérateur Tevatron

#### 2.1.1 Introduction

Le Tevatron est le collisionneur installé au *Fermi National Accelerator Laboratory* (**FNAL**) [35] à Batavia, dans la banlieue de Chicago. Ce synchrotron est le dernier élément d’une chaîne d’accélération (voir figure 2.1) qui permet de réaliser des collisions entre protons et antiprotons avec une énergie dans le centre de masse  $\sqrt{s} = 1.96$  TeV (0.98 TeV par faisceau). Il existe deux zones de collision sur l’anneau qui abritent les détecteurs **DØ** [36] et **CDF** (*Collider Detector at Fermilab*) [37]. Ces deux expériences permettent d’étudier les processus physiques produits lors des collisions en détectant et en identifiant les particules (électrons, muons, jets de quarks et de gluons) produites lors d’interactions dures. La physique des collisions hadroniques a été introduite au § 1.3.2.4. Tevatron a connu deux périodes d’opération. Le Run I, entre 1992 et 1996, a accumulé une luminosité intégrée d’environ  $120 \text{ pb}^{-1}$  par expérience. L’analyse de ces données a permis, entre autre, la découverte du quark top en 1995 par DØ et CDF [8]. Une nouvelle phase de prise de données, le Run IIa, a débuté en 2001 après une interruption de plusieurs années consacrées à améliorer les performances des détecteurs et de l’accélérateur. L’énergie de collision est passée de 1.8 à 1.96 TeV et la luminosité instantanée a augmenté d’un facteur 5. Environ  $1.5 \text{ fb}^{-1}$  de données ont été délivrés par le Tevatron d’août 2002 à février 2006. A partir de juin 2006, une troisième phase de prise de données le Run IIb devrait démarrer avec pour objectif de recueillir d’ici à 2009 un total de 4 à  $8 \text{ fb}^{-1}$  pour l’ensemble du Run II (IIa+IIb). La luminosité instantanée et intégrée délivrée par le Tevatron pour le Run IIa est donnée sur la figure 2.2.



Fermilab Tevatron Accelerator With Main Injector

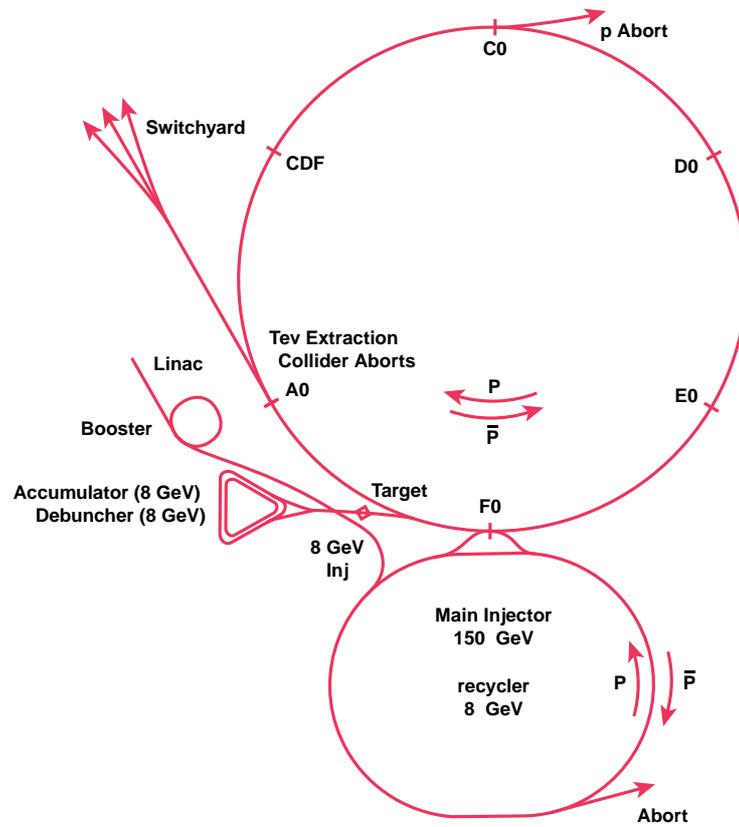


FIG. 2.1: En haut : Vue aérienne du complexe d'accélération et du collisionneur Tevatron.  
 En bas : Représentation schématique de la chaîne d'accélération.

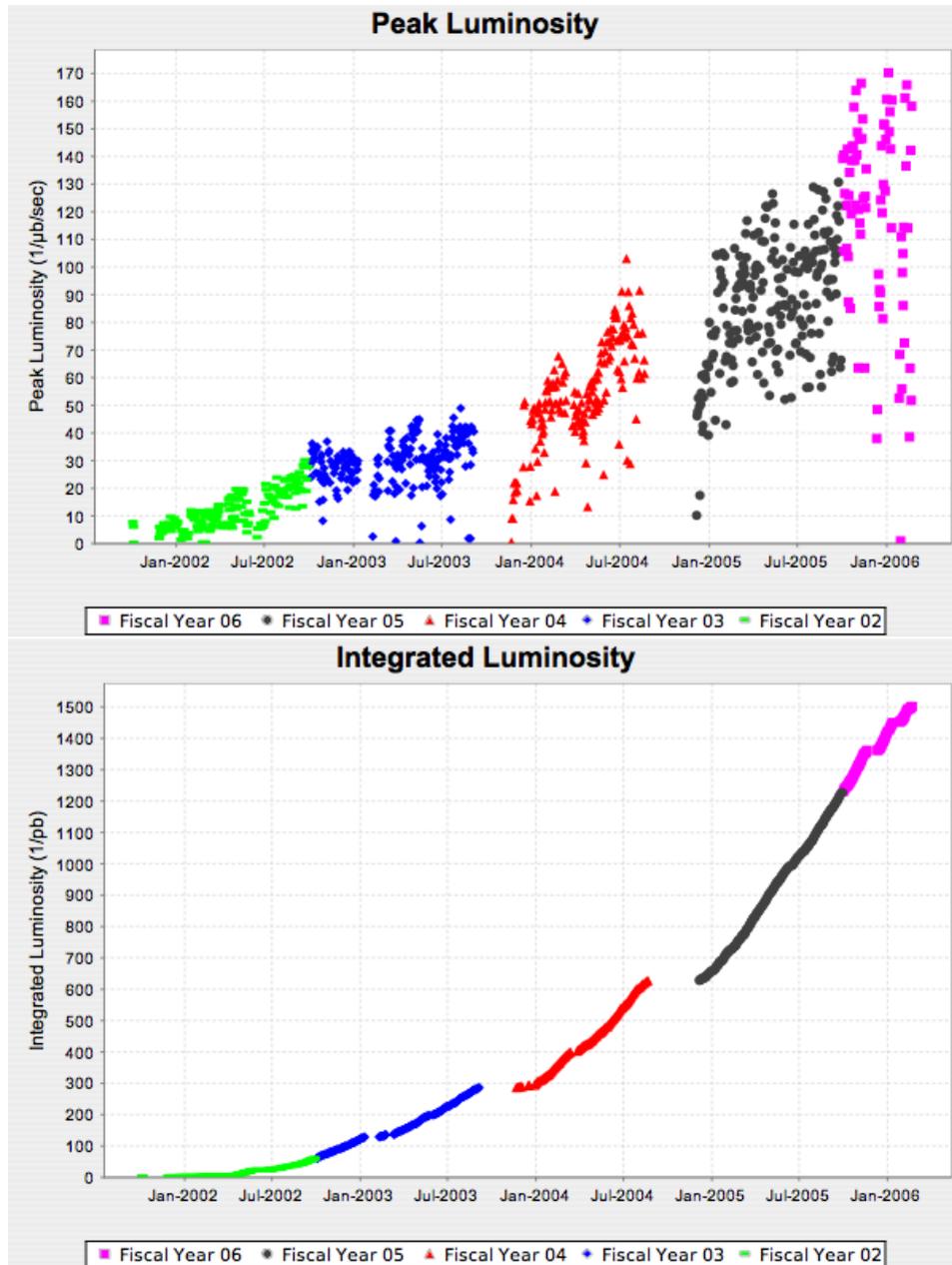


FIG. 2.2: A gauche : luminosité instantanée maximale pour le Run IIa, d'octobre 2001 à février 2006.

A droite : luminosité intégrée délivrée par le Tevatron pour le Run IIa.

### 2.1.2 Le faisceau de protons

Le faisceau de protons est produit à partir d'hydrogène gazeux puis accéléré par différents dispositifs. La chaîne d'accélération des protons consiste en [38] :

- *La production d'ions  $H^-$*  : de l'hydrogène gazeux est ionisé par un champ électrique de 25 keV pour former des ions  $H^-$ .
- *L'accélération des ions  $H^-$*  : les ions  $H^-$  sont accélérés successivement par deux dispositifs. Un accélérateur électrostatique **Cockcroft-Walton** confère aux ions une énergie de 750 keV. Les ions sont ensuite injectés dans un accélérateur linéaire de 130 m de long (**Linac**) dont les cavités radio-fréquences permettent au faisceau d'ions d'atteindre une énergie de 400 MeV.
- *La production du faisceau de protons* : Les ions  $H^-$  sont envoyés sur une cible de carbone, qui leur arrache leurs électrons. Il en résulte un faisceau de protons.
- *L'accélération des protons* : Les protons sont ensuite accélérés par trois synchrotrons et groupés en paquets :
  - Le **Booster** est un petit synchrotron de 120 m de diamètre qui permet d'accélérer le faisceau de protons à une énergie de 8 GeV.
  - L'**injecteur principal** est un synchrotron de 500 m de diamètre qui joue plusieurs rôles. Il augmente l'énergie du faisceau issu du **Booster** jusqu'à 150 GeV et regroupe les protons en paquets de  $2.10^{11}$  pour leur injection vers le **Tevatron**. Il permet également de produire un faisceau de 120 GeV de haute intensité (paquets  $2.10^{12}$  protons) utilisé pour les expériences sur cibles fixes, dont un cas particulier est la production des antiprotons. Enfin l'injecteur principal est utilisé pour l'accélération des antiprotons à 150 GeV avant leur injection dans le **Tevatron** et pour recueillir et décélérer les antiprotons restant après une série de collisions.
  - Le **Tevatron** est le dernier synchrotron de la chaîne d'accélération. Les faisceaux de protons et d'antiprotons y sont amenés à une énergie de 960 GeV et stockés pendant plusieurs heures (environ 30 heures au maximum) pour réaliser des collisions entre paquets de protons et d'antiprotons au centre des détecteurs CDF et DØ. Chaque faisceau se compose de 36 paquets, regroupés en trois super-paquets séparés par un temps de  $2 \mu s$ . **Tevatron** utilise 722 aimants supraconducteurs refroidis par hélium liquide produisant un champ magnétique de 4 T afin de courber la trajectoire des faisceaux. Les principales caractéristiques du **Tevatron** sont regroupées dans le tableau 2.1

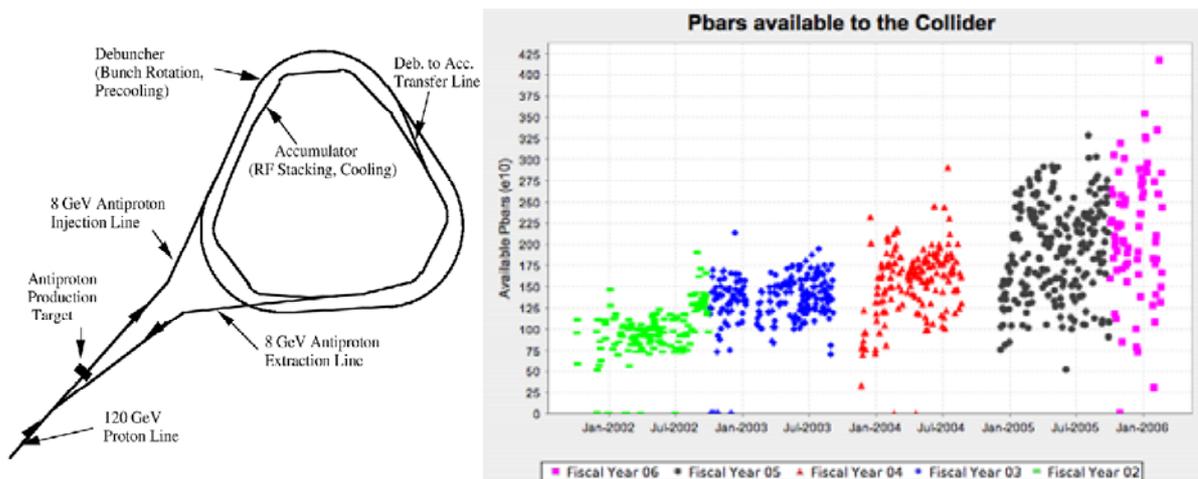
### 2.1.3 Production et recyclage du faisceau d'antiprotons

Les antiprotons sont produits par la collision du faisceau de protons de 120 GeV issu de l'injecteur principal sur une cible de nickel. Les antiprotons sont ensuite focalisés par une lentille magnétique. Il faut environ  $10^5$  protons pour produire un antiproton. Les antiprotons ainsi créés possèdent une énergie moyenne d'environ 8 GeV, avec une

	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 ( $\text{cm}^{-2}\text{s}^{-1}$ )	$1.6 \times 10^{30}$	$1 \times 10^{32}$	$3 \times 10^{32}$
Luminosité intégrée ( $\text{fb}^{-1}$ )	0.15	1.4	4-8
Nombre de paquets $p \times \bar{p}$	$6 \times 6$	$36 \times 36$	$36 \times 36$
Nombre d'interactions	2.5	2.3	8
Temps de croisement (ns)	3500	396	396

TAB. 2.1: Principales caractéristiques du Tevatron aux Run I, IIa et IIb.

grande dispersion. Le **debuncher**, un accélérateur triangulaire, permet de réduire cette dispersion en énergie. Il utilise des cavités radiofréquences qui ralentissent les antiprotons d'énergie supérieure à 8 GeV et accélère ceux d'énergie inférieure en effectuant une rotation d'espace de phase. Le faisceau perd sa structure en paquet au profit d'une meilleure résolution en énergie. L'énergie du faisceau est également stabilisée par refroidissement stochastique : la dispersion est mesurée à chaque tour et corrigée à l'aide d'un champ électromagnétique. Après 1.5 s le faisceau continu ainsi produit est envoyé dans l'**accumulateur**, situé dans le même tunnel que le **debuncher**. Les antiprotons y sont stockés et accumulés pendant plusieurs heures jusqu'à atteindre une intensité suffisante. Le refroidissement stochastique se poursuit dans l'accumulateur. Finalement, les antiprotons sont regroupés en paquets puis envoyés vers l'injecteur principal puis dans le **Tevatron**. La chaîne de production des antiprotons ainsi que l'évolution du nombre d'antiprotons produits au cours du Run IIa sont représentées sur la figure 2.1.3.

FIG. 2.3: A gauche : la chaîne de production des antiprotons comportant la cible, le **debuncher** et l'**accumulateur**.

A droite : nombre d'antiprotons disponibles au cours du temps.

La faible efficacité de production limite donc la quantité d'antiprotons disponibles. Le nombre d'antiprotons est ainsi un facteur limitant de la luminosité. Pour augmenter le nombre d'antiprotons disponible, un **recycleur** a été mis en place. Seule une faible partie des particules circulant dans l'anneau participe effectivement aux collisions. Le recycleur devait permettre de récupérer et de stocker les antiprotons non utilisés pour les réinjecter ultérieurement dans le Tevatron. Ceci c'est révélé impossible à mettre en œuvre pratiquement et la recycleur est utilisé comme simple anneau de stockage supplémentaire. Le recycleur consiste en un synchrotron à aimants permanents conçu pour stocker un faisceau de 8 GeV sur une longue période. Il est situé dans le même tunnel que l'injecteur principal. Depuis l'été 2005, un faisceau d'électrons de 4.8 MeV est utilisé pour améliorer le refroidissement et ainsi augmenter l'intensité du faisceau d'antiprotons dans le recycleur.

## 2.2 L'expérience DØ

### 2.2.1 Description générale

L'expérience DØ est l'un des deux détecteur installés auprès du Tevatron. Ce détecteur possède une géométrie cylindrique autour de l'axe des faisceaux. Il est conçu pour reconstruire les objets physiques à grande impulsion transverse produits lors des collisions  $p\bar{p}$ .

Le système de coordonnées utilisé est un repère cartésien orthonormé  $(O, x, y, z)$ . L'origine  $O$  du repère est placé au centre du détecteur. L'axe du faisceau est l'axe  $(Oz)$ , orienté dans le sens de circulation du faisceau de protons. L'axe  $(Oy)$  est vertical, dirigé vers le haut. Pour repérer l'impulsion  $\mathbf{p}$  d'une particule dans le détecteur les coordonnées sphériques  $(p_T, \theta, \phi)$  sont également utilisées où  $p_T$  est l'impulsion transverse (projetée dans le plan  $xy$ ),  $\theta$  l'angle polaire relativement à l'axe  $z$  et  $\phi$  l'angle azimutal relativement à l'axe  $x$ . Plutôt que l'angle  $\theta$ , on préfère utiliser la pseudorapidité  $\eta$  définie par :

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

Pour des particules ultra-relativistes, la pseudorapidité est équivalente à la rapidité  $y$  qui est un invariant de Lorentz :

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

La géométrie du détecteur et le système de coordonnées sont illustrés sur la figure 2.4. Après définition du vertex primaire de l'interaction (voir § 2.2.2.4), on distingue deux définition de la pseudorapidité d'un objet : la pseudorapidité physique  $\eta$  qui correspond à la quadri-impulsion propagée au vertex primaire et la pseudorapidité géométrique  $\eta_{det}$  relative au centre du détecteur  $(0,0)$ . La pseudorapidité géométrique décrit la position

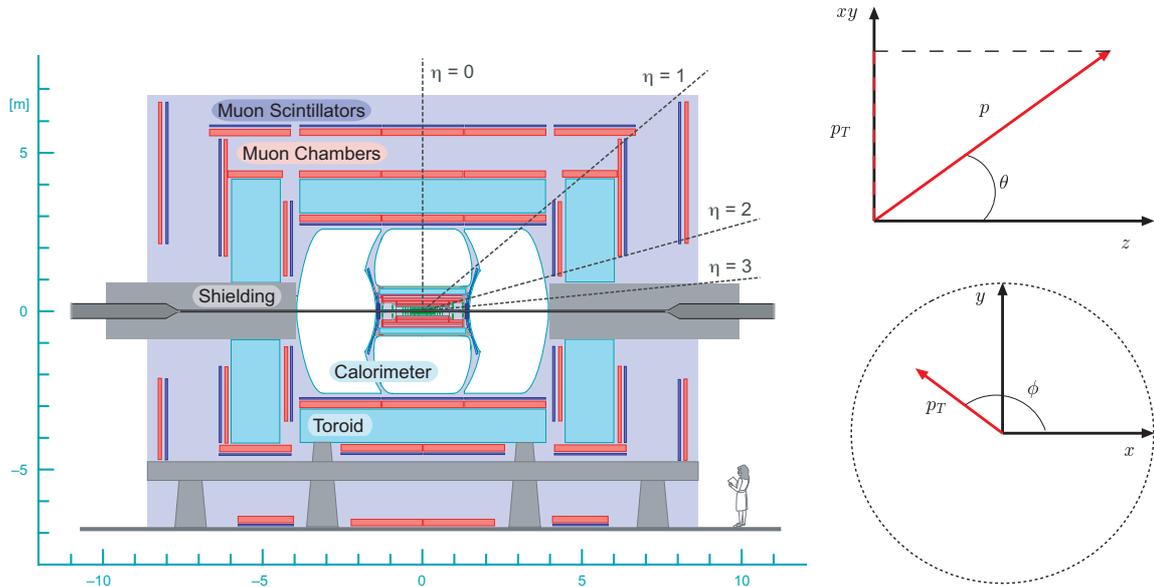


FIG. 2.4: Représentation schématique du détecteur DØ (coupe longitudinale) et du système de coordonnées. L'axe  $z$  est l'axe du faisceau et le plan transverse est le plan  $(x,y)$ . Les dimensions sont indiquées en mètres.

de l'objet dans le détecteur. Sauf indication contraire, la pseudorapacité physique sera systématiquement utilisée par la suite. Le détecteur DØ se compose de multiples sous-détecteurs concentriques que l'on peut regrouper en trois ensembles :

- au plus près du faisceau, un détecteur de traces pour reconstruire les trajectoires des particules chargées.
- un calorimètre qui permet l'identification des électrons, des photons et des jets.
- un spectromètre externe, destiné à la reconstruction des muons.

Le détecteur DØ utilisé au Run II reprend certains éléments du détecteur du Run I [39] en particulier le calorimètre et une partie du système à muons. plusieurs améliorations y ont été apportées [40], notamment :

- Le trajectographe et le solénoïde supraconducteur ont été ajouté.
- L'électronique du calorimètre a été entièrement remplacée pour s'adapter au nouveau temps de croisement (396 ns au lieu de 3500 ns).
- Des détecteurs de pieds de gerbe ont été ajoutés afin d'améliorer la résolution sur la reconstruction des électrons.
- La partie avant du spectromètre à muons a été remplacée pour offrir une meilleure résistance aux radiations.
- Des détecteurs de protons à grande rapidité ont été installés pour étudier les processus diffractifs.
- Le système de déclenchement et d'acquisition a été entièrement renouvelé.

Nous allons maintenant décrire plus en détail chaque sous-détecteur.

## 2.2.2 Le détecteur de traces

Le détecteur de traces constitue la partie la plus interne du détecteur. Il se compose de deux sous-détecteurs : un détecteur de vertex à micropistes de silicium (*Silicon Microstrip Tracker, SMT*) et un détecteur à fibres scintillantes (*Central Fiber Tracker, CFT*). Ces deux détecteurs sont installés à l'intérieur d'un solénoïde supraconducteur de 2 T qui permet de mesurer l'impulsion des traces reconstruites. Le schéma d'ensemble du trajectographe est donné sur la figure 2.5

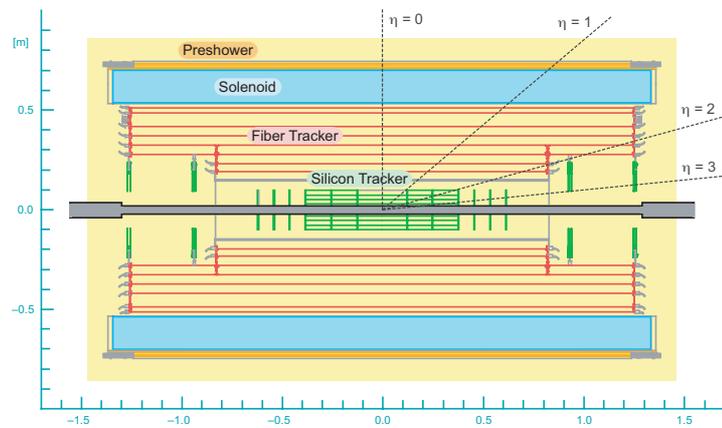


FIG. 2.5: Représentation en coupe du détecteur de traces complet et du solénoïde. Les dimensions sont indiquées en mètres.

### 2.2.2.1 Le détecteur de vertex (*SMT*)

La figure 2.6 donne une représentation complète du détecteur de vertex *SMT*.

Il se compose de capteurs à micropistes de silicium organisés en six barillets d'un rayon interne de 2.7 cm et d'un rayon externe de 10.5 cm placés autour du tube à vide. Chaque barillet est composé de 4 couches concentriques de détecteurs simple et double face. Sur chaque couche, les capteurs sont agencés en quinconce afin d'obtenir une bonne herméticité. Cette structure est illustrée sur la figure 2.7. Pour améliorer l'acceptance du détecteur, 6 disques verticaux (F) composés de détecteurs double face sont intercalés entre les barillets aux positions  $|z| = 12.5, 25.4$  et  $38.2$  cm. Six autres disques F sont installés aux extrémités de l'ensemble aux positions  $|z| = 43.1, 48.1$  et  $53.1$  cm. Chaque disque F a un rayon interne de 2.6 cm et un rayon externe de 10 cm. Ce dispositif permet la reconstruction des traces jusqu'à  $|\eta_{det}| = 2$ . Enfin 4 disques verticaux (H), plus larges (rayon interne de 9.6 cm, rayon externe de 26 cm), sont installés aux extrémités du détecteur aux positions  $|z| = 100.4$  et  $121.0$  cm et permettent d'étendre la couverture

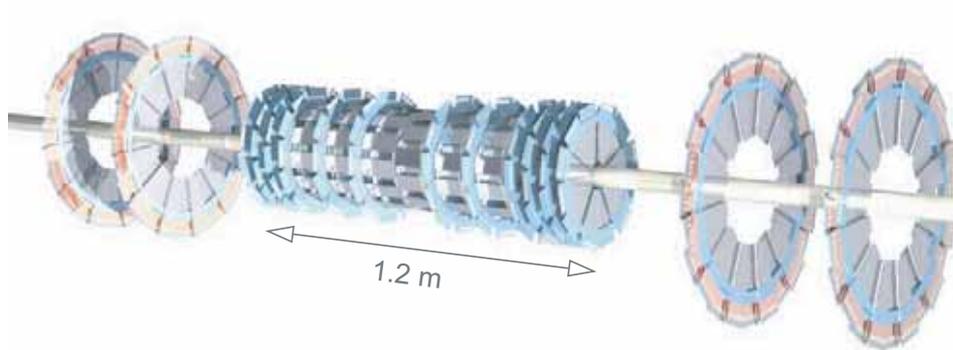


FIG. 2.6: Représentation en tridimensionnelles du *SMT* : barillets, disques F (internes) et disques H (externes).

en pseudorapidité jusqu'à  $|\eta_{det}| = 3$ . Le *SMT* compte un total de 793000 canaux de lecture, regroupés en quatre super-couches (*superlayer*). La résolution spatiale du *SMT* atteint  $10 \mu\text{m}$  en  $r - \phi$  et  $40 \mu\text{m}$  selon l'axe  $Oz$ .

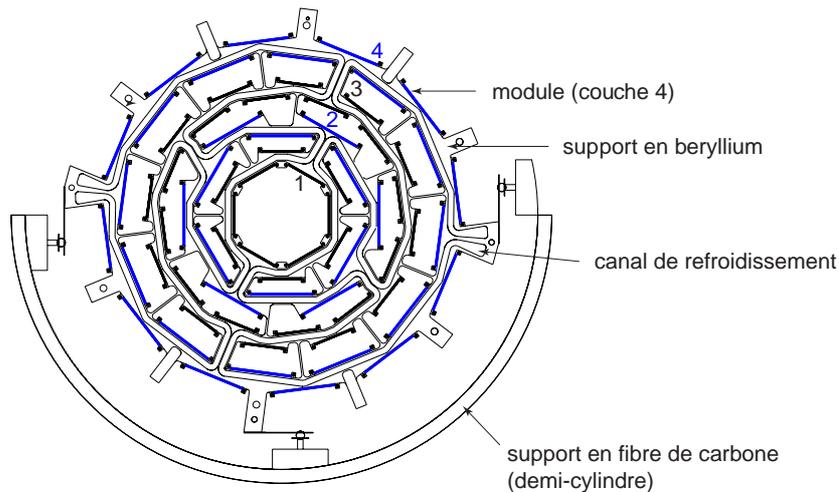


FIG. 2.7: Coupe transversale d'un tonneau du *SMT*. Les couches 1 et 3 sont indiquées en noir, les couches 2 et 4 en bleu.

Pour le Run IIb, une couche de silicium supplémentaire vient d'être insérée à l'intérieur de la première couche du *SMT*. Cette *couche 0* doit permettre de compenser la dégradation de la résolution due au vieillissement du détecteur sous l'effet des radiations et à la perte de canaux de lecture et d'améliorer notablement la résolution en paramètre d'impact des traces.

### 2.2.2.2 Le détecteur à fibres scintillantes (*CFT*)

Le *CFT* est un détecteur à fibres scintillantes qui complète le *SMT* pour la reconstruction des traces dans la partie centrale du détecteur ( $|\eta_{det}| < 1.7$ ). Il consiste en huit cylindres concentriques en fibre de carbone portant chacun deux couches de fibres scintillantes de  $835 \mu\text{m}$  de diamètre. La première couche est orientée selon l'axe  $Oz$ , la seconde faisant un angle de  $+3^\circ$  (cylindres impairs) ou  $-3^\circ$  (cylindres pairs) avec la première. Une vue en coupe d'une partie cylindre est présentée sur la figure 2.8. Les deux couches les plus internes ont une longueur de 1.66 m, les six autres mesurent 2.52 m. Une représentation schématique complète du *CFT* est donnée sur la figure 2.5. Des guides de lumière transparents prolongent chacune des 74000 fibres scintillantes. La lecture du signal est assurée par des photodiodes à avalanche (*Visible Light Photo Counters, VLPC*) fonctionnant à très basse température (10 K). La résolution spatiale de chaque couche du *CFT* est de l'ordre de  $100 \mu\text{m}$  dans le plan  $r - \phi$ .

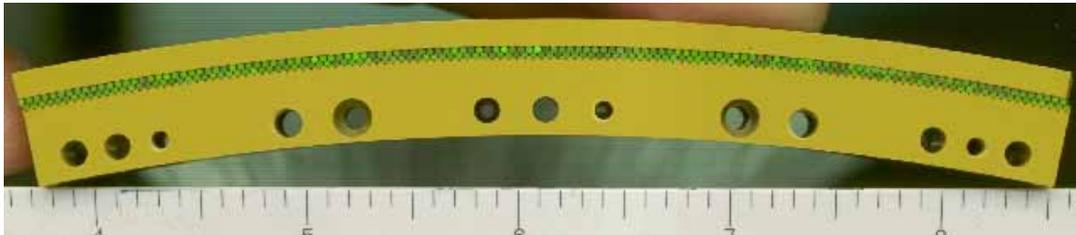


FIG. 2.8: Coupe transversale d'une partie de cylindre du *CFT*, avec la structure en fibre de carbone (en jaune) et les deux couches de fibres scintillantes (en vert).

### 2.2.2.3 La reconstruction des traces

Le *SMT* et le *CFT* sont insérés dans un solénoïde supraconducteur de 60 cm de rayon. La matière formant l'épaisseur de cet aimant correspond à une longueur de radiation d' $1 X_0$ <sup>1</sup>. Le champ magnétique à l'intérieur du détecteur de traces est uniforme, d'une valeur de 2 T. La carte du champ magnétique est donnée par la figure 2.9.

Trois domaines en pseudorapacité sont différenciés lors de la reconstruction des trajectoires des particules chargées pour tenir compte de l'acceptance des différents sous-détecteurs :

- $|\eta_{det}| < 1.7$  : Dans cette région, la couverture du *CFT* est totale. Les coups dans le *CFT* et le *SMT* sont utilisés pour reconstruire les traces, qui doivent toutes avoir des coups dans chaque couche du *CFT* (16 coups).
- $1.7 < |\eta_{det}| < 2$  : La couverture du *CFT* n'est pas totale et les traces ont moins de coups dans ce sous-détecteur. Dans cette région, le *SMT* et le *CFT* sont utilisés.

<sup>1</sup>Une longueur de radiation d' $1 X_0$  correspond à une perte d'énergie moyenne par bremsstrahlung pour un électron de  $1/e$ .

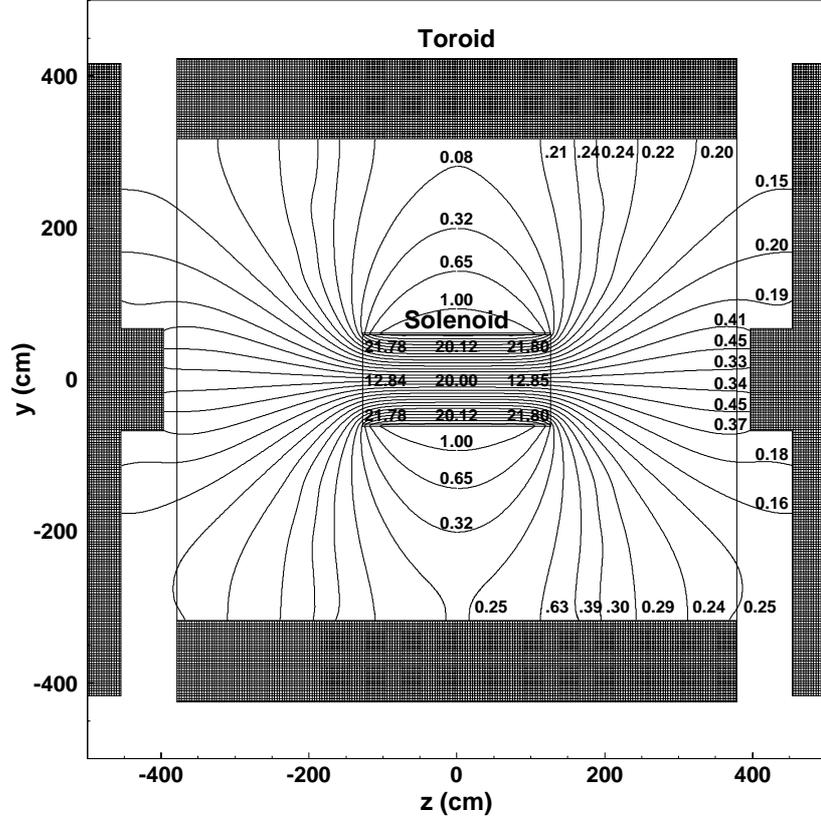


FIG. 2.9: Carte du champ magnétique du solénoïde de DØ . Les valeurs du champ sont indiquées en kG (1 kG = 0.1 T).

–  $|\eta_{det}| > 2$  : Seul le *SMT* couvre ce domaine de pseudorapidité. Les trajectoires des particules chargées sont reconstruites relativement au centre géométrique du détecteur. La trajectoire d'une particule chargée dans un champ magnétique uniforme est une hélice. Une description plus complète peut être trouvée dans la référence [41]. Les déformations de la trajectoire hélicoïdale idéale dues à la matière traversée par les particules et aux défauts d'alignement sont prises en compte. L'ajustement des paramètres et de la matrice d'erreur des traces est réalisé au moyen d'algorithmes de Kalman [42], qui constituent une version améliorée de l'ajustement par la méthode des moindres carrés. Le  $\chi^2$  de l'ajustement est utilisé comme critère de qualité de la reconstruction d'une trace. La limite inférieure sur l'impulsion transverse d'une trace reconstruite est  $p_T > 183$  MeV, Le trajectographe ne permet que de reconstruire la trajectoire et donc l'impulsion d'une particule chargée, pas son énergie. Par défaut, la masse du pion ( $m_{\pi^+} = 139.6$  MeV/ $c^2$  [5]) est imposée pour reconstruire la quadri-impulsion d'une trace. L'efficacité de reconstruction des traces est illustrée par la figure 2.10. Les résolutions en impulsion et en impulsion transverse des traces reconstruites sont données

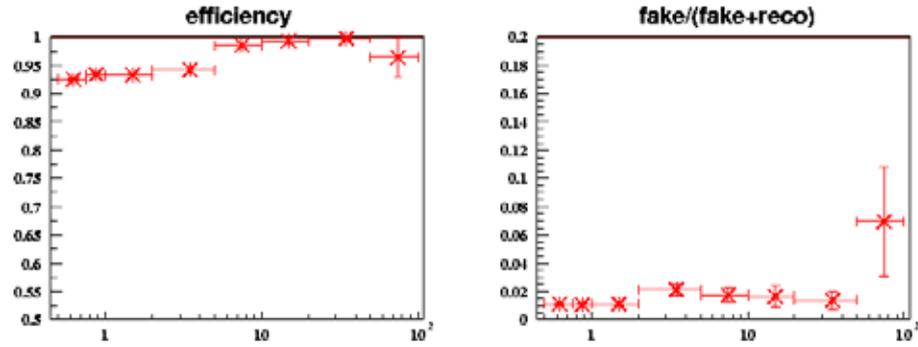


FIG. 2.10: Efficacité et taux de mauvaise reconstruction des traces dans DØ en fonction de l'impulsion transverse.

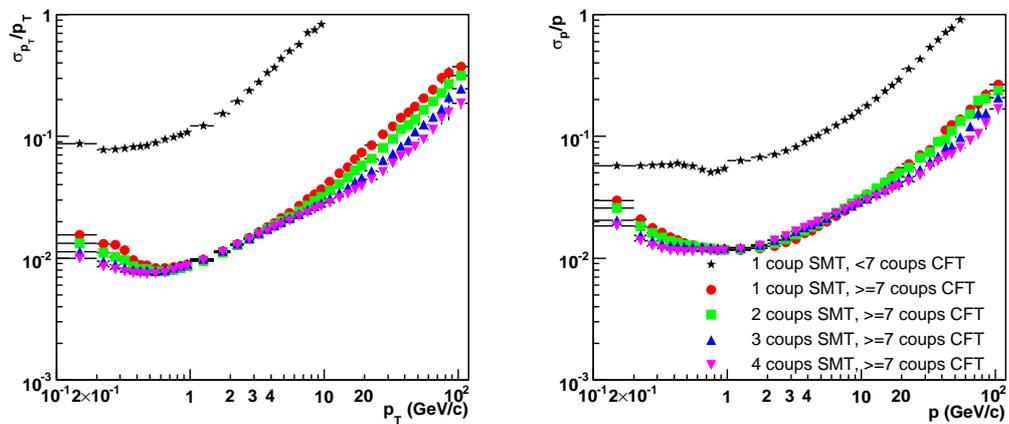


FIG. 2.11: Résolution en impulsion transverse (à gauche) et en impulsion (à droite) des traces reconstruites. Les étoiles noires correspondent aux traces vers l'avant où la couverture de la *CFT* est moins bonne (1 coups dans la *SMT*, moins de 7 coups dans la *CFT*). Les traces avec au moins 7 coups dans la *CFT* et de 1 à 4 coups dans une couche de la *SMT* sont représentées respectivement par des disques rouges, des triangles verts, des triangles bleus et des carrés violets.

sur la figure 2.11. Cette résolution est de l'ordre du pourcent pour des traces d'impulsion comprise entre une centaine de MeV/c et quelques GeV/c. La dégradation de la résolution pour les grandes impulsions est liée à l'augmentation de l'incertitude sur la courbure. À faible impulsion, cette dégradation provient des diffusions multiples dans le détecteur.

#### 2.2.2.4 La reconstruction du vertex primaire

Un ajustement de l'intersection des traces reconstruites permet de reconstruire le vertex primaire d'interaction de l'événement. La position de ce vertex primaire est alors

utilisée pour calculer correctement les quadri-impulsions des objets physiques reconstruits (traces, jets, électrons, muons,...) et doit donc être connue avec précision. La résolution sur la position du vertex primaire intervient également dans les algorithmes d'identification de jets de quark  $b$  qui recherchent des traces déplacées par rapport au vertex primaire (voir chapitre 3).

Seules les traces satisfaisant une qualité suffisante sont utilisées pour reconstruire les vertex. Elles doivent avoir au moins deux coups dans le *SMT* et leur impulsion transverse doit vérifier  $p_T > 0.5 \text{ GeV}/c$ . L'algorithme de recherche du vertex primaire fonctionne en deux étapes. Un vertex dans le plan transverse ( $xy$ ), correspondant à la position du faisceau, est reconstruit en utilisant toutes les traces dont la signification en paramètre d'impact transverse (cf. § 3.2.1) relativement au centre du détecteur  $(0, 0)$  est inférieure à 100. La quadri-impulsion de chaque trace ainsi que sa matrice d'erreur sont alors propagées à ce vertex.

Une recherche plus fine du vertex est alors effectuée en utilisant uniquement les traces dont la signification transverse relativement au vertex précédemment calculé est inférieure à 4. Cette dernière coupure permet de rejeter des traces issues de vertex secondaires déplacés, créés par des particules à grand temps de vol (hadrons beaux, charmés, ou étranges, conversions de photon, leptons  $\tau$  ...) ou des traces mal reconstruites.

Des amas de traces sont formés en associant itérativement les traces dont la position selon l'axe  $z$  se situe à moins de 2 cm de la position de l'amas. Les traces de plus grande impulsion transverse servent de point de départ et la position de l'amas est recalculée à chaque ajout de trace. Un vertex est alors ajusté en utilisant un filtre de Kalman [43]. Les quadri-impulsions et les matrices d'erreur des traces sont également réajustées lors de cette procédure.

A cause de l'empilement des événements et d'éventuelles interactions multiples, il peut y avoir plusieurs vertex reconstruits pour un même événement. Une sélection probabiliste basée sur la distribution en impulsion transverse des traces associées permet de déterminer le vertex primaire de l'interaction [44].

La figure 2.12 illustre certaines caractéristiques des vertex primaires reconstruits. La distribution longitudinale du vertex primaire est une gaussienne de largeur 27 cm, ce qui justifie la longueur du *SMT*. La résolution sur la position du vertex primaire dans le plan transverse est de  $38 \mu\text{m}$  dans la partie centrale du détecteur (position selon  $z$  du vertex inférieure à 40 cm).

Le même type d'algorithme peut être utilisé pour reconstruire des vertex secondaires déplacés par rapport au vertex primaire. Ces vertex secondaires servent notamment à identifier les jets de saveurs lourdes ou à reconstruire des modes de désintégrations spécifiques de hadrons. Un exercice intéressant consiste à identifier les vertex à deux traces en imposant la masse de l'électron pour chaque trace. Les vertex ainsi déterminés correspondant à une masse invariante voisine de 0 proviennent majoritairement de conversions de photon ( $\gamma \rightarrow e^+e^-$ ) dans la matière du détecteur. Ces conversions se produisant uniquement quand le photon traverse de la matière, la position de ces vertex est fortement corrélée à la géométrie du détecteur. Une telle tomographie du détecteur de vertex est

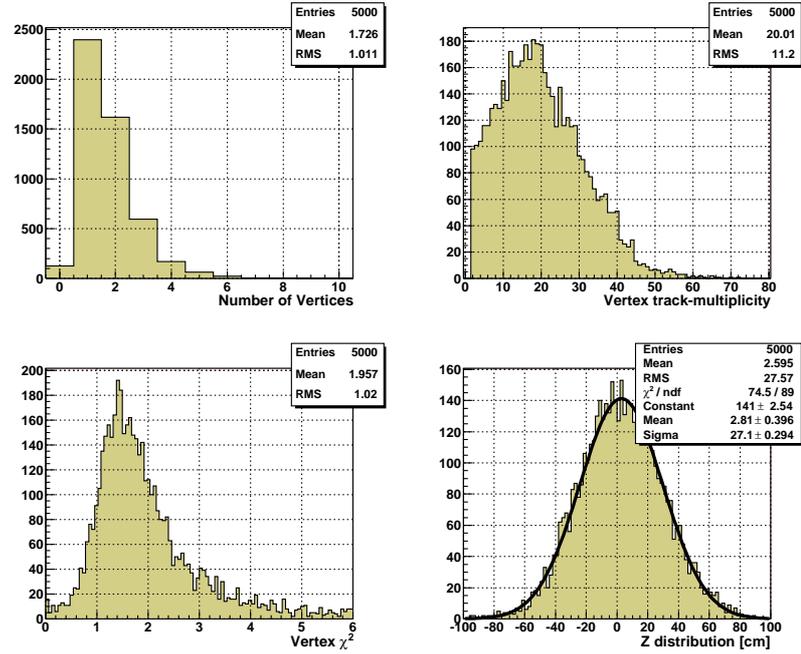


FIG. 2.12: Quelques caractéristiques des vertex reconstruits dans des événements multijet : de gauche à droite et de bas en haut, la multiplicité de vertex par événement, la multiplicité de traces par vertex, le  $\chi^2$  par degré de liberté de l'ajustement du vertex et la position selon  $z$  du vertex.

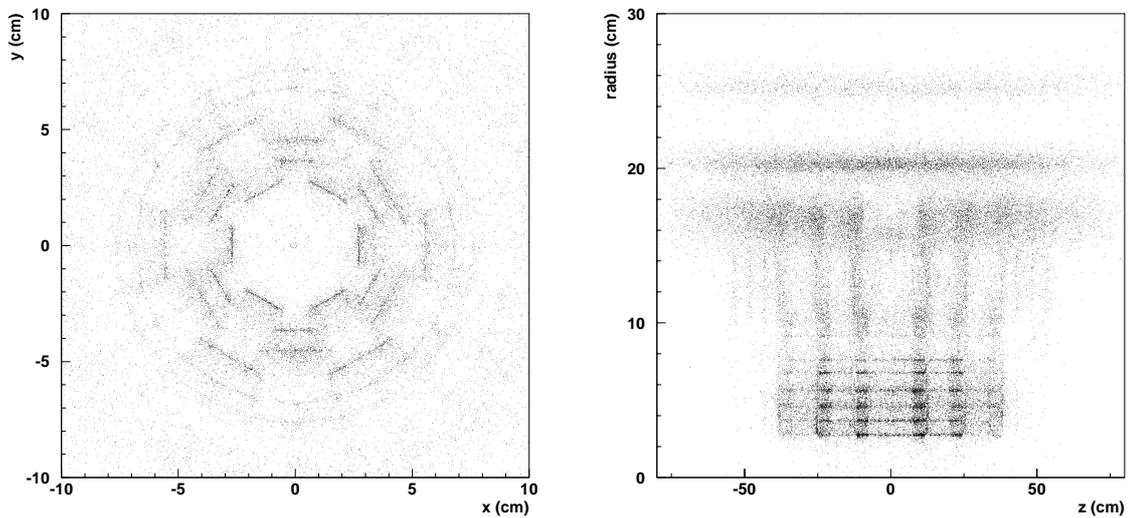


FIG. 2.13: Tomographie du SMT par reconstruction des vertex  $\gamma \rightarrow e^+e^-$ . La position des vertex reconstruits dans le plan transverse  $xy$  (à gauche) et dans le plan  $rz$  (à droite) permet de mettre en évidence les différentes couches du détecteur au silicium, les premiers disques F ainsi que les premiers cylindres du  $CFT$  (cf. figures 2.5 et 2.7).

présentée par la figure 2.13. Cet exercice illustre les capacités de reconstruction des traces et des vertex du trajectographe de DØ .

## 2.2.3 La calorimétrie

### 2.2.3.1 Les détecteurs de pieds de gerbe

Le solénoïde correspond à une longueur de radiation de  $1X_0$  et les gerbes électromagnétiques peuvent commencer à se développer dans cette région. Pour améliorer la reconstruction des électrons, des détecteurs de pieds de gerbe ont été installés entre le solénoïde et le calorimètre. Le détecteur central *CPS* (*C*entral *P*re*S*hower) est cylindrique et situé autour du solénoïde. Le détecteur avant *FPS* (*F*orward *P*re*S*hower) est installé sur les bouchons du calorimètre, dans la région intercryostat (cf. § 2.2.3.3). Ils sont composés de barreaux de scintillateurs organisés en 2 (*FPS*) ou 3 (*CPS*) couches. Le coeur de chaque barreau est extrudé pour accueillir une fibre plastique qui décale la longueur d'onde de la lumière émise et assure sa collecte. Les guides de lumière transparents qui prolongent ces fibres ainsi que les détecteurs *VLPC* et l'électronique de lecture sont identiques à ceux du *CFT*. La géométrie des barreaux et la structure d'une couche est décrite par la figure 2.14. Les détecteurs de pieds de gerbe font à la fois office

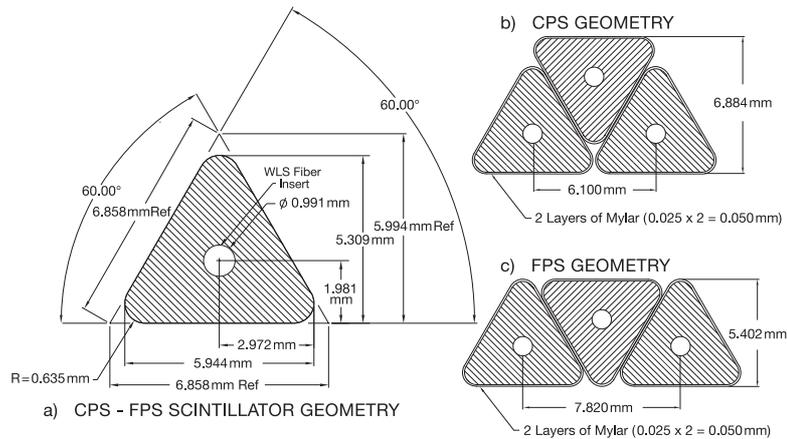


FIG. 2.14: A gauche : un barreau du détecteur de pieds de gerbe. A droite : structure d'une couche du *CPS* (en haut) et du *FPS* (en bas).

de trajectographe et de calorimètre. Malheureusement, faute de temps et de moyens, ils ne sont pas exploités dans la reconstruction et l'analyse des données utilisées dans cette thèse. L'information du *CPS* est incluse dans la reconstruction des électrons de basse impulsion ( $p_T \lesssim 15 \text{ GeV}/c$ ) dans les versions les plus récentes du programme de reconstruction.

### 2.2.3.2 Le calorimètre

Le calorimètre de DØ est composé d'un tonneau central ( $|\eta_{det}| \lesssim 1$ ) et de deux bouchons ( $1.3 \lesssim |\eta_{det}| \lesssim 4.2$ ). Une représentation du calorimètre est donnée par la figure 2.15.

Le calorimètre se compose de cellules disposées en couches de lecture. Les cellules

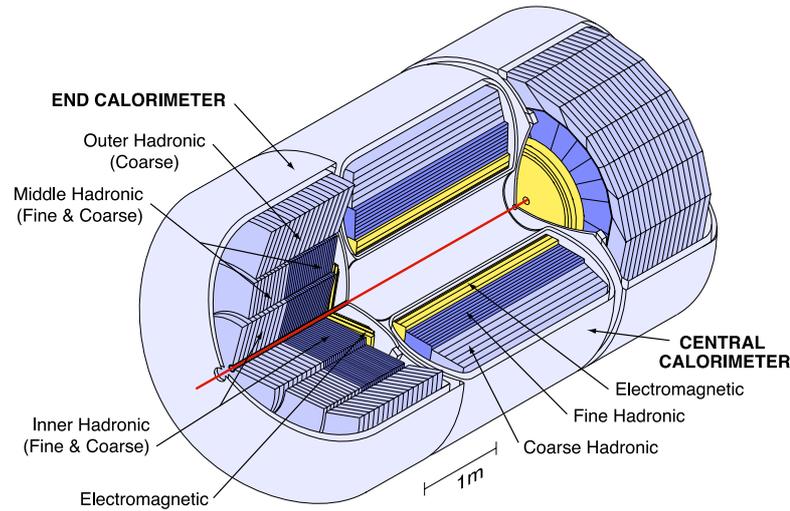


FIG. 2.15: Représentation schématique du calorimètre avec la partie centrale et les deux bouchons. Les couches électromagnétiques du calorimètre sont indiquées en jaune, les couches hadroniques *FH* et *CH* en bleu foncé et clair respectivement.

sont organisées en tours pseudoprojectives de segmentation  $\Delta\eta_{det} \times \Delta\phi = 0.1 \times 0.1$ . Cette structure est illustrée par la figure 2.16. Au total le calorimètre compte environ 5000 tours. C'est un calorimètre à échantillonnage qui utilise l'argon liquide comme milieu actif. Chaque partie du calorimètre est enfermée dans un cryostat qui maintient la température interne à environ 80 K. Il est compensé, c'est-à-dire que les réponses électromagnétiques et hadroniques sont similaires, ce qui assure que la réponse totale du calorimètre est linéaire en fonction de l'énergie des hadrons. Le calorimètre se compose de trois sections concentriques où différents matériaux sont utilisés comme milieu passif :

- Les quatre couches les plus internes utilisent de l'uranium appauvri pur et forment le **calorimètre électromagnétique** (EM). La troisième couche, où les gerbes électromagnétiques développent leur amplitude maximale, est plus finement segmentée ( $\Delta\eta_{det} \times \Delta\phi = 0.05 \times 0.05$ ) et permet ainsi une meilleure identification des particules électromagnétiques.
- Les trois couches suivantes (quatre dans les bouchons) forment la **région hadronique à granularité fine** (*Fine Hadronic, FH*) du calorimètre. Un alliage d'uranium-niobium (2%) est utilisé comme matériau passif. C'est dans cette région que se développent les gerbes hadroniques.

- La dernière couche de lecture forme le **calorimètre hadronique à faible granularité** (*Coarse Hadronic, CH*) qui utilise comme matériau passif du cuivre dans le tonneau central ou de l'acier dans les bouchons. Cette région possède une granularité moins fine que le *FH*.

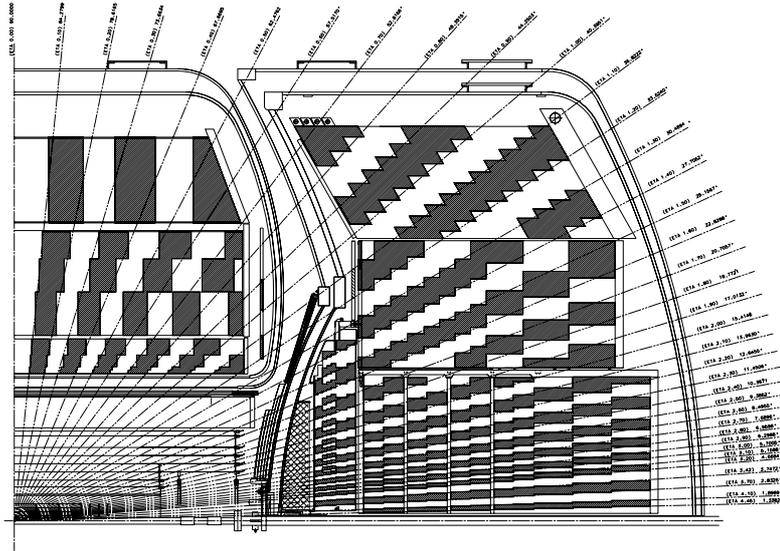


FIG. 2.16: Structure en couches et en tours pseudoprojectives d'un quart du calorimètre.

### 2.2.3.3 La région inter-cryostat

L'espace situé entre le calorimètre central et chaque bouchon ( $1 < \eta_{det} < 1.3$ ) est peu instrumenté. Cette région inter-cryostat ou *ICR* est principalement occupée par les câbles de lectures du *SMT* et du *CFT*. Des scintillateurs sont installés dans cette région du détecteur afin de compenser l'absence de cellules calorimétriques. Ce **détecteur inter-cryostat**, l'*ICD*, est complété par des cellules calorimétriques sans absorbeur situées le long de la paroi interne des cryostats, appelées *massless gaps*. Ces détecteurs offrent une faible résolution et ne permettent qu'une récupération partielle de l'information perdue dans cette région. En particulier les électrons ne peuvent être identifiés dans cette région.

### 2.2.3.4 La reconstruction des électrons et des photons

Les électrons sont reconstruits avec deux algorithmes différents, correspondant à des domaines d'impulsion transverse différents.

A basse impulsion transverse, les électrons sont reconstruits par la méthode de la route [47] qui utilise comme point de départ les traces isolées en leur associant des cellules du calorimètre.

A grand  $p_T$  ( $p_T > 15$  GeV/c) les électrons sont reconstruits avec un algorithme de cône [48]. Cet algorithme utilise uniquement les tours du calorimètre et reconstruit également les photons. Des amas calorimétriques sont formés à partir des tours du calorimètre situées dans un cône de rayon  $= \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.2$  autour d'une tour d'énergie transverse supérieure à 0.5 GeV/c. Un objet électromagnétique se différencie d'un jet par sa fraction électromagnétique et son isolation. La fraction électromagnétique,  $Emf$  correspond à la fraction d'énergie de l'amas déposée dans les couches du calorimètre électromagnétique. L'isolation  $iso$  est caractérisée par la quantité relative d'énergie située dans un cône d'ouverture  $0.4 < \Delta R < 0.2$  autour de l'amas :

$$iso = \frac{E^{\Delta R=0.4} - E_{EM}^{\Delta R=0.2}}{E_{EM}^{\Delta R=0.2}} \quad (2.3)$$

avec  $E^{\Delta R=0.4}$  ( $E_{EM}^{\Delta R=0.2}$ ) l'énergie totale (électromagnétique) déposée dans un cône de rayon 0.4 (0.2) autour de la direction de l'amas. Les coupures standard d'identification d'un objet électromagnétique sont  $Emf > 0.9$  et  $iso < 0.15$ .

La forme de la gerbe permet également de séparer un objet électromagnétique d'un pion. Un ensemble de 7 variables corrélées est utilisé pour caractériser cette forme : les fractions d'énergie dans chacune des 4 couches du calorimètre électromagnétique, l'énergie électromagnétique totale, la largeur angulaire de la gerbe selon  $\phi$  et la position du vertex primaire selon l'axe  $Oz$ . La matrice H est définie comme l'inverse de la matrice de covariance des 7 variables. Le  $\chi^2$  de cette matrice caractérise le critère électromagnétique de l'amas. Un objet électromagnétique satisfait  $\chi_{Hmx7}^2 < 50$ .

Les électrons sont différenciés des photons par l'association d'une trace à l'amas électromagnétique. Un amas sans trace associée est un photon.

Finalement, une dernière variable d'identification des électrons est définie par une fonction de vraisemblance [49] combinant des variables calorimétriques et trajectographiques :  $\chi_{Hmx7}^2$ ,  $Emf$ , le rapport de l'énergie transverse de l'amas à l'impulsion transverse de la trace  $E_T/p_T$ , le nombre de traces dans un cône de rayon  $= 0.05$  autour de la trace de l'électron, la somme des impulsions transverses des traces dans un cône de rayon  $\Delta R = 0.4$  autour de la trace de l'électron, le  $\chi^2$  de l'association entre l'amas et la trace et le paramètre d'impact de la trace par rapport au vertex primaire. L'énergie de l'électron ou du photon reconstruit est corrigée de la réponse du calorimètre. La résolution en énergie du calorimètre pour des objets électromagnétiques peut être paramétrisée par :

$$\frac{\sigma_E}{E} = \frac{N}{E} + \frac{S}{\sqrt{E}} + C \quad (2.4)$$

où  $N$  est lié aux bruits (uranium et électronique),  $S$  aux fluctuations d'échantillonnage et  $C$  est un terme constant. Ces termes, mesurés dans la région centrale [45], valent :

$$N = 0.23 \pm 0.10 \text{ GeV}, \quad S = 0.202 \pm 0.006 \text{ GeV}^{1/2}, \quad C = 0.004 \pm 0.002 \quad (2.5)$$

### 2.2.3.5 La reconstruction des jets

Un jet résulte de la fragmentation d'un quark ou d'un gluon. Il peut être défini à plusieurs niveaux : au niveau partonique, au niveau des particules stables (pions et kaons) résultant des désintégrations successives des hadrons produits lors de la fragmentation et de l'hadronisation du parton initial, et enfin au niveau des tours du calorimètre où les particules déposent leur énergie. Les jets sont reconstruits au niveau calorimétrique. L'algorithme de reconstruction doit vérifier certaines propriétés :

- Les jets reconstruits sont similaires aux niveaux partonique, particule et calorimétrique,
- Il ne doit pas être sensible aux singularités collinéaire et infrarouge (radiations de gluons mous ou durs).
- Les jets reconstruits sont invariants sous une transformation de Lorentz.
- Il ne doit dépendre ni de la géométrie du détecteur ni de la luminosité, c'est-à-dire être peu sensible à une augmentation du bruit d'événements sous-jacents (interaction multiples, partons spectateurs, ...).
- Il doit satisfaire les contraintes imposées par l'environnement informatique, c'est-à-dire être facile à implémenter et peu gourmand en ressources (mémoire, temps de calcul,...).
- Il doit être compatible avec l'algorithme du Run I pour permettre de comparer les résultats.

L'algorithme utilisé par la plupart des analyses du Run II est un algorithme de cône, l'*Improved Legacy Cone Algorithm* [50] utilisant uniquement l'information calorimétrique. Avant la reconstruction, les cellules calorimétriques bruyantes ou isolées sont supprimées par les algorithmes NADA [51] et T42 [52]. Une quadri-impulsion est déterminée pour chaque cellule à partir de sa position  $(\eta, \phi)$  dans le détecteur et de l'énergie déposée en imposant une masse nulle. La quadri-impulsion d'une tour est définie par la somme vectorielle des quadri-impulsions de toutes les cellules la composant. Les tours du calorimètre dont l'énergie transverse excède 500 MeV/c servent de points de départ (graines) à la formation des proto-jets. Les tours situées dans un cône de rayon donné ( $\Delta R = 0.5$  ou  $0.7$ ) sont alors associées itérativement au proto-jet en recalculant sa direction à chaque ajout de tour. Un tel algorithme avec graine est sensible aux radiations de gluons. Pour s'en affranchir une deuxième série de proto-jets est construite en utilisant comme graines les barycentres dans l'espace  $(\eta, \phi)$  de chaque paire de proto-jets. Finalement les proto-jets partageant des tours en commun sont soit fusionnés soit séparés selon que la fraction d'énergie commune aux deux proto-jets soit supérieure ou inférieure à 50%. En cas de séparation, chaque tour commune est associée au proto-jet dont elle est le plus proche dans l'espace  $(\eta, \phi)$ . Seul les jets reconstruits d'énergie transverse  $p_T > 8$  GeV/c sont conservés.

Les jets reconstruits doivent également satisfaire certaines coupures de qualité. En particulier, jets reconstruits incluent des électrons et des photons qu'il faut rejeter. Ces coupures portent sur :

- La fraction d'énergie dans le calorimètre électromagnétique :  $0.05 < Emf < 0.95$ .
- La fraction d'énergie dans le  $CH$  :  $CHf < 0.4$ . Cette coupure permet de supprimer une partie du bruit instrumental.

La résolution en impulsion transverse des jets centraux ( $|\eta_{det}| < 0.4$ ) peut se modéliser par :

$$\frac{\sigma_{p_T}}{p_T} = \sqrt{\frac{N^2}{E^2} + \frac{S^2}{E}} + C^2 \quad (2.6)$$

avec

$$N = 3.57 \pm 1.34 \text{ GeV}, \quad S = 0.893 \pm 0.121 \text{ GeV}^{1/2}, \quad C = 0.058 \pm 0.007 [46]. \quad (2.7)$$

### 2.2.3.6 La correction en énergie des jets

Du fait de biais introduits par le détecteur et par l'algorithme de cône, l'énergie mesurée du jet calorimétrique peut différer sensiblement de l'énergie du jet de particules. C'est cette dernière grandeur qui est importante pour les analyses de physique. L'énergie de chaque jet est ainsi corrigée par un facteur multiplicatif. Cette correction est souvent citée sous son acronyme anglais **JES** (**J**et **E**nergy **S**cale). Nous décrivons ici la correction définie dans la version v5.3 du programme JETCORR [53]. C'est cette version qui est utilisée pour notre analyse. L'énergie du jet de particules  $E_{part}$  est reliée à celle du jet calorimétrique  $E_{cal}$  par :

$$E_{part} = \frac{E_{cal} - E_{offset}}{R_{jet}S_{ooc}} \quad (2.8)$$

où :

- $E_{offset}$  est la valeur moyenne de l'énergie dans le calorimètre ne provenant pas du vertex d'interaction primaire. Ce terme inclut le bruit de l'uranium, le bruit électronique et les événements sous-jacents. Ce terme est déterminé à partir d'événements de biais minimal et paramétrisé en fonction de la pseudorapidité géométrique  $\eta_{det}$  et la multiplicité de vertex primaires reconstruits.
- $R_{jet}$  est la réponse hadronique du calorimètre, c'est-à-dire le facteur reliant l'énergie déposée et l'énergie mesurée. Elle est estimée dans les données réelles en étudiant des événements  $\gamma + jet$  où le jet et le photon sont dos-à-dos. Elle dépend de  $\eta_{det}$ ,  $\phi$  et de l'énergie du jet.
- $S_{ooc}$  représente la fraction de l'énergie du jet déposée en dehors du cône. Cette correction est nécessaire car l'algorithme de reconstruction des jets utilise un cône de taille fixe. Elle est estimée dans la simulation et calculée en fonction de  $\eta_{det}$  et de l'énergie du jet.

La correction peut également prendre en compte les muons présents dans le jet ainsi que les neutrinos associés, notamment pour les jets de saveurs lourdes. Cette correction n'est pas utilisée dans notre analyse. La correction totale et l'erreur systématique associée en fonction de l'énergie et de la pseudorapidité géométrique du jet est donnée sur la

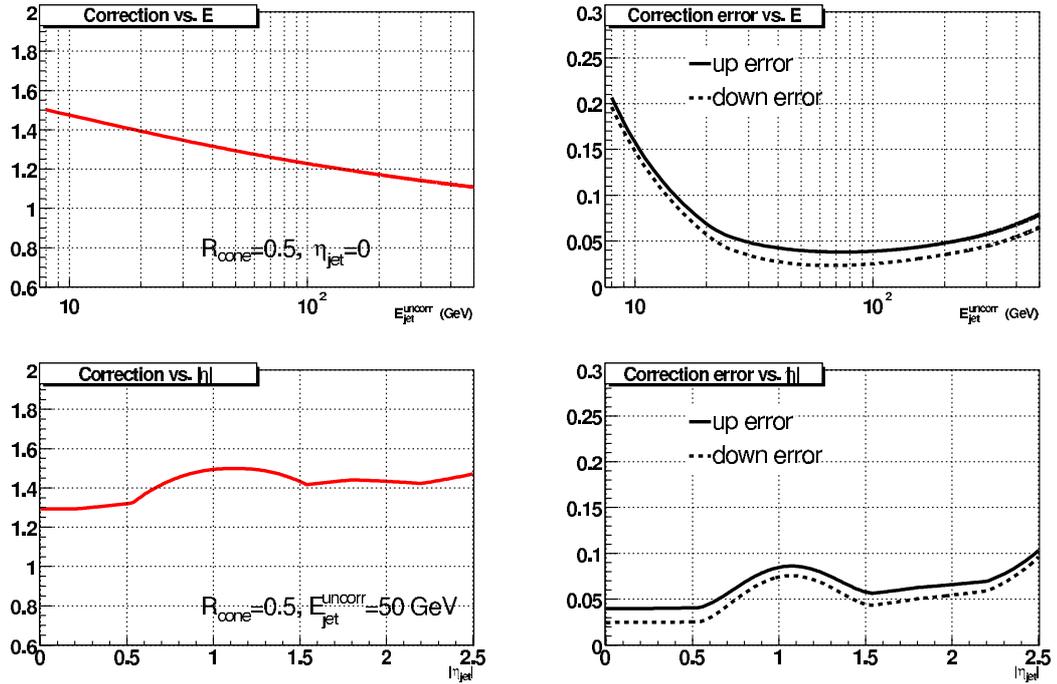


FIG. 2.17: La correction d'énergie des jets (à gauche) et l'erreur absolue asymétrique associée (à droite) en fonction de l'énergie (en haut) et de la pseudorapidité géométrique (en bas) du jet.

figure 2.17. L'erreur sur  $JES$  est une composante importante de l'erreur systématique totale de nombreuses analyses. En particulier c'est un facteur limitant dans la mesure de la masse du quark top. Afin de réduire cette incertitude, les analyses de masse du top dans le mode lepton+jets utilisent la masse du boson  $W$  hadronique pour calibrer la correction en énergie [54]. Pour diminuer cette erreur, il est également possible d'appliquer une correction relative dans la simulation [55].

Cette correction est utilisée pour tout les jets indépendamment de la saveur du parton initial. Pour des jets de saveurs lourdes (principalement les quarks  $b$ ), la fragmentation est plus dure et on s'attend à une correction différente notamment pour l'énergie en dehors du cône. La correction d'énergie pour les saveurs lourdes est actuellement à l'étude dans DØ.

### 2.2.3.7 L'énergie transverse manquante

L'impulsion totale des partons qui interagissent est très faible dans le plan transverse. La conservation de l'énergie impulsion impose donc que le bilan des impulsions transverses de toutes les particules produites lors de l'interaction soit quasiment nul. Les neutrinos de grande impulsion transverse, qui ne sont pas détectés, apparaissent dans ce

bilan sous forme d'énergie transverse manquante. Une grande énergie manquante peut également être la signature d'autres particules stables qui n'interagissent pas dans le détecteur, tel le superpartenaire le plus léger dans certains modèles supersymétriques avec conservation de la R-parité. Cette grandeur est reconstruite à partir des cellules du calorimètre c'est pourquoi on parle d'énergie transverse manquante  $\cancel{E}_T$  plutôt que d'impulsion transverse manquante bien que ces deux notions soit équivalentes [56].

La projection selon  $x$  ou  $y$  de l'énergie transverse manquante brute est définie par l'opposé de la somme des composantes  $E_x$  ou  $E_y$  de la quadri-impulsion de chaque cellule du calorimètre électromagnétique et du  $FH$  :

$$\cancel{E}_{T_{x,y}}^{brute} = - \sum_{\substack{i \in \text{cellules} \\ \text{avec } E > 0}} E_{x,y}^i \quad (2.9)$$

$$\cancel{E}_T^{brute} = \sqrt{\cancel{E}_{T_x}^{brute2} + \cancel{E}_{T_y}^{brute2}} \quad (2.10)$$

Les cellules du  $CH$ , trop bruyantes, ne sont pas utilisées. L'énergie manquante brute est ensuite corrigée pour prendre en compte les corrections d'énergie des objets reconstruits (jets, électrons et photons), de la fraction d'énergie dans le  $CH$  des jets et de l'énergie déposée dans le calorimètre par les muons. Enfin, elle est corrigée de l'impulsion transverse des muons reconstruits par le spectromètre à muons (cf. 2.2.4).

## 2.2.4 Le système à muons

Les muons déposent peu d'énergie dans le calorimètre. Ce sont les seules particules chargées à pouvoir s'en échapper. Un spectromètre à muons est installé à l'extérieur du calorimètre pour les détecter.

### 2.2.4.1 Le spectromètre à muons

Le spectromètre à muons [57] se compose d'une section centrale d'acceptance  $|\eta| < 1$  et de deux sections à l'avant et à l'arrière du détecteur qui étendent la couverture jusqu'à  $|\eta| < 2$ .

Chaque section est composée de trois couches concentriques : A, B et C consistant chacune en un ensemble de tubes à dérive. Dans la partie centrale, les tubes à dérive (*Proportional Drift Tubes, PDT*) sont ceux du Run I. Ils sont orientés selon l'axe  $Oz$ . Dans les sections avant, les *PDT* du Run I ont été remplacés par des mini-tubes à dérive (*Mini-Drift Tubes, MDT*) plus résistants. Un blindage a également été ajouté à l'avant pour protéger le spectromètre des radiations. La partie inférieure du détecteur est moins instrumentée car elle est occupée par la structure supportant l'ensemble du spectromètre.

Trois aimants toroïdaux (un par section) sont installés entre les couches A et B du spectromètre et permettent une mesure de l'impulsion des muons avec un résolution d'environ

30%. La structure en acier des toroïdes sert au retour du champ du solénoïde (cf. figure 2.9).

Des compteurs à scintillation (trois couches à l'avant et deux au centre), situés de part et d'autre des toroïdes complètent le spectromètre. Ces détecteurs à lecture rapide sont utilisés à la fois dans le système de déclenchement et pour la reconstruction des muons. L'information temporelle des scintillateurs joue un rôle important dans la réjection des muons du rayonnement cosmique.

L'agencement des tubes à dérive et des scintillateurs est présenté sur la figure 2.18. L'ensemble du spectromètre à muons est présenté sur la figure 2.4.

#### 2.2.4.2 La reconstruction des muons

La reconstruction est réalisée dans un premier temps avec le seul spectromètre à muons, définissant ainsi les **muons locaux**. La reconstruction des muons locaux est réalisée en trois étapes :

- Dans chaque couche A, B et C, les coups sont regroupés pour former des segments droits, le champ magnétique étant nul dans les couches B et C (situées à l'extérieur du toroïde) et négligeable dans la couche A.
- Les segments des couches B et C sont ensuite reliés en un segment BC plus long. Si cette association échoue, les segments B et C sont abandonnés.
- Les segments A et BC sont associés à leur tour. La déviation de direction entre les segments A et BC causée par le champ magnétique dans le toroïde est pris en compte lors de l'ajustement et permet la mesure de l'impulsion du muon. Les muons de basse impulsion peuvent ne pas traverser le toroïde c'est pourquoi les segments A célibataires sont conservés. Les segments BC sont conservés dans la région où l'acceptance de la couche A est réduite (pied du détecteur).

Différentes qualités pour un muon sont définies par le groupe de travail "muon-id" [58] en fonction du nombre de segments associés au muon local. Pour améliorer significativement la résolution en impulsion, un muon local peut être associé à une trace reconstruite dans le trajectographe. La qualité de cette association est caractérisée par une variable de  $\chi^2$ . On obtient alors les **muons globaux**. L'énergie déposée dans le calorimètre par un muon reconstruit est également utilisée pour corriger sa trajectoire et son énergie.

#### 2.2.5 La mesure de la luminosité

Deux détecteurs circulaires composés de 24 secteurs de scintillateurs permettent de compter le nombre d'interactions inélastiques  $p\bar{p}$ . La lecture du signal de chaque secteur est assurée par un photomultiplicateur. Ces détecteurs sont installés sur les bouchons du calorimètre, à l'intérieur du rayon interne du *FPS* et couvrent l'intervalle de pseudorapidité  $2.7 < |\eta| < 4.4$ . Le luminomètre est représenté sur la figure 2.19. La

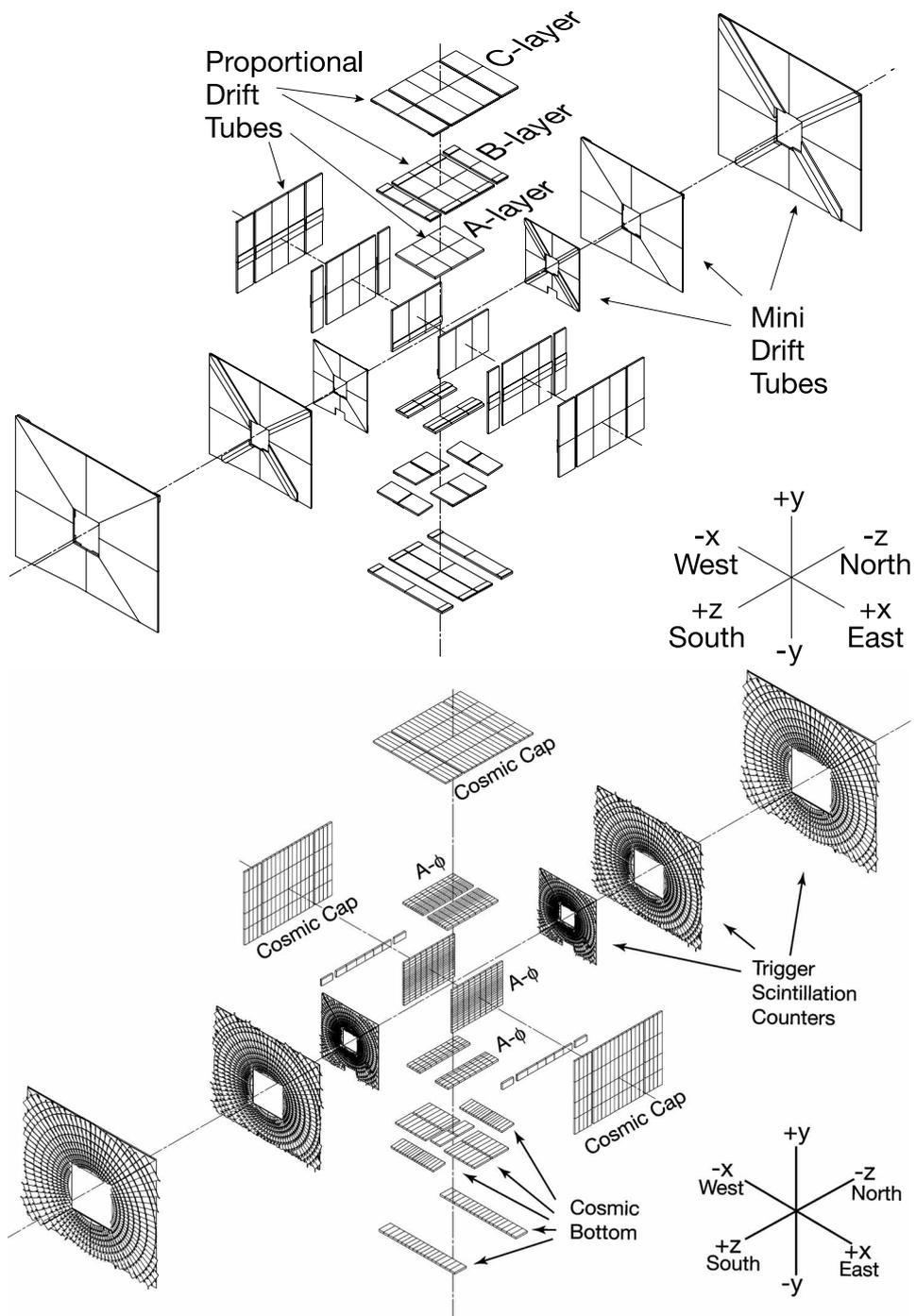


FIG. 2.18: Représentation éclatée du spectromètre à muons : les tubes à dérive (en haut) et les scintillateurs (en bas).

luminosité est estimée à partir du nombre moyen d'interactions inélastiques  $\langle N_{inel} \rangle$ , de la fréquence de croisement des faisceaux  $f$  et de la section efficace effective des interactions inélastiques  $\sigma_{pp}^{eff}$  [59] :  $\mathcal{L} = \frac{f \langle N_{inel} \rangle}{\sigma_{pp}^{eff}}$ . Cette section efficace tient compte de l'acceptance et de l'efficacité du luminomètre. La mesure de la luminosité à DØ est détaillée dans la référence [60]. L'erreur relative sur la luminosité intégrée est de 6.5%.

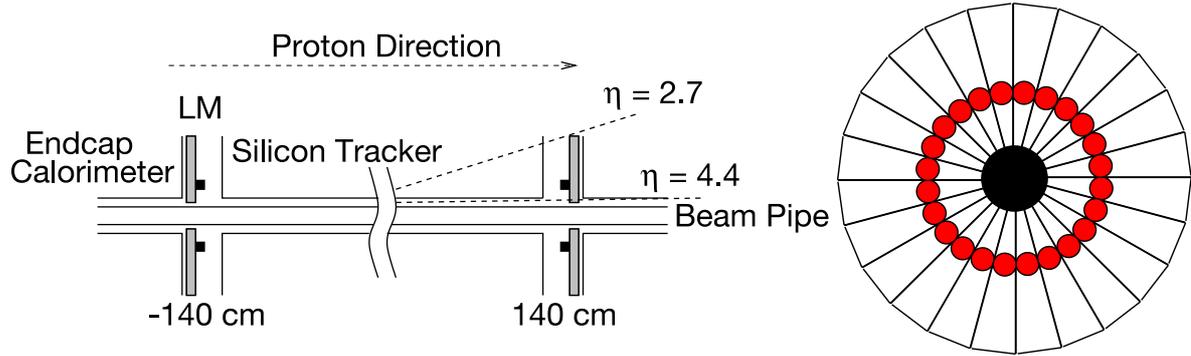


FIG. 2.19: Le luminomètre de DØ .

A gauche : Position du luminomètre entre le trajectographe et les bouchons du calorimètre.

A droite : Schéma de face du luminomètre avec sa structure en 24 segments. Les disques rouges indiquent la position des photomultiplicateurs.

### 2.2.6 Le système de déclenchement

La section efficace totale d'interaction  $p\bar{p}$  vaut environ 70 mb, soit dix milliards de fois la section efficace de production de paires  $t\bar{t}$  (cf. 1.3.2.5). Compte tenu de la luminosité instantanée et de l'acceptance géométrique, le taux d'événements dans le détecteur atteint environ 2.5 MHz. Si l'on prend en compte l'espacement entre les superpaquets, ce taux se ramène à environ 1.7 MHz. Il est impossible d'enregistrer et de reconstruire tous ces événements, la bande passante admissible n'excédant pas 50 Hz. Il est nécessaire de définir des critères précis pour déclencher l'acquisition des données qui sélectionnent majoritairement les événements correspondant à des processus physiques intéressants. Il faut signaler que, vu le large programme de physique d'une expérience du type de DØ et la faible bande passante, le choix de ces critères est souvent délicat. Le système de déclenchement de DØ se compose de trois niveaux successifs :

- Le **niveau 1, L1** : Le premier niveau doit pouvoir traiter rapidement (environ  $3.5 \mu s$ ) et efficacement le très fort taux d'événements provenant du détecteur. Il consiste uniquement en un traitement des données par une électronique dédiée. A ce niveau, seul l'énergie transverse déposée dans les tours du calorimètre (*L1CAL*) et les traces chargées reconstruites à partir des coups dans le *CFT*, le *CPS* et le *FPS* (*L1CTT*) et dans le spectromètre à muons (*L1Muon*) sont prises en compte.

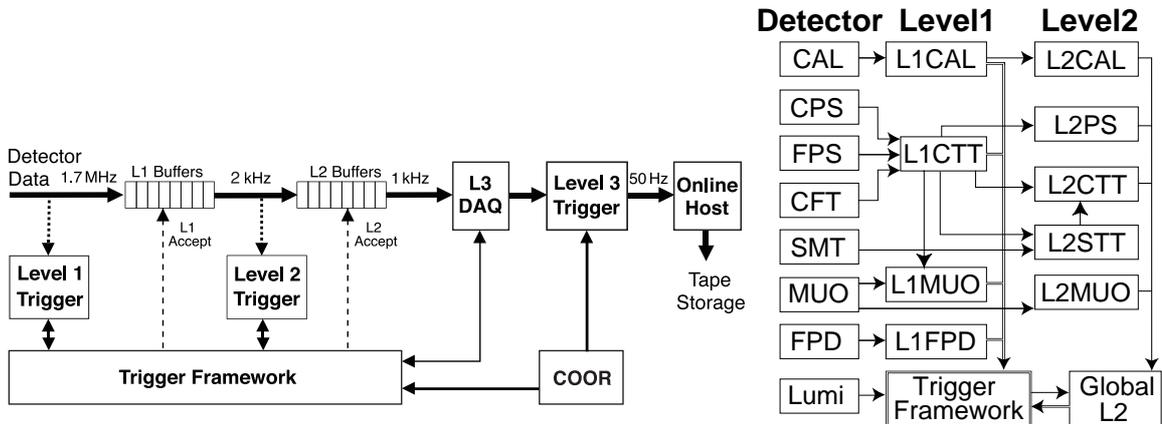


FIG. 2.20: Organigrammes de la chaîne de déclenchement.

A gauche : La chaîne complète.

A droite : Les différentes composantes des niveaux 1 et 2 avec leurs interactions.

Un événement est sélectionné si l'énergie transverse dans une ou plusieurs tours du calorimètre et/ou l'impulsion transverse d'une ou de plusieurs traces dépassent un seuil préalablement fixé. Le taux de sortie du niveau 1 est d'environ 1.6 kHz.

- Le **niveau 2, L2** : Le second niveau réduit le taux d'événements à 1 kHz et dispose d'un temps de décision de  $100 \mu s$ . Différents micro-processeurs sont utilisés pour reconstruire les informations de chaque sous-détecteur qui sont alors corrélées pour former des objets physiques (traces, amas calorimétriques, électrons, jets, muons).
- Le **niveau 3, L3** : Le troisième niveau est la dernière étape du déclenchement. Ce niveau de décision est entièrement informatique et utilise une ferme de processeurs pour exécuter des algorithmes simplifiés de reconstruction des objets physiques. Le temps de décision est de 100 ms et les événements retenus au niveau 3 sont ensuite enregistrés sur bandes, avec un taux de sortie de 50 Hz.

La chaîne de déclenchement est représentée sur la figure 2.20.

## 2.2.7 Le programme de reconstruction et la simulation du détecteur

Les données enregistrées sur bandes sont traitées par le programme DØRECO [61] qui assure la reconstruction des objets physiques. DØRECO produit des fichiers aux formats *DST* et *ThumbNail* [62]. Les données analysées au cours de cette thèse ont été traitées avec la version p14 de DØRECO. Les données sont ensuite réparties en différents lots en fonction de leur contenu physique (événements multijet, événements contenant un ou plusieurs électrons, etc...). Ces lots de données sont ensuite convertis en fichiers exploitables par le programme ROOT [63] qui est utilisée comme base des codes

d'analyse. Les corrections, notamment la correction d'énergie de jets et la calibration des électrons ainsi que les critères d'identification et de qualité des objets physiques sont appliqués lors de cette conversion de format. Deux types de fichiers ROOT ont été utilisés dans cette thèse : TMBTree [64] et TopTree [65], le format d'analyse du groupe de travail "top".

La simulation des processus physiques est initialement réalisée par un générateur d'événements (par exemple PYTHIA [15]). Les événements engendrés sont ensuite processés par le programme DOGSTAR [66] qui simule la réponse du détecteur au moyen du logiciel GEANT 3 [67]. Le programme DOSIM [66] assure finalement la digitisation des événements. A ce stade les événements simulés ont la même structure que les données réelles et la même chaîne de reconstruction est utilisée. Par défaut le système de déclenchement n'est pas simulé. Le programme TRIGSIM [68] peut être utilisé à cet effet. Néanmoins ce programme n'est pas utilisé dans les analyses de physique à cause des différences importantes d'efficacité du système de déclenchement observées entre les données réelles et simulées. Le traitement du déclenchement dans l'analyse sera détaillé au § 4.6.2. Finalement, les résolutions en impulsion transverse des électrons [69], des muons [70] et des jets [70] reconstruits dans la simulation sont dégradées pour reproduire les distributions observées dans les données réelles. Ces modifications sont répercutées sur l'énergie transverse manquante qui présente alors un bon accord avec les données.



Romeo —

*O, she doth teach the torches to burn bright !  
It seems she hangs upon the cheek of night  
Like a rich jewel in an Ethiopie's ear ;  
Beauty too rich for use, for earth too dear !*

William Shakespeare, “*Romeo and Juliet*”, Act I, Scene 5.

# 3

## Identification des jets de quarks beaux

### 3.1 Introduction

Les quarks beaux apparaissent dans de nombreux processus du Modèle Standard étudiés au Tevatron : production QCD de paire  $b\bar{b}$ , production associée de bosons de jauge et de jets ( $Wb\bar{b}$ ,  $Zb\bar{b}$ ,  $\gamma b\bar{b}$ ), désintégration du  $Z$  ( $Z \rightarrow b\bar{b}$ ) et du quark top ( $t \rightarrow Wb$ ) ou encore dans la recherche du boson de Higgs ( $H \rightarrow b\bar{b}$ ). De plus, des quarks  $b$  peuvent également apparaître dans des états finals de physique au-delà du MS, par exemple dans les désintégrations de particules supersymétriques (stop, sbottom). Il est donc important de pouvoir différencier les jets issus de l’hadronisation d’un quark  $b$  des autres jets. Ceci peut se faire grâce aux propriétés particulières du quark  $b$  :

- Les hadrons beaux ne se désintègrent que par interaction faible. De plus cette désintégration est ralentie par la faible valeur du terme de la matrice  $CKM$   $|V_{cb}| = 0.0413 \pm 0.0015$  [5]. Un méson  $B$  a une durée de vie d’environ 1.5 ps soit  $c\tau \approx 400 \mu\text{m}$  [5]. Pour un hadron de 40 GeV/c, cela se traduit par une distance de vol dans le référentiel du laboratoire d’environ 3 mm. Les traces issues des produits de désintégration du  $B$  ne convergent donc plus au vertex primaire mais à un vertex secondaire déplacé, qui peut alors être reconstruit. Par conséquent ces traces possèdent un grand paramètre d’impact (distance minimale d’approche) par rapport au vertex primaire. A DØ, deux algorithmes utilisent cette information pour identifier les jets  $b$ . Le premier, appelé **SVT** (**S**econdary **V**ertex **T**agger) [71], reconstruit des vertex secondaires et identifie les jets  $b$  en fonction

- de la distance de vol de ces vertex. Le second algorithme, nommé *JLIP* (*Jet Lifetime Probability*), qui sera détaillé au § 3.2, combine le paramètre d'impact des traces en une probabilité que le jet soit issu du vertex primaire. *JLIP* a été entièrement développé par le groupe DØ de l'IPHC.
- Le quark  $b$  est plus lourd que les autres quarks (à l'exception du quark top), avec  $m_b \approx 4.2 \text{ GeV}/c^2$  [5]. Il en résulte que d'une part la masse invariante du vertex secondaire de désintégration d'un hadron beau est plus élevée et que d'autre part la fragmentation est plus dure. Les jets de quark  $b$  ont ainsi un cône plus ouvert et les particules associées à un tel jet possèdent une plus grande impulsion transverse relativement à l'axe du jet ( $p_{Trel}$ ).
  - Les hadrons beaux ont un taux de désintégration semi-leptonique plus élevé que les hadrons charmés ou étranges. De plus, l'impulsion transverse relative du lepton est plus élevée que pour une désintégration semi-leptonique d'un hadron charmé ou étrange, comme l'illustre la figure 3.1. Les jets de quark  $b$  peuvent ainsi être identifiés en recherchant un lepton dans le cône du jet, en imposant éventuellement une coupure inférieure supplémentaire sur son  $p_{Trel}$ . Cette méthode d'identification est également utilisée à DØ, particulièrement pour les désintégrations semi-muoniques de hadrons beaux. Il est en effet plus aisé de reconstruire un muon dans un jet qu'un électron. Electrons et jets déposent leur énergie dans le calorimètre et il est difficile de séparer l'électron du reste du jet. Cette méthode n'est donc applicable que dans le cas d'une désintégration semi-muonique, ce qui correspond à environ 18% des jets de  $b$  (en incluant les décroissances directe  $b \rightarrow c\mu\nu_\mu$  et indirecte  $b \rightarrow Xc \rightarrow Xs\mu\nu_\mu$ ) et dépend principalement de l'efficacité de reconstruction des muons. Cette méthode est donc globalement moins efficace que celles basées sur le temps de vol, qui s'applique à tous les jets  $b$ .

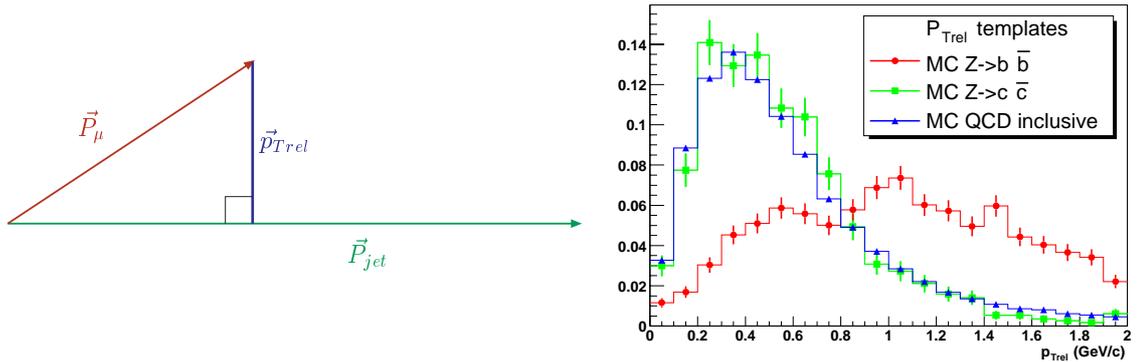


FIG. 3.1: A gauche : la définition de l'impulsion transverse du muon relativement à l'axe du jet,  $p_{Trel}$ . A droite : le spectre en  $p_{Trel}$  pour des muons dans des jets simulés de quarks  $b$  (disques rouges), de quarks  $c$  (carrés verts) et de quarks légers (triangles bleus).

Que ce soit les algorithmes basés sur le temps de vol ou sur la reconstruction de muons dans les jets, ces méthodes identifient les saveurs lourdes : principalement les jets  $b$  mais aussi des jets de quarks  $c$  ( $c\tau \approx 100 \mu\text{m}$ ,  $m_c \approx 1.4 \text{ GeV}/c^2$ ) dans une moindre mesure. Plus marginalement, certaines désintégrations hadroniques du lepton  $\tau$  ( $c\tau = 87 \mu\text{m}$ ,  $m_\tau = 1.8 \text{ GeV}/c^2$ ) peuvent également contribuer. Nous allons maintenant détailler l'algorithme *JLIP* au travers de sa description théorique et de son implémentation dans l'expérience  $D\bar{O}$ . Le processus de validation de l'algorithme pour son utilisation dans des analyses de physique sera expliqué. Cet algorithme sera par la suite utilisé dans l'étude de la production électrofaible du quark top, présentée au § 4.

## 3.2 L'algorithme *JLIP*

L'algorithme *JLIP* est basé sur une méthode d'identification des jets  $b$  utilisant le paramètre d'impact des traces associées à un jet. Pour chaque trace, la probabilité de temps de vol, c'est à dire la probabilité que cette trace provienne du vertex primaire, est dérivée de son paramètre d'impact. La probabilité de chaque trace du jet est ensuite globalement combinée pour obtenir une probabilité de temps de vol par jet. Cette méthode a déjà été mise en œuvre dans plusieurs expériences du LEP [72].

### 3.2.1 Significance en paramètre d'impact

Le paramètre d'impact d'une trace chargée est la distance minimale entre la trajectoire de la trace, représentée par une hélice, et le vertex primaire de l'interaction. Compte tenu de la résolution de reconstruction des traces et du vertex, le paramètre d'impact est proche de 0 pour des traces dues à des particules provenant du vertex primaire. C'est une grandeur définie dans l'espace à trois dimensions, qui peut se décomposer en une composante transverse et une composante longitudinale, selon l'axe du faisceau. La moins bonne résolution selon l'axe  $z$  du faisceau nous contraint à n'utiliser que l'information dans le plan transverse. La trajectoire hélicoïdale d'une particule chargée dans un champ magnétique orienté selon un axe  $z$  est caractérisée par cinq paramètres :

- $IP_0$  : le paramètre d'impact transverse par rapport à l'origine des coordonnées,
- $IP_z$  : le paramètre d'impact longitudinal,
- $\phi_0$  : l'angle azimutal au péricée,
- $\tan \lambda$  : le pas de l'hélice,
- $q/p_T$  : la charge de la particule divisée par son impulsion transverse,

Le paramètre d'impact dans le plan transverse dépend à la fois de la position du vertex primaire  $\vec{v} = (v_x, v_y, v_z)$  et de la position de la trace dans le plan transverse,  $IP_0$  :

$$IP_{xy} = IP_0 - (\vec{e} \cdot \vec{v}), \quad (3.1)$$

où  $\vec{e} = (\sin \phi_0, -\cos \phi_0, 0)$  est le vecteur unitaire perpendiculaire à la direction de la trace dans le plan transverse. Pour une trace associée à un jet,  $IP_{xy}$  est signé par rapport à

l'axe du jet, en fonction du signe de  $\cos \theta$  où  $\theta$  désigne l'angle entre le paramètre d'impact et l'axe du jet, comme illustré sur la figure 3.2.

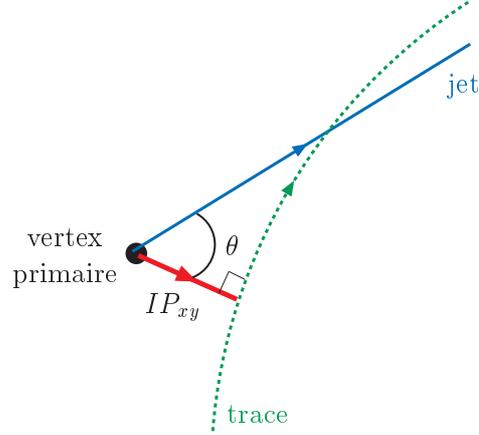


FIG. 3.2: Définition du paramètre d'impact transverse  $IP_{xy}$  signé par l'angle  $\cos \theta$ .

La résolution sur le paramètre d'impact dépend donc de la qualité de reconstruction du vertex et de la trace. Plutôt que  $IP_{xy}$ , il est préférable d'utiliser la signification en paramètre d'impact, c'est-à-dire le paramètre d'impact divisé par son erreur :

$$S_{xy} = IP_{xy} / \sigma_{IP_{xy}}, \quad (3.2)$$

afin d'atténuer certains effets de résolution. La distribution en signification signée pour des traces issues du vertex primaire doit être symétrique et centrée en 0, la dispersion ne provenant que de la résolution sur les traces et le vertex. Au contraire, des traces provenant d'un vertex secondaire déplacé doivent avoir une grande signification positive. La figure 3.3 représente la distribution en signification signée pour différents saveurs de jet. Les jets de saveurs lourdes ( $b$  et  $c$ ) présentent comme prévu un excès dans la région de signification positive. C'est également le cas, dans une moindre mesure, pour les jets légers ( $u$ ,  $d$ ,  $s$  et gluons). Ceci résulte de traces provenant de  $V^0$  (particules à grand temps de vol qui se désintègrent avant le détecteur de traces, telles le  $\Lambda$  ou le  $K_S^0$ ) présents dans les jets légers. La partie négative ne contient à priori que des traces du vertex primaire et devrait avoir la même forme quelle que soit la saveur des jets. Néanmoins on observe un petit excès aux grandes significances négatives pour les jets  $b$  qui s'explique par des traces de signification positive mal signée.

### 3.2.2 Probabilité de temps de vol par trace et par jet

La partie négative de la distribution en signification signée dans un lot d'événements dominés par des jets légers ne contient que des traces issues du vertex primaire. Partant du principe que cette distribution en signification est symétrique pour les traces provenant du vertex primaire, la partie négative peut être interprétée comme une densité de

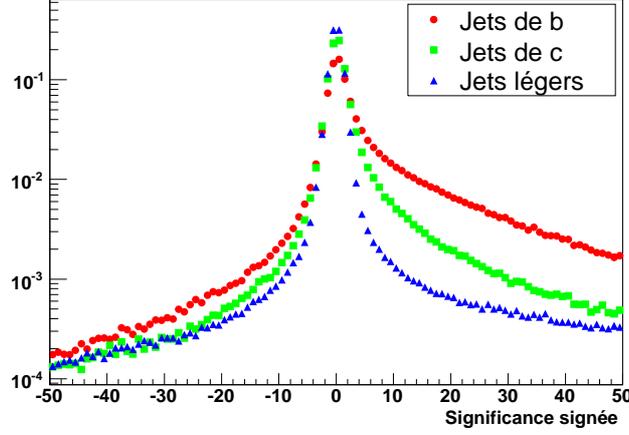


FIG. 3.3: Distributions normalisées de la signification signée pour des jets de quarks  $b$  (disques rouges), de quarks  $c$  (carrés verts) et des jets légers (triangles bleus) dans des événements simulés.

probabilité  $R(|S_{xy}|)$  d'avoir une trace de signification  $\pm|S_{xy}|$  issue du vertex primaire. Cette densité de probabilité peut être estimée directement sur des données réelles ou simulées. La probabilité pour une trace quelconque, de signification  $S$  de venir du vertex primaire est l'intégrale :

$$\mathcal{P}_{trace}(S) = \int_{|S|}^{+\infty} R(s) ds \quad (3.3)$$

Par construction, la distribution de la probabilité par trace est uniforme entre 0 et 1 pour des traces provenant du vertex primaire. Les traces issues de vertex déplacés ayant une grande signification, leur distribution de probabilité doit piquer à 0. Ensuite, la probabilité par jet, combinant  $N$  traces, s'obtient par [72] :

$$\mathcal{P}_{jet} = \Pi \sum_{j=0}^{N-1} \frac{(-\log \Pi)^j}{j!}, \text{ avec } \Pi = \prod_{i=1}^N \mathcal{P}_{trace}(S_i) \quad (3.4)$$

Cette probabilité par jet peut également être construite avec seulement des traces de même signe de signification, soit :

$$\mathcal{P}_{jet}^{\pm} = \Pi^{\pm} \sum_{j=0}^{N-1} \frac{(-\log \Pi^{\pm})^j}{j!}, \text{ avec } \Pi^{\pm} = \prod_{S_i \gtrless 0} \mathcal{P}_{trace}(S_i) \quad (3.5)$$

Comme les traces issues de particules à grand temps de vie possèdent des significances positives, c'est la probabilité positive  $\mathcal{P}_{jet}^+$  qui est utilisée pour identifier les jets  $b$ , en imposant une coupure inférieure,  $\mathcal{P}_{jet}^+ < C$ . La coupure  $C$  fixe le point de fonctionnement de l'algorithme. Comme la distribution de  $\mathcal{P}_{jet}^+$  est presque constante pour des

jets légers (à l'exception de quelques jets contenant des  $V^0$ ), la valeur de  $C$  donne approximativement l'efficacité d'étiqueter un jet léger, c'est à dire le taux de mauvaise identification. Cette valeur peut être ajustée afin d'optimiser l'efficacité d'étiquetage et le taux de mauvaise identification pour une analyse donnée. La probabilité négative  $\mathcal{P}_{jet}^-$  ne sert pas pour l'identification des jets  $b$ . En revanche elle joue un rôle important dans le processus de certification de l'algorithme pour estimer précisément le taux de mauvaise identification.

### 3.3 Lots de données pour la calibration et la certification

Les données utilisées ont été reconstruites avec les versions p14.03.00  $\rightarrow$  p14.06.01 du programme de reconstruction de  $D\bar{O}$  [61]. Ces données ont été enregistrées entre juillet 2002 et août 2004 (runs 162000 à 197000) et couvrent l'ensemble des données p14 pass2, soit 350 à 400 pb $^{-1}$ .

La calibration de l'algorithme requiert des événements multijet, contenant principalement des jets légers :

- Les données réelles utilisées correspondent à environ 830000 événements du lot EM1TRK traités au format TopTree avec la version lpanema du programme TopAnalyze [73]. La sélection de ce lot demande au moins un objet électromagnétique d'impulsion transverse  $p_T > 8$  GeV associé à une trace. Ce lot contient un nombre important de jets.
- Les événements simulés utilisés représentent 640000 événements multijet, générés avec PYTHIA, dans deux gammes d'impulsion transverse partonique ( $40 < p_T < 80$  GeV et  $80 < p_T < 160$  GeV). Ces événements ont été traités au format TopTree avec la version Stradivarius updated de TopAnalyze.

Le choix de ces lots de données ne résulte pas uniquement d'arguments physiques mais également des contraintes liées au fonctionnement de l'expérience. La calibration de l'algorithme sur les données est la première étape d'une longue procédure de certification. Afin d'avoir des algorithmes d'identification de jets  $b$  utilisables dans les analyses le plus rapidement possible, cette calibration doit se faire dans des conditions où l'ensemble des outils d'analyse ne sont pas forcément stables. En fait, seule la reconstruction des traces et dans une moindre mesure l'identification des jets calorimétriques doivent être finalisées, puisque ce sont les traces dans des jets qui sont utilisées. A l'époque où cette calibration a été réalisée seul le lot EM1TRK était disponible avec une statistique suffisante pour fournir des événements multijet. Le désaccord entre les versions de TopAnalyze entre les données réelles et simulées n'affecte pas la reconstruction des traces. La certification de l'algorithme nécessite beaucoup plus de données pour les mesures d'efficacités. Tous les outils d'identification doivent être finalisés, en particulier la reconstruction et la calibration en énergie des jets et l'identification des muons. Les données

utilisées proviennent des lots du *Common Sample Group* [74]. Les lots de données simulées servent principalement à déterminer des facteurs correctifs aux efficacités mesurées dans les données. Tous ces lots ont été processés au format TMBTree avec les corrections du programme `d0correct v8.2` [70], à l'exception des événements Monte-Carlo  $t\bar{t}$  processés au format TopTree avec la version `lpanema` de TopAnalyze. Les critères d'identification des objets sont consistants pour ces versions de TMBTree et de TopTree.

- Les lots de données réelles utilisés pour la certification de *JLIP* sont :
  - EM1TRK (environ 1.8 millions d'événements). Ces données sont similaires à celles du lot de calibration avec une statistique plus grande. Ce lot est utilisé pour mesurer le taux de mauvaise identification.
  - QCD (environ 1.4 millions d'événements). Ces données sont constituées d'événements déclenchés sur des jets. Ce lot est également utilisé pour mesurer le taux de mauvaise identification.
  - BID (environ 21 millions d'événements). Ce lot est constitué d'événements contenant un muon *loose*<sup>1</sup> de  $p_T > 4$  GeV/c associé à un jet dans un cône de rayon  $R = 0.5$ . De plus, chaque événement doit avoir au moins deux jets de  $p_T > 15$  GeV/c et de pseudorapidité  $|\eta| < 2.5$ . La présence d'un muon dans un jet est souvent due à la présence de saveur lourde. Ce lot est donc enrichi en jet de  $b$  et sert aux mesures d'efficacité d'étiquetage.
- Tous les lots d'événements simulés ont été produits avec le générateur PYTHIA à l'exception des événements  $t\bar{t}$  produits avec ALPGEN (pour la formulation théorique des éléments de matrice) et PYTHIA (pour les phases de radiations initiales et finales, d'interactions multiples et de fragmentation) :
  - Environ 640000 événements multijet, dans deux gammes d'impulsion transverse partonique ( $40 < p_T < 80$  GeV/c et  $80 < p_T < 160$  GeV/c) utilisés pour déterminer des facteurs correctifs au taux de mauvaise identification.
  - Environ 100000 événements  $Z \rightarrow b\bar{b} \rightarrow \mu X$  et 63000 événements  $t\bar{t} \rightarrow b\bar{b} \rightarrow \mu X$ , où chaque événement contient un  $b$  avec décroissance semi-leptonique. (Le lot  $t\bar{t}$  se compose de 26000 événements  $\ell^+ \ell^- \nu_\ell \bar{\nu}_\ell b\bar{b}$  et 36000  $\ell \bar{\nu}_\ell j j b\bar{b}$ , avec  $\ell = e$  ou  $\mu$ ).
  - Environ 100000 événements  $Z \rightarrow b\bar{b}$ , 95000 événements  $t\bar{t}$  (dont 36000  $\ell^+ \ell^- \nu_\ell \bar{\nu}_\ell b\bar{b}$  et 59000  $\ell \bar{\nu}_\ell j j b\bar{b}$ ) et 200000 événements dijets  $b\bar{b}$  dans deux gammes d'impulsion transverse partonique ( $40 < p_T < 80$  GeV/c et  $80 < p_T < 160$  GeV/c).
  - Environ 100000 événements  $Z \rightarrow c\bar{c}$  et 200000 événements dijets  $c\bar{c}$  ( $40 < p_T < 80$  GeV/c et  $80 < p_T < 160$  GeV/c).

Dans les lots Monte-Carlo, la saveur des jets est déterminée en recherchant des hadrons beaux ou charmés générés dans un cône de rayon  $R = 0.5$  autour de l'axe du jet. Un jet sera traité comme jet  $b$  s'il contient au moins un hadron beau, comme un jet  $c$  s'il contient au moins un hadron charmé mais aucun hadron beau et comme un jet de

<sup>1</sup>Les muons reconstruits sont classés en *tight*, *medium* ou *loose* suivant des critères d'identification plus ou moins sévères

saveurs légères dans tous les autres cas. Les jets légers incluent de fait les jets de quarks  $u$ ,  $d$  et  $s$  ainsi que les jets de gluons.

## 3.4 Calibration de l'algorithme *JLIP*

### 3.4.1 Qualité des jets

Les jets dans  $D\bar{O}$  sont reconstruits à partir de la seule information calorimétrique. L'algorithme *JLIP* repose uniquement sur les traces chargées associées à un jet. Il faut donc introduire un critère de qualité supplémentaire pour s'assurer que les jets auxquels l'algorithme est appliqué possèdent des traces. C'est le critère de **taggabilité**. Cette *taggabilité* permet de découpler les effets physiques des effets d'inefficacité du détecteur de traces dans les mesures d'efficacité de l'algorithme. Elle doit ainsi, dans une certaine mesure réduire d'éventuelles différences d'efficacité entre données réelles et événements simulés. La *taggabilité* est définie en associant un jet de traces au jet calorimétrique. Les jets de traces sont reconstruits à partir de traces satisfaisant :

- $p_T > 0.5$  GeV/c,
- au moins un coups dans une couche du détecteur de vertex *SMT* (barillets et disques F),
- $|IP_z| < 0.4$  cm, relativement au vertex primaire,
- $|IP_{xy}| < 0.2$  cm, relativement au vertex primaire.

Les traces de  $p_T > 1$  GeV/c servent de point de départ à la construction des jets de traces. Chacune de ces traces constitue le jet de traces initial. Ensuite toutes les traces vérifiant  $\Delta R(\text{trace}, \text{jet de traces}) < 0.5$  et  $|\Delta z(\text{trace}, \text{jet de traces})| < 0.2$  cm sont associées de manière itérative au jet, en recalculant la quadri-impulsion du jet à chaque itération. Un jet calorimétrique est *taggable* s'il est associé dans un cône de rayon  $R = 0.5$  à un jet de trace constitué d'au moins deux traces. La figure 3.4 illustre la différence d'efficacité de *taggabilité* entre des données réelles (environ 80%) et simulées (environ 90%). La reconstruction des traces dans les jets étant sensiblement meilleure dans la simulation que dans les données réelles, les efficacités de *taggabilité* y sont plus élevées. Entre différents lots de données réelles, l'efficacité de *taggabilité* peut varier en fonction de la période de prise de données ou en fonction de biais introduits par le système de déclenchement. L'efficacité moyenne de *taggabilité* est comprise entre 0.7 et 0.8. Ce critère de *taggabilité* est commun à tous les algorithmes d'identification des jets de quarks  $b$  de  $D\bar{O}$  mais n'est pas optimal pour *JLIP* car ce dernier ne nécessite qu'une seule trace de signification positive associée au jet. Dans l'implémentation de *JLIP* à  $D\bar{O}$ , seules les traces de  $p_T > 1$  GeV/c sont utilisées. La *taggabilité* n'impliquant pas que le jet possède des traces de signification positive, il existe de rares cas de jets *taggables* pour lesquels la probabilité  $P_{jet}^+$  ne peut être calculée. Inversement, un jet possédant une unique trace de  $p_T > 1$  GeV/c et de signification positive n'est pas *taggable* mais peut avoir une probabilité *JLIP* définie.

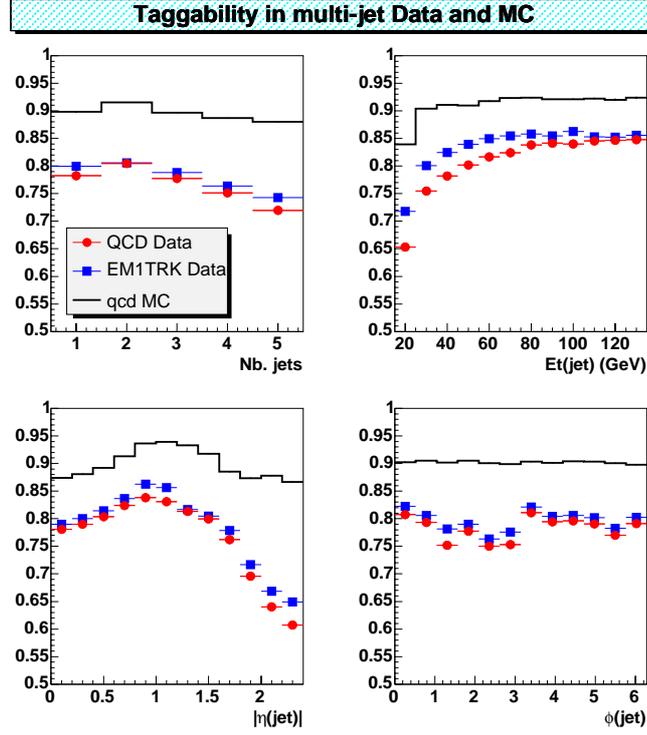


FIG. 3.4: Efficacités de *taggabilité* en fonction de la multiplicité de jets dans l'événement, l'énergie transverse, la pseudorapidité et l'angle azimutal du jet, pour des données multijet (QCD Data, disques rouges), des données multijet avec au moins un objet électromagnétique (EM1TRK, carrés bleus) et des événements multijet simulés (qcd MC, ligne noire).

### 3.4.2 Critères de sélection

Les événements utilisés pour la calibration doivent avoir un vertex primaire dans la région effective du détecteur *SMT* ( $|PV_z| < 60$  cm où  $PV_z$  est la position du vertex primaire sur l'axe du faisceau) et posséder au moins trois traces. Une trace est associée à un jet calorimétrique si elle se trouve dans un cône de rayon  $R = 0.5$  autour de l'axe du jet. Seul les traces associées à un jet *taggable* de  $p_T > 15$  GeV/c et de pseudorapidité  $|\eta| < 2.5$  sont prises en compte. De plus, les traces doivent satisfaire les mêmes critères que ceux utilisés dans la définition d'un jet *taggables* (voir § 3.4.1), avec une coupure supplémentaire en impulsion transverse,  $p_T > 1$  GeV/c. Suivant cette définition les traces retenues pour l'algorithme ne sont pas exactement les mêmes que celles utilisées pour définir la *taggabilité*. La résolution sur le paramètre d'impact d'une trace est sensible à certains paramètres :

- le nombre d'coups dans le *SMT* et le *CFT*,
- la qualité de l'ajustement de la trace, caractérisée par son  $\chi^2$ ,
- l'impulsion transverse,

– la pseudorapidité, qui est fortement corrélée à l'acceptance du *SMT* et du *CFT*. Pour prendre en compte ces effets et ainsi améliorer les performances de l'algorithme les traces sont classées en 29 catégories définies dans le tableau 3.1.

coups <i>SMT</i>	coups <i>CFT</i>	$ \eta $	$\chi^2$	$p_T$ (GeV/c)
1 coups dans la première couche	$\leq 6$	$[1.6, 2.0]$ ou $> 2.0$	$> 0$	$> 1$
1 <i>superlayer</i>	$\geq 7$	$< 1.2$	$[0, 2]$ ou $> 2$	$> 1$
"	"	$> 1.2$	$> 0$	$> 1$
2, 3 ou 4 <i>superlayers</i>	$\geq 7$	$< 1.2$	$[0, 2]$	$[1, 2], [2, 4]$ ou $> 4$
"	"	"	$[2, 4]$ ou $> 4$	$> 1$
"	"	$[1.2, 1.6]$	$[0, 2]$ ou $> 2$	$> 1$
"	"	$> 1.6$	$> 0$	$> 1$

TAB. 3.1: Définition des 29 catégories de traces intervenant dans la calibration. La première ligne correspond aux traces hors de l'acceptance du *CFT* ( $|\eta| > 1.6$ )

### 3.4.3 Réjection des $V^0$

Les jets issus de quarks légers peuvent également contenir des particules neutres à grand temps de vol nommées de manière générique  $V^0$ . Ces particules produisant des vertex déplacés sont de trois types : des kaons neutres  $K_S^0 \rightarrow \pi^+\pi^-$ , des baryons étranges (principalement des  $\Lambda \rightarrow p\pi$ ) et des conversions de photon  $\gamma \rightarrow e^+e^-$ . L'objectif étant d'identifier les jets de saveur lourde, il est utile de rejeter de l'algorithme les traces provenant de  $V^0$ . Pour cela, toutes les traces reconstruites sont utilisées. Pour chaque paire de traces de signe opposé un vertex est calculé et la quadri-impulsion des deux traces est propagée au vertex pour trois hypothèses de masse : deux pions, un proton (associé à la trace de plus grand  $p_T$ ) et un pion et une paire electron-positron. Pour chaque hypothèse, la masse invariante des deux traces est déterminée. Les paires de traces satisfaisant l'un des critères suivants :

$$\begin{aligned}
0.473 < m_{\pi\pi} < 0.517 \text{ GeV}/c^2 & : \text{fenêtre de masse du } K_S^0 \text{ (} m_{K_S^0} = 0.498 \text{ GeV}/c^2\text{)}, \\
1.109 < m_{p\pi} < 1.122 \text{ GeV}/c^2 & : \text{fenêtre de masse du } \Lambda \text{ (} m_{\Lambda} = 1.116 \text{ GeV}/c^2\text{)}, \\
0 < m_{ee} < 0.025 \text{ GeV}/c^2 & : \text{conversion de photon.},
\end{aligned}$$

sont rejetées de l'algorithme d'étiquetage. La figure 3.5 présente les pics de masse de  $K_S^0$  et de  $\Lambda$  observés dans une petite fraction des données. Cette procédure identifie un  $V^0$  dans environ 7% des jets *taggables* et 23% des jets identifiés avec *JLIP*. Ces fractions sont détaillés sur la figure 3.6. Seul 10 à 20% de l'ensemble des  $V^0$  est ainsi rejeté.

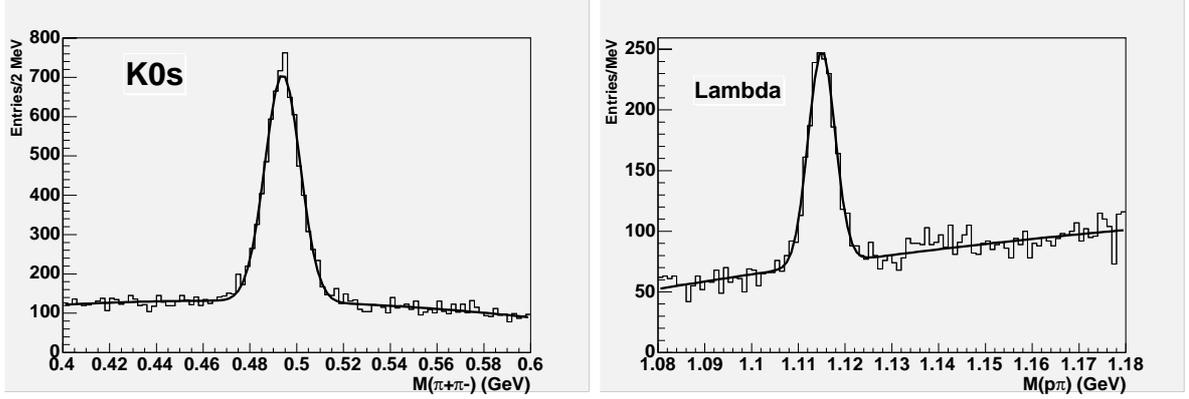


FIG. 3.5: Masses invariantes reconstruites avec deux traces chargées pour les hypothèses  $K_S^0 \rightarrow \pi^+\pi^-$  (à gauche) et  $\Lambda \rightarrow p\pi$  (à droite). Les ajustements correspondent à une gaussienne pour le pic de masse et un polynôme pour le fond.

### 3.4.4 Correction de l'erreur sur le paramètre d'impact

La distribution en signification des traces peut être paramétrée par une composante gaussienne et une queue non-gaussienne (voir figure 3.3). Si l'erreur sur le paramètre d'impact est correctement déterminée, la largeur de la partie gaussienne devrait être égale à l'unité. Ce n'est le cas ni dans les données, ni dans la simulation. Une correction sur l'erreur est donc introduite pour prendre en compte cet effet. La résolution sur le paramètre d'impact dépend à la fois de la résolution sur la trace et sur le vertex primaire. Chacun de ces effets est pris en compte séparément. La partie gaussienne de la distribution provient plus particulièrement des diffusions multiples dans le détecteur de traces. Dans le plan transverse à l'axe du faisceau, la dispersion due aux diffusions est inversement proportionnelle à l'impulsion transverse  $p_T$  et proportionnelle à la racine carrée de la distance parcourue par la particule. Comme le détecteur possède une géométrie cylindrique, selon l'axe du faisceau, cette distance est inversement proportionnelle à  $\sin\theta$ . Cette correction est donc paramétrée en fonction de la variable  $p_{scat}$  définie par :

$$p_{scat} = p_T \sin^{1/2} \theta = p \sin^{3/2} \theta, \text{ où } p \text{ est l'impulsion de la trace.} \quad (3.6)$$

Les 29 catégories de traces sont regroupées en 5 supercatégories en fonction du nombre de coups dans le *SMT* et le *CFT* :

- Une catégorie pour les traces vers l'avant du détecteur avec moins de 7 coups dans le *CFT*, un coup dans la couche interne du *SMT*, et  $|\eta| > 1.6$ ,
- Quatre catégories pour les traces avec au moins 7 coups dans le *CFT* et 1, 2, 3 ou 4 coups dans le *SMT*.

Le découpage en supercatégories et en 15 intervalles de  $p_{scat}$  (de 0 à 35 GeV/c) permet de prendre en compte la dépendance de la correction en fonction de la résolution sur la trace.

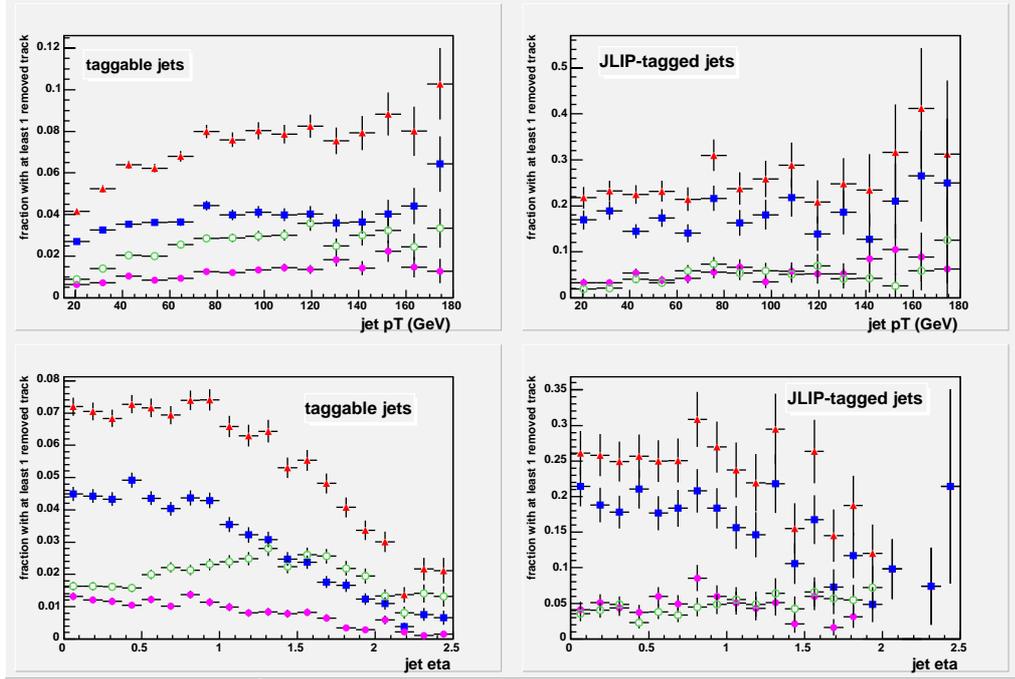


FIG. 3.6: Fractions de  $V^0$  dans des données multijet pour des jets *taggables* (à gauche) et identifiés avec *JLIP* sans réjection des  $V^0$  (à droite) en fonction de l'impulsion transverse (en haut) et de la pseudorapidité (en bas) des jets. Les  $K_S^0$  sont représentés par des carrés bleus, les conversions de photon par des cercles verts, les  $\Lambda$  par des disques roses et l'ensemble des  $V^0$  par des triangles rouges. Le nombre de conversions de photon croît avec la pseudorapidité (jusqu'à  $|\eta| < 2$ ) car la quantité de matière traversée augmente.

Pour chaque supercatégorie et chaque intervalle de  $p_{scat}$ , la partie centrale de la distribution en signficance ( $-15 < S_{xy} < 15$ ) est ajustée avec la somme d'une gaussienne et d'une exponentielle pour modéliser la queue de distribution. La figure 3.7 présente les résultats de tels ajustements. L'écart type de la gaussienne donne la valeur de la correction due à la résolution sur la trace,  $\alpha_{trace}(p_{scat}, catégorie)$ . La paramétrisation en fonction de  $p_{scat}$  de cette correction dans chaque catégorie est donnée sur la figure 3.8. Enfin une dépendance en fonction du nombre de traces du vertex primaire est introduite pour prendre en compte la résolution sur le vertex primaire. Pour cela, le même type d'ajustements que précédemment est réalisé sur les distributions en signficance corrigée du facteur  $\alpha_{trace}$ . La correction d'erreur résiduelle  $\alpha_{vertex}(N_{PV})$  liée au vertex primaire ainsi obtenue est également illustrée par la figure 3.8. Finalement, la signficance corrigée s'écrit :

$$S_{xy}^{corr} = \frac{IP_{xy}}{\alpha_{trace}(p_{scat}, catégorie)\alpha_{vertex}(N_{PV})\sigma_{IP}} \quad (3.7)$$

La résolution corrigée sur le paramètre d'impact est donnée par la figure 3.9. L'effet sur la résolution  $\sigma_{IP}$  des diffusions multiples peut être modélisée en fonction de  $p_{scat}$  sous

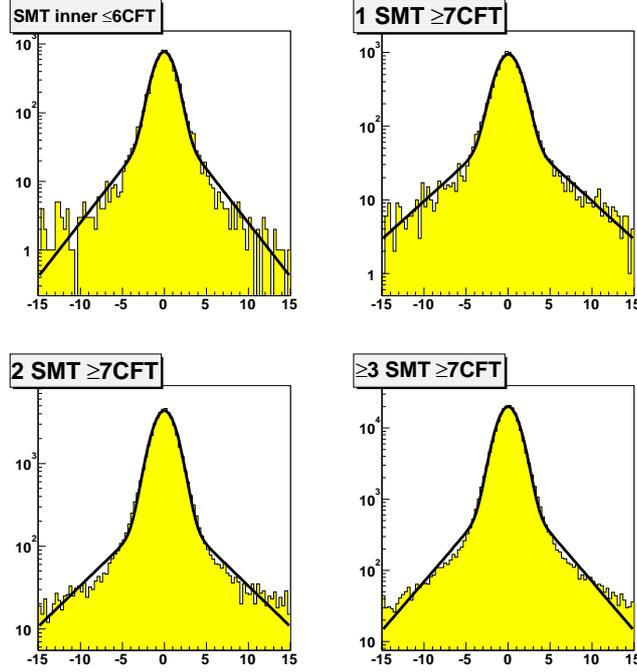


FIG. 3.7: Distributions de signficance pour  $2.5 < p_{scat} < 3$  GeV/c dans les cinq supercatégories (les deux dernières sont regroupées dans le dernier graphe) dans les données EM1TRK avec leurs ajustements.

la forme :

$$\sigma_{IP}^{corr} = \frac{\xi}{p_{scat}} + \zeta, \quad (3.8)$$

où  $\xi$  caractérise l'amplitude des diffusions multiples et  $\zeta$  la résolution asymptotique en l'absence de diffusion, qui varie de 20 à 40  $\mu\text{m}$  selon les catégories de traces.

### 3.4.5 Fonctions de résolution

La dernière étape de la calibration consiste à déterminer les densités de probabilité utilisées dans le calcul de la probabilité de temps de vol (fonction  $R(s)$  de l'équation 3.3). Comme indiqué au § 3.2.2, ces fonctions sont obtenues à partir de la partie négative des distributions de signficance signées pour des événements contenant principalement des jets légers. Pour les données réelles, cette opération s'effectue en utilisant les événements multijet du lot EM1TRK tandis que les événements Monte-Carlo multijet sont utilisés pour les événements simulés. Dans les événements simulés, seul les jets de quarks légers et de gluons sont utilisés.

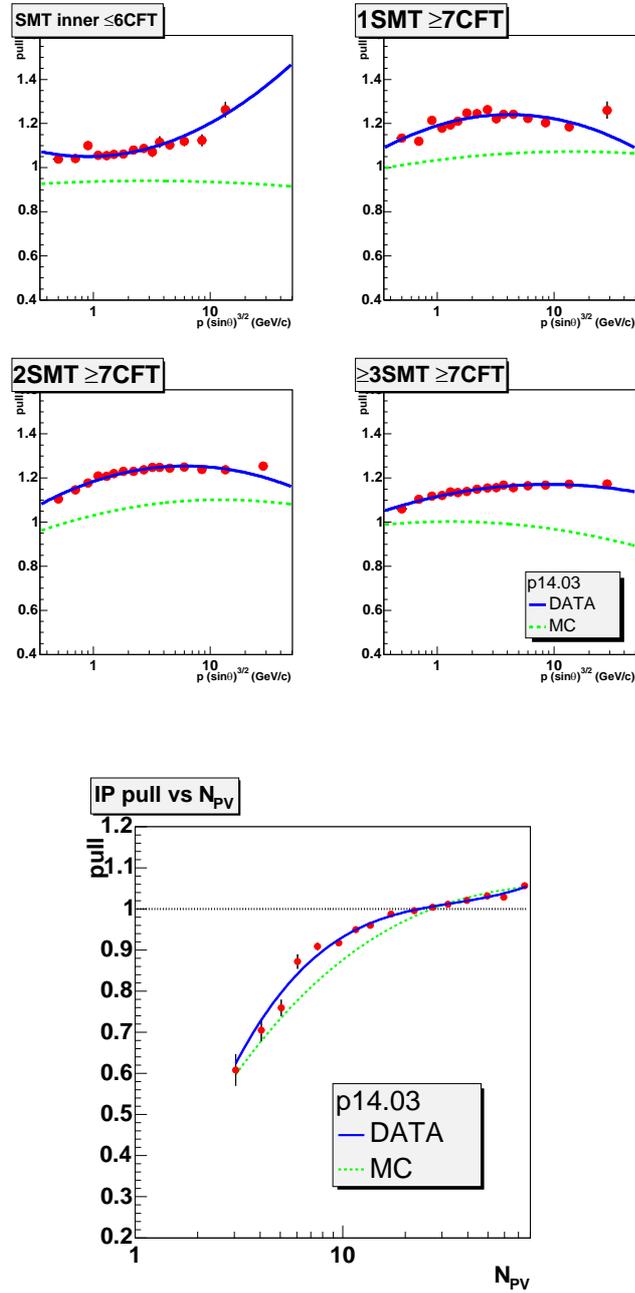


FIG. 3.8: Paramétrisations des corrections d'erreur. En haut, correction liée à la résolution des traces pour les cinq supercatégories (les deux dernières sont regroupées dans le dernier graphe) en fonction de  $p_{scat}$ . En bas, correction résiduelle liée à la résolution sur le vertex primaire en fonction du nombre de traces du vertex. Les données réelles sont en trait continu bleu, la simulation en pointillés verts.

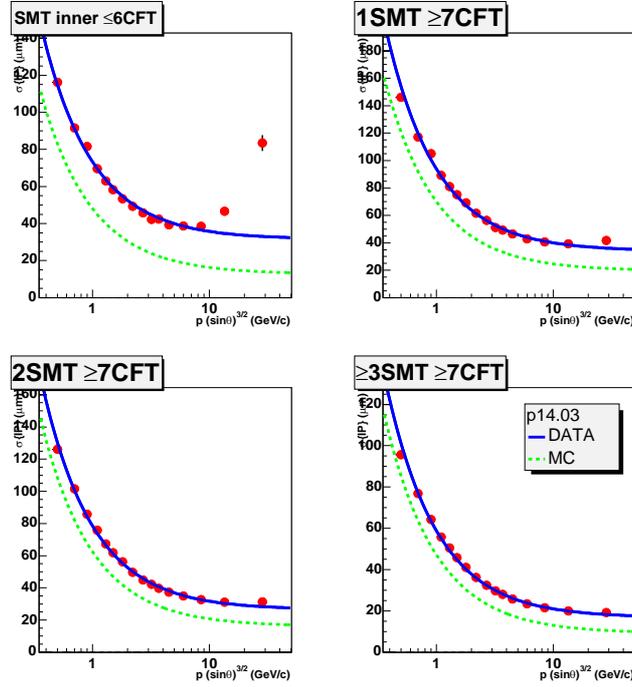


FIG. 3.9: Résolution corrigée sur le paramètre d'impact des traces pour les cinq supercatégories (les deux dernières sont regroupées dans le dernier graphe). Les données réelles sont en trait continu bleu, la simulation en pointillés verts.

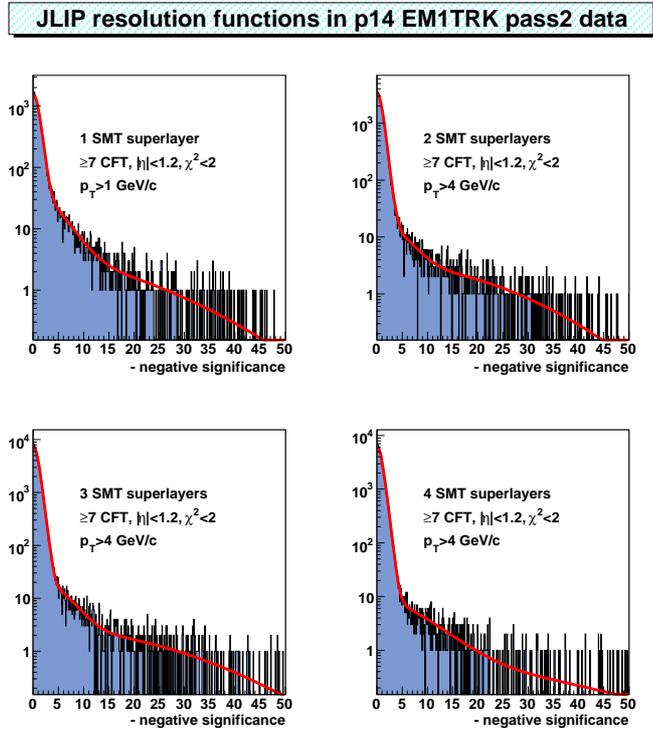


FIG. 3.10: Fonctions de résolution dans 4 des 29 catégories pour les données EM1TRK et leur paramétrisation à l'aide de 4 gaussiennes.

La partie négative des distributions en signification signée des traces sélectionnées est ajustée dans chacune des 29 catégories par une somme de quatre gaussiennes, centrées en 0. Ceci permet une description satisfaisante des queues de distributions, comme illustré sur la figure 3.10. La statistique disponible ne permet pas de modéliser ces fonctions de résolution au delà de  $S_{xy} = 50$ . De plus, les traces de très grande signification peuvent également souffrir de défauts de reconstruction. Les traces ayant une signification supérieure à 50 sont également rejetées de l'algorithme. Cette coupure n'affecte pas les performances de l'algorithme. Dans ce contexte la probabilité par trace se réécrit :

$$\mathcal{P}_{trace}(S_{xy}^{corr}) = \frac{\int_{|S|}^{50} R(s) ds}{\int_0^{50} R(s) ds} \quad (3.9)$$

Les probabilités positives et négatives par trace et par jet obtenues pour des jets simulés de saveur connue et pour les jets du lot de données EM1TRK sont présentées sur la figure 3.11. La probabilité négative par trace  $\mathcal{P}_{trace}^-$  sert à contrôler la qualité des ajustements : elle doit être uniformément distribuée pour les données EM1TRK et les jets légers simulés, qui ont servi à déterminer les fonctions de résolution. Les probabilités positives par jet  $\mathcal{P}_{jet}^+$  présentent le pic à 0 attendu pour les saveurs lourdes ( $b$  et  $c$ ).

### 3.5 Certification de l'algorithme *JLIP*

Afin de valider l'algorithme pour un usage dans l'ensemble de la collaboration, il est nécessaire d'estimer son efficacité avec précision. La mauvaise description des traces de jets dans le Monte-Carlo **p14** conduit à d'importantes différences d'efficacité entre données réelles et simulation. Pour cette raison, les efficacités sont déterminées principalement dans les données et les événements simulés ne servent qu'à la détermination de facteurs correctifs qui sont des rapports d'efficacités. Cette contrainte complique considérablement le processus de certification et est une source d'erreurs systématiques. Au niveau des analyses, deux choix se présentent à l'utilisateur : soit appliquer *JLIP* aux jets simulés et introduire un facteur correctif entre données et Monte-Carlo ; soit ne pas appliquer l'algorithme mais pondérer chaque événement par la probabilité qu'il possède un ou plusieurs jets étiquetés pour reproduire de manière statistique l'effet de l'algorithme. Le calcul de ces pondérations est effectué à partir des efficacités d'identification d'un jet de saveur donnée ( $b$ ,  $c$  ou léger) dans les données. Pour assurer une meilleure description, ces efficacités sont paramétrées en fonction de variables correctement reproduites dans la simulation : l'impulsion transverse  $p_T$  et de la pseudorapidité du jet  $\eta$ . L'impulsion transverse est sensible à la multiplicité et à l'impulsion transverse des traces dans le jet ; La pseudorapidité permet d'introduire une dépendance en fonction de l'acceptance du *SMT* et du *CFT*. Les paramétrisations des efficacités sont appelées **TRF** (**T**ag **R**ate **F**unctions). L'une des tâches de la certification consiste donc à fournir des paramétrisations des efficacités d'identification et des facteurs correctifs accompagnés de leurs erreurs systématiques. Comme pour la calibration, seuls les jets de

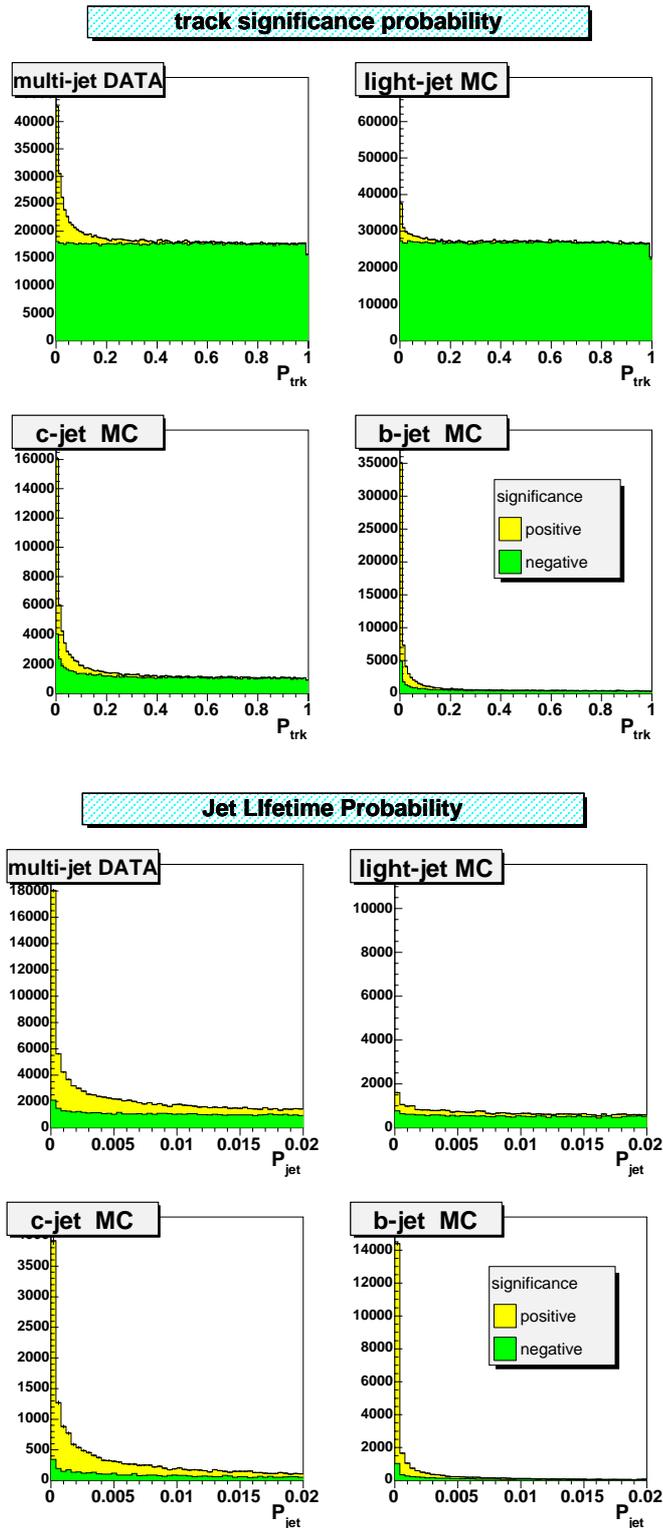


FIG. 3.11: Distributions de probabilité par trace (en haut) et par jet (en bas). Les probabilités par jet sont coupées à 0.02 ; elle sont uniformes entre 0.02 et 1. La probabilité positive est en jaune, la probabilité négative en vert. Pour chaque série d'histogrammes, les jets utilisés sont : les jets du lot EM1TRK (en haut à gauche), des jets légers simulés (en haut à droite), des jets  $c$  simulés (en bas à gauche) et des jets  $b$  simulés (en bas à droite).

d'impulsion transverse  $p_T > 15$  GeV/c et de pseudorapidité  $|\eta| < 2.5$  sont utilisés pour la certification. Cette restriction est due d'une part à la statistique limitée d'événements à grande pseudorapidité, d'autre part aux limites de validité de la correction en énergie des jets. L'algorithme *JLIP* fournit une variable continue ( $\mathcal{P}_{jet}^+$ ), ce qui permet a priori d'optimiser le point de fonctionnement pour chaque analyse. La nécessité de recourir à des paramétrisations d'efficacité oblige à sélectionner un nombre limité de points de fonctionnement pour lesquels ces paramétrisations sont déterminées. Les points de fonctionnement retenus, classés par ordre croissant de sévérité d'étiquetage, sont :

- *ExtraLoose* :  $\mathcal{P}_{jet}^+ < 0.04$
- *SuperLoose* :  $\mathcal{P}_{jet}^+ < 0.02$
- *Loose* :  $\mathcal{P}_{jet}^+ < 0.01$
- *Medium* :  $\mathcal{P}_{jet}^+ < 0.005$
- *Tight* :  $\mathcal{P}_{jet}^+ < 0.003$
- *UltraTight* :  $\mathcal{P}_{jet}^+ < 0.001$

Nous allons maintenant décrire les différentes méthodes utilisées dans cette procédure de certification.

### 3.5.1 La méthode SystemD

SystemD est une méthode de mesure d'efficacité qui a été développée pour répondre au besoin de s'affranchir le plus possible du Monte-Carlo.

#### 3.5.1.1 Un peu de mathématiques

Considérons un cas général où nous disposons de  $n$  lots de données indépendants contenant différentes fractions de signal et d'un ou plusieurs bruits de fond,  $f_{i=0..b}^{j=1..n}$  avec  $b$  le nombre de bruits de fond ( $i = 0$  pour le signal). Pour chaque lot de données  $j$ , nous pouvons écrire :

$$1 = \sum_{i=0}^b f_i^j, \quad (3.10)$$

soit un total de  $b+1$  équations pour  $n \cdot (b+1)$  inconnues. Ajoutons maintenant  $c$  critères de sélection, indépendants les uns des autres, ayant chacun des efficacités de sélection  $\varepsilon_{i=0..b}^{k=1..c}$  du signal et des fonds différentes. Seule une fraction  $Q_j^k$  d'événements passe le critère  $k$  dans lot de données  $j$ . Nous pouvons donc introduire une nouvelle série d'équations :

$$Q_j^k = \sum_{i=0}^b \varepsilon_i^k f_i^j \quad (3.11)$$

Enfin, pour des critères de sélection indépendants, l'efficacité de la composition de plusieurs critères est le produit des efficacités de chaque critère. Pour chaque combinaison

de  $r$  critères de sélection ( $r \leq c$ ), nous avons ainsi une équation :

$$Q_j^{k_1, \dots, k_r} = \sum_{i=0}^b \left( \prod_{v=1}^r \varepsilon_i^{k_v} \right) f_i^j \quad (3.12)$$

L'équation 3.12 généralise les équations 3.10 ( $r = 0$ ) et 3.11 ( $r = 1$ ). Chaque lot de données introduit  $b + 1$  inconnues (fractions de signal et de fonds). De même, chaque critère de sélection ajoute  $b + 1$  inconnues supplémentaires pour les efficacités. Au total on a donc  $(n + c)(b + 1)$  inconnues. Le nombre d'équations disponibles dépend du nombre de combinaisons possibles des  $c$  critères, soit un total de  $n \sum_{r=0}^c \binom{r}{c} = n 2^c$  équations. Pour avoir un système qui admette un nombre fini de solutions, il faut nécessairement que  $n$ ,  $b$  et  $c$  satisfassent l'inéquation :

$$(n + c).(b + 1) \leq n.2^c \quad (3.13)$$

Les premières solutions non triviales, s'obtiennent pour :

- $n = 1, c = 3, b = 1$  : 8 équations et 8 inconnues.
- $n = 2, c = 2, b = 1$  : 8 équations et 8 inconnues.
- $n = 1, c = 4, b = 2$  : 16 équations et 15 inconnues.
- $n = 2, c = 3, b = 2$  : 16 équations et 15 inconnues.

Dans tous les cas, le système d'équations obtenu n'est pas linéaire et peut posséder plusieurs solutions. Nous nous limiterons dans notre étude au cas d'un système de huit équations, baptisé *System8*. Si le système possède des solutions, il n'y en a que deux. Elles sont alors symétriques par rapport à la permutation du signal et du fond de par la structure même du système. Une connaissance a priori d'au moins une des inconnues permet de résoudre l'ambiguïté. Les seuls paramètres du système sont les fractions d'événements  $Q_i^{k_0, \dots, k_r}$  qui sont déterminées sur des données. La résolution du système permet d'extraire les fractions de signal et de fond ainsi que toutes les efficacités de sélection.

Dans un modèle plus réaliste, les critères de sélection peuvent être partiellement corrélés, auquel cas il faut rajouter des facteurs correctifs dans les équations couplant plusieurs critères. Pour un fond  $i$  donné, le facteur correctif  $\kappa_i^{k_1, k_2}$  pour deux critères  $k_1$  et  $k_2$  est défini par :

$$\varepsilon_i^{k_1, k_2} = \kappa_i^{k_1, k_2} \varepsilon_i^{k_1} \varepsilon_i^{k_2} \quad (3.14)$$

Il peut être déterminé à partir d'événements simulés en comptant le nombre total d'événements  $N_i$  et les nombres d'événements passant le critère  $k_1$  ( $N_i^{k_1}$ ), le critère  $k_2$  ( $N_i^{k_2}$ ) et les deux critères ( $N_i^{k_1, k_2}$ ). Le facteur correctif s'écrit alors :

$$\kappa_i^{k_1, k_2} = \frac{\varepsilon_i^{k_1, k_2}}{\varepsilon_i^{k_1} \varepsilon_i^{k_2}} = \frac{N_i N_i^{k_1, k_2}}{N_i^{k_1} N_i^{k_2}} \quad (3.15)$$

Le facteur correctif pour trois critères  $k_1$ ,  $k_2$  et  $k_3$  peut être approximé avec les facteurs correctifs à deux critères si ces facteurs correctifs sont voisins de l'unité par :

$$\kappa_i^{k_1, k_2, k_3} = \kappa_i^{k_1, k_2} \cdot \kappa_i^{k_2, k_3} \cdot \kappa_i^{k_3, k_1} \quad (3.16)$$

Les solutions du système ne sont alors plus symétriques. Il faut toujours recourir à une connaissance a priori d'une solution pour résoudre l'ambiguïté. Dans la pratique la méthode n'est applicable que pour des valeurs de  $\kappa_i^{k_1, k_2}$  proche de l'unité, c'est-à-dire pour de faibles corrélations.

### 3.5.1.2 Application à l'étiquetage des jets de quark $b$

Dans la pratique il est difficile de rassembler toutes les conditions nécessaires à la résolution du système. Il est néanmoins possible, sous certaines hypothèses, d'utiliser la méthode *System8* pour extraire l'efficacité d'un algorithme d'identification des jets  $b$ . Le lot de données utilisé est l'ensemble des jets du lots BID qui vérifient :

- $p_T > 15$  GeV/c,
- $|\eta| < 2.5$ ,
- le jet est *taggable*,
- le jet contient un muon. Seuls les muons *medium*, de  $p_T > 2$  GeV/c, associés à une trace avec un  $\chi^2 < 100$  sont pris en compte. Un muon est associé au jet si  $\Delta R(\mu, jet) < 0.5$ . Si plusieurs muons sont associés au jet, seul celui avec le plus petit  $\chi^2$  est considéré. Finalement le jet est sélectionné si le muon associé a un  $p_T > 4$  GeV/c. Cette dernière coupure est ajoutée pour retirer un biais dans le spectre en  $p_T$  des muons, lié à la définition du lot BID.

L'identification par *JLIP* est naturellement l'un des trois critères de sélection utilisés. Le choix de deux autres critères qui modifient le contenu en jet  $b$  est limité. Le premier est une coupure sur l'impulsion transverse relative (voir figure 3.1) du muon associé au jet. La valeur de la coupure ( $p_{Trel} > 0.7$  GeV/c) est choisie dans un domaine tel que son efficacité soit similaire pour les jets  $c$  et légers. Les jets  $b$  sont généralement produit par paire, la probabilité que le jet muonique soit issu d'un quark  $b$  augmente donc s'il existe dans le même événement un autre jet étiqueté. Cette propriété est utilisée comme dernier critère de sélection. Les trois critères ainsi retenus sont :

- $J$  :  $\mathcal{P}_{jet}^+ < C$ ,  $C$  correspondant à l'un des six points de fonctionnement définis précédemment.
- $M$  :  $p_{Trel} > 0.7$  GeV/c, pour le muon.
- $B$  : présence dans l'événement d'un autre jet *taggable* avec  $\mathcal{P}_{jet}^+ < 0.01$ .

La trace associée au muon peut aussi être utilisée dans le calcul de  $\mathcal{P}_{jet}^+$  introduisant une faible corrélation entre  $J$  et  $M$ . De plus, les paramètres d'impact utilisés par  $J$  et  $B$  sont calculés par rapport au même vertex primaire, ce qui conduit à une autre corrélation entre  $J$  et  $B$ . Ces deux effets sont pris en compte sous forme de facteurs correctifs déterminés dans les événements simulés. A priori, il n'existe pas de corrélation entre  $B$  et  $M$ . Ceci se vérifie dans la simulation et ces facteurs seront fixés à 1 par la suite et n'apparaîtront pas.

Les jets  $b$  constituent le signal recherché. Les jets  $c$  et légers sont regroupés dans le bruit de fond. Ceci n'est possible que si les efficacités sur ces types de jets sont voisines. Ce n'est le cas que pour la coupure sur le  $p_{Trel}$  du muon. Pour les critères  $J$  et  $B$ , cette

hypothèse sur le bruit de fond n'est plus vérifiée car *JLIP* identifie également les jets  $c$ . Les efficacités ne peuvent donc pas être utilisées pour estimer le taux de mauvaise identification. Chacun des deux critères modifiant la composition en saveur, une nouvelle corrélation est introduite qui contribue au facteur correctif de corrélation entre  $J$  et  $B$  déjà évoqué. Comme il n'existe pas de simulation permettant d'estimer les fractions de jets  $c$  et légers dans le lot  $BID$ , le facteur  $\kappa_{cl}^{(B,J)}$  ne peut pas être déterminé. Ce facteur est fixé arbitrairement à 1. Cependant l'efficacité de  $J$  sur les jets  $b$  est peu sensible aux variations de ce facteur correctif comme le montre la figure 3.12. Les variations de la figure 3.12 sont prises en compte dans l'estimation des erreurs systématiques (voir § 3.5.1.4).

*System8* peut donc être utilisé pour mesurer l'efficacité de *JLIP* dans les données

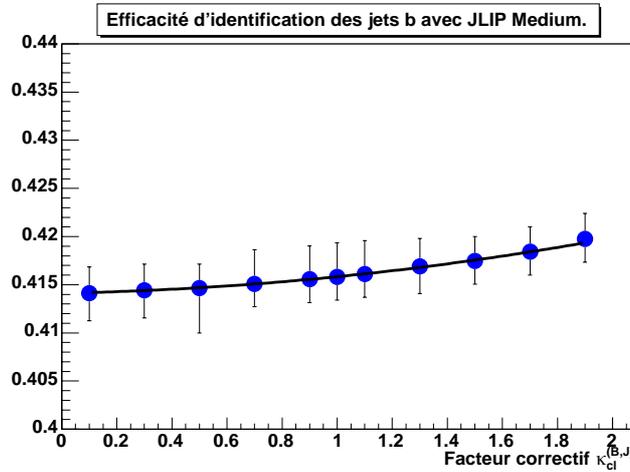


FIG. 3.12: Dépendance de l'efficacité de *JLIP Medium* déterminée avec *System8* en fonction du facteur correctif  $\kappa_{cl}^{(B,J)}$  dans l'intervalle  $[0;2]$ . Les autres paramètres d'entrée du système sont ceux du tableau 3.2.

réelles. Cependant les sept autres résultats de la résolution du système vont être affectés par l'hypothèse sur le fond. Leurs valeurs donnent néanmoins un ordre de grandeur intéressant. Le système d'équations à résoudre est alors :

$$\left\{ \begin{array}{l} f_b + f_{cl} = 1 \\ \varepsilon_b^{(J)} f_b + \varepsilon_{cl}^{(J)} f_{cl} = Q^{(J)} \\ \varepsilon_b^{(M)} f_b + \varepsilon_{cl}^{(M)} f_{cl} = Q^{(M)} \\ \varepsilon_b^{(B)} f_b + \varepsilon_{cl}^{(B)} f_{cl} = Q^{(B)} \\ \kappa_b^{(J,M)} \varepsilon_b^{(J)} \varepsilon_b^{(M)} f_b + \kappa_{cl}^{(J,M)} \varepsilon_{cl}^{(J)} \varepsilon_{cl}^{(M)} f_{cl} = Q^{(J,M)} \\ \varepsilon_b^{(M)} \varepsilon_b^{(B)} f_b + \varepsilon_{cl}^{(M)} \varepsilon_{cl}^{(B)} f_{cl} = Q^{(M,B)} \\ \kappa_b^{(B,J)} \varepsilon_b^{(B)} \varepsilon_b^{(J)} f_b + \kappa_{cl}^{(B,J)} \varepsilon_{cl}^{(B)} \varepsilon_{cl}^{(J)} f_{cl} = Q^{(B,J)} \\ \kappa_b^{(J,M)} \kappa_b^{(B,J)} \varepsilon_b^{(J)} \varepsilon_b^{(M)} \varepsilon_b^{(B)} f_b + \kappa_{cl}^{(J,M)} \kappa_{cl}^{(B,J)} \varepsilon_{cl}^{(J)} \varepsilon_{cl}^{(M)} \varepsilon_{cl}^{(B)} f_{cl} = Q^{(J,M,B)} \end{array} \right. \quad (3.17)$$

Le tableau 3.2 donne les nombres de jets sélectionnés dans l'ensemble du lot BID, les facteurs correctifs et la solution du système 3.17. Le choix de la bonne solution est fait en imposant pour les trois critères que l'efficacité sur le signal soit supérieure à celle sur le fond, ce qui est le comportement attendu :

$$\varepsilon_b^{(J)} > \varepsilon_{cl}^{(J)} \quad (3.18)$$

$$\varepsilon_b^{(M)} > \varepsilon_{cl}^{(M)} \quad (3.19)$$

$$\varepsilon_b^{(B)} > \varepsilon_{cl}^{(B)} \quad (3.20)$$

Total				
	Nombre de jets $N^0$	fraction $Q^0$	fraction $f_b$	fraction $f_{cl}$
	7827900	1	<b>0.26</b>	<b>0.74</b>
Critères de sélection simples				
(x)	Nombre de jets $N^{(x)}$	fraction $Q^{(x)}$	efficacité $\varepsilon_b$	efficacité $\varepsilon_{cl}$
<i>J</i>	999516	0.13	<b>0.42</b>	<b>0.025</b>
<i>M</i>	3492770	0.45	<b>0.66</b>	<b>0.37</b>
<i>B</i>	484023	0.062	<b>0.14</b>	<b>0.033</b>
Critères de sélection combinés				
(x)	Nombre de jets $N^{(x)}$	fraction $Q^{(x)}$	correction $\kappa_b$	correction $\kappa_{cl}$
<i>J, M</i>	606321	0.077	0.99	0.90
<i>M, B</i>	263916	0.034	1	1
<i>B, J</i>	128511	0.016	1.016	1
<i>B, J, M</i>	82382	0.011	1.006	0.90

TAB. 3.2: Application numérique de *System8* pour l'ensemble du lot BID pour *JLIP Medium* ( $\mathcal{P}_{jet}^+ < 0.005$ ). Les deux premières colonnes donnent le nombre et la fraction de jets satisfaisant chaque critère de sélection, simple et combiné. Les fractions  $Q^{(x)}$  sont les paramètres initiaux du système.

Les huit cases en haut à droite (en gras) présentent le résultat de la résolution du système. L'efficacité d'étiquetage des jets *b* mesurée pour *JLIP Medium* est de 42%.

Les huit cases en bas à droite présentent les facteurs correctifs estimés dans la simulation. Le facteur  $\kappa_{cl}^{(B,J)}$  est fixé arbitrairement à 1. Les facteurs correctifs  $\kappa_b^{B,J,M}$  et  $\kappa_{cl}^{B,J,M}$  sont approximés par le produit des facteurs correctifs pour deux critères

Les dépendances en fonction de  $p_T$  et de  $|\eta|$  des jets ont été étudiées séparément dans 6 intervalles d'impulsion transverse et 6 intervalles de pseudorapidité. Ces efficacités sont présentées sur la figure 3.13. Pour chaque intervalle les facteurs correctifs sont recalculés et l'efficacité de *JLIP* est extraite de la résolution du système. Les facteurs correctifs sont calculés dans les différents lots Monte-Carlo disponibles. Les facteurs

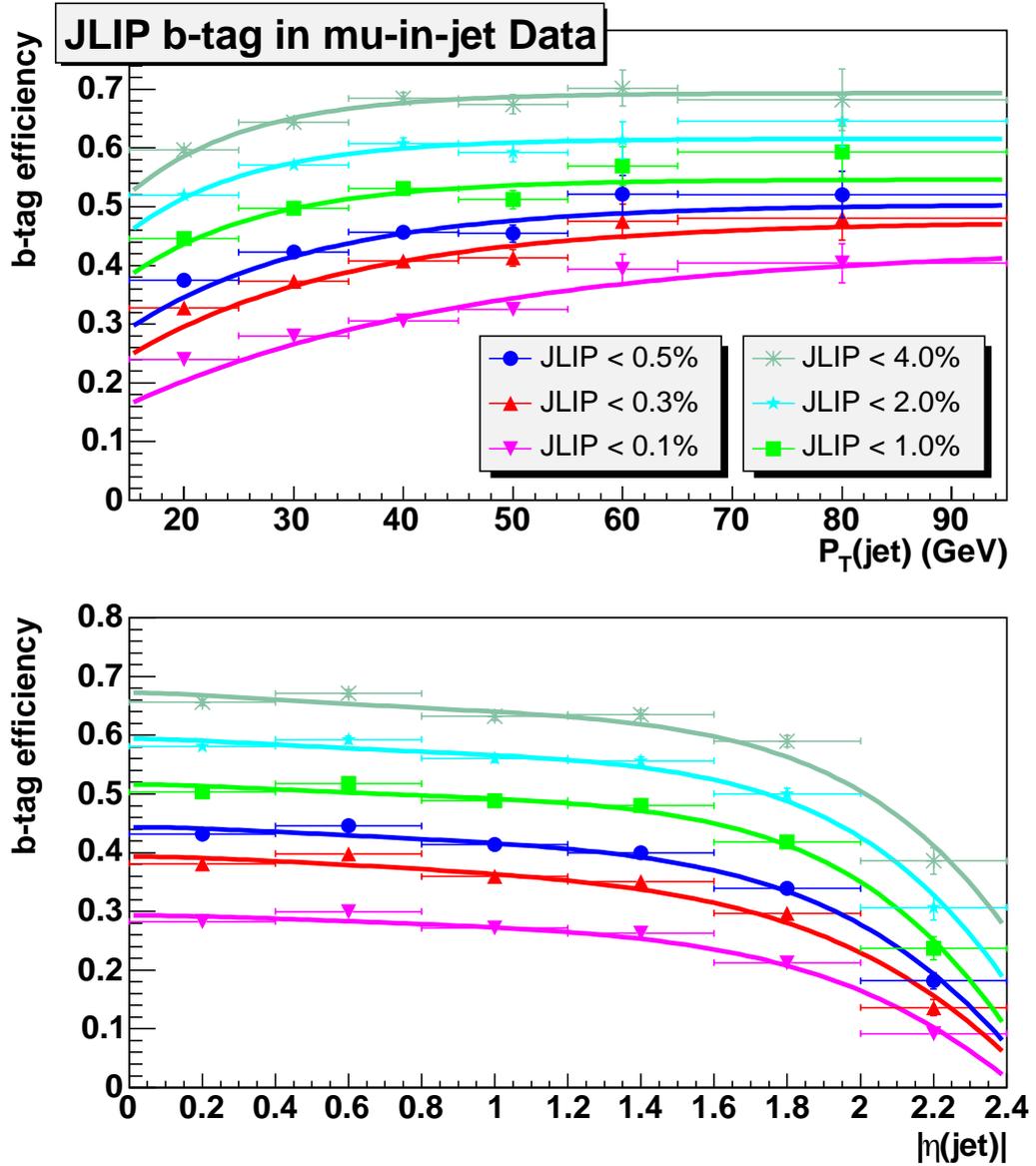


FIG. 3.13: Efficacités de *JLIP* mesurées avec *System8* dans le lot BID, pour les 6 points de fonctionnement : *ExtraLoose* (croix grises), *SuperLoose* (étoiles cyan), *Loose* (carrés verts), *Medium* (disques bleus), *Tight* (triangles rouges) et *UltraTight* (triangles roses), en fonction de l'impulsion transverse (en haut) et de la pseudorapidité du jet (en bas).

$\kappa_b^{(B,J)}$  (figure 3.14) et  $\kappa_{cl}^{(J,M)}$  sont constants et indépendants du processus physique simulé. Le facteur  $\kappa_{cl}^{(J,M)}$  est similaire pour des jets  $c$  et les jets légers. Le facteur  $\kappa_b^{(J,M)}$  (figure 3.15) possède une dépendance en impulsion transverse qui est prise en compte dans la résolution du système par intervalles de  $p_T$ . La valeur moyenne sur le spectre en  $p_T$  de jets est utilisée pour l'étude de la dépendance en pseudorapidité. Il a déjà été vérifié qu'il n'y a pas de corrélations entre  $B$  et  $M$  dans la simulation. Finalement les facteurs correctifs utilisés sont :

$$\begin{aligned}
\kappa_b^{(J,M)} &= (1.03 - 0.00117p_T) \pm 0.004 \\
\kappa_b^{(B,J)} &= 1.016 \pm 0.003 \\
\kappa_{cl}^{(J,M)} &= 0.90 \pm 0.05 \\
\kappa_{cl}^{(B,J)} &= 1 \pm 0.8
\end{aligned} \tag{3.21}$$

Les erreurs sur  $\kappa_b^{(J,M)}$  et  $\kappa_b^{(B,J)}$  proviennent uniquement de la statistique des lots Monte-Carlo. L'erreur sur  $\kappa_{cl}^{(J,M)}$  inclut les variations du facteur correctif entre jets  $c$  et légers. Enfin l'erreur sur  $\kappa_{cl}^{(B,J)}$  est arbitraire et permet de prendre en compte l'incertitude liée à l'hypothèse sur le fond.

### 3.5.1.3 Test de *System8*

Etant donné la complexité de la méthode et le nombre d'hypothèses émises, il est important de vérifier la validité de *System8* sur un lot d'événements simulés. Pour un tel lot, la composition en saveur est connue et il est possible de calculer directement les efficacités de chaque critère de sélection. Ces nombres peuvent ensuite être confrontés aux résultats de *System8*.

Le lot test est composé d'un mélange d'événements  $t\bar{t}$ ,  $Z \rightarrow b\bar{b} \rightarrow \mu X$ ,  $Z \rightarrow c\bar{c} \rightarrow \mu X$  et multijet. Le nombre de jets légers contenant un muon étant très faible, une trace de grande impulsion transverse ( $p_T > 8$  GeV/c) est sélectionnée aléatoirement pour représenter un muon venant de la désintégration de pions ou de kaons. Le lot ainsi constitué compte 430000 jets muoniques (75% de jets légers, 5% de jets  $c$  et 20% de jets  $b$ ). *System8* est appliqué dans 7 intervalles d'impulsion transverse des jets de 15 à 135 GeV/c. Les facteurs correctifs utilisés sont ceux de l'équation 3.21. En particulier, le facteur  $\kappa_{cl}^{(B,J)}$  est maintenu à 1. Les huit résultats de *System8* sont en bon accord avec les valeurs attendues comme le montre la figure 3.16, en particulier l'efficacité de *JLIP*.

### 3.5.1.4 Estimation des erreurs systématiques

De nombreuses sources d'erreurs systématiques sont prises en compte lors de l'évaluation des efficacités d'identification des jets  $b$ . Elle sont estimées de manière conservatrice et indépendante.

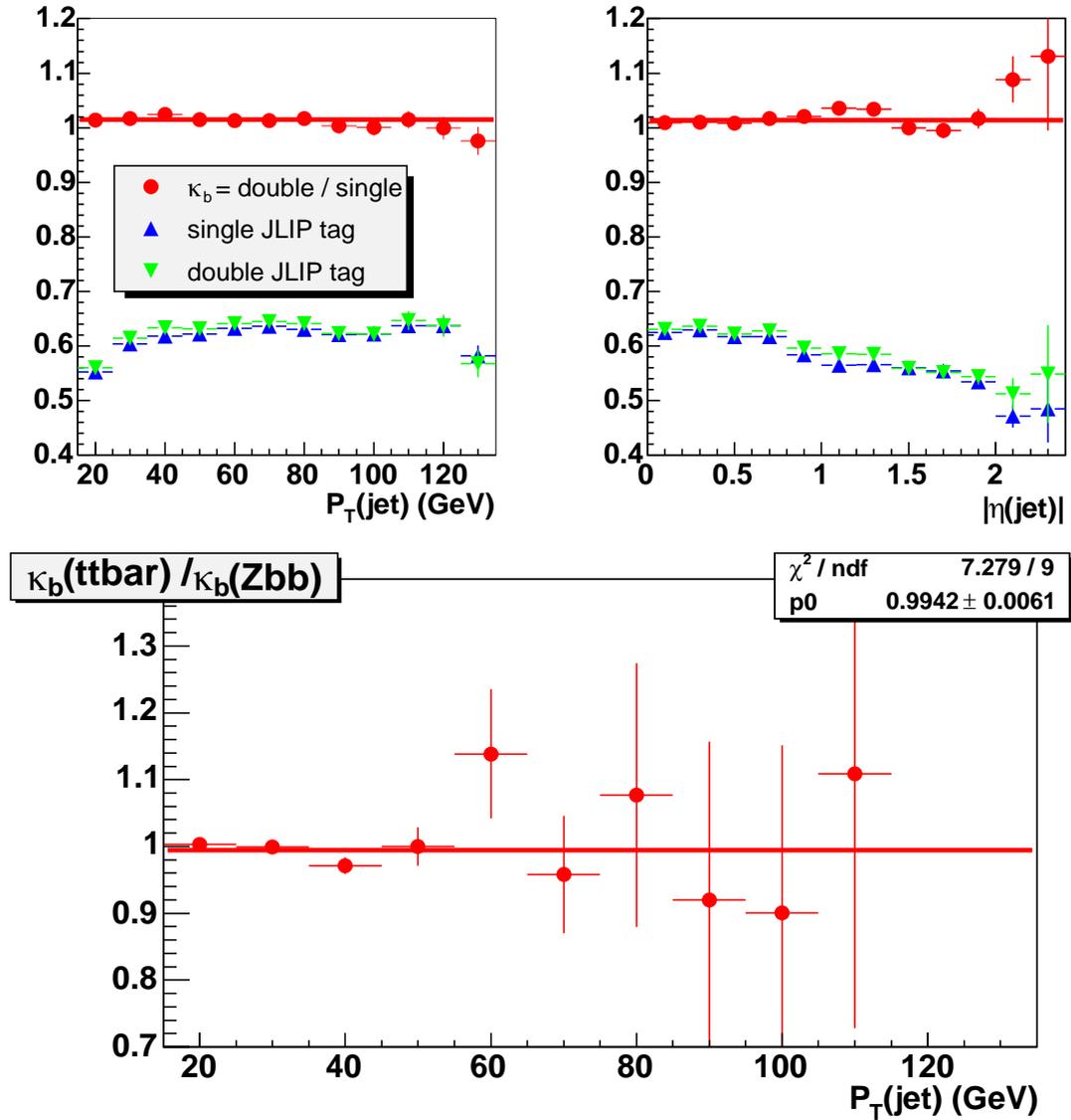
**double/single JLIP tag correlation in  $b \rightarrow \mu$  MC (JLIP < 0.5%)**


FIG. 3.14: En haut : dépendance en fonction de l'impulsion transverse (à gauche) et de la pseudorapidité (à droite) du jet muonique du facteur de corrélation  $\kappa_b^{(B,J)}$ . Les triangles bleus représentent l'efficacité de  $J$ , les triangles verts l'efficacité combinée ( $B, J$ ) et les disques rouges le rapport des deux graphes précédents ( $\kappa_b^{(B,J)}$ ). Aucune dépendance significative n'est observée. En bas : rapport entre les valeurs de  $\kappa_b^{(B,J)}$  entre des lots de données simulées  $t\bar{t}$  et  $Z \rightarrow b\bar{b}$ . Ce rapport ne dépend pas du processus physique.

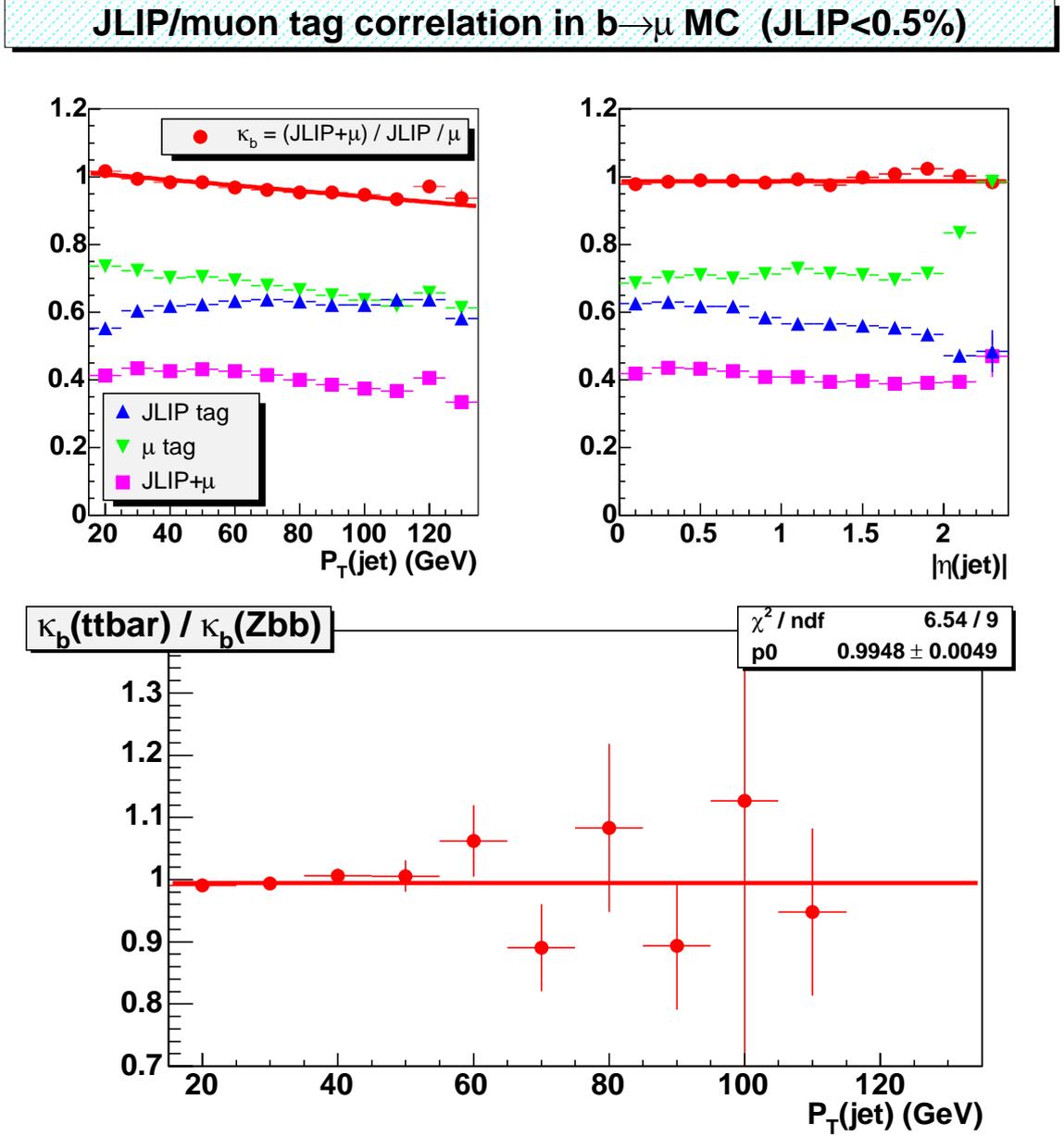


FIG. 3.15: En haut : dépendance en fonction de l'impulsion transverse (à gauche) et de la pseudorapidité (à droite) du jet muonique du facteur de corrélation  $\kappa_b^{(J,M)}$ . Les triangles bleus représentent l'efficacité de  $J$ , les triangles verts l'efficacité de  $M$  et les carrés roses l'efficacité combinée  $(J, M)$ . Les disques rouges représentent le facteur correctif  $\kappa_b^{(J,M)}$ . La dépendance en fonction de  $p_T$  est paramétrisée par une droite.

En bas : rapport entre les valeurs de  $\kappa_b^{(J,M)}$  entre des lots de données simulées  $t\bar{t}$  et  $Z \rightarrow b\bar{b}$ . Ce rapport ne dépend pas du processus physique.

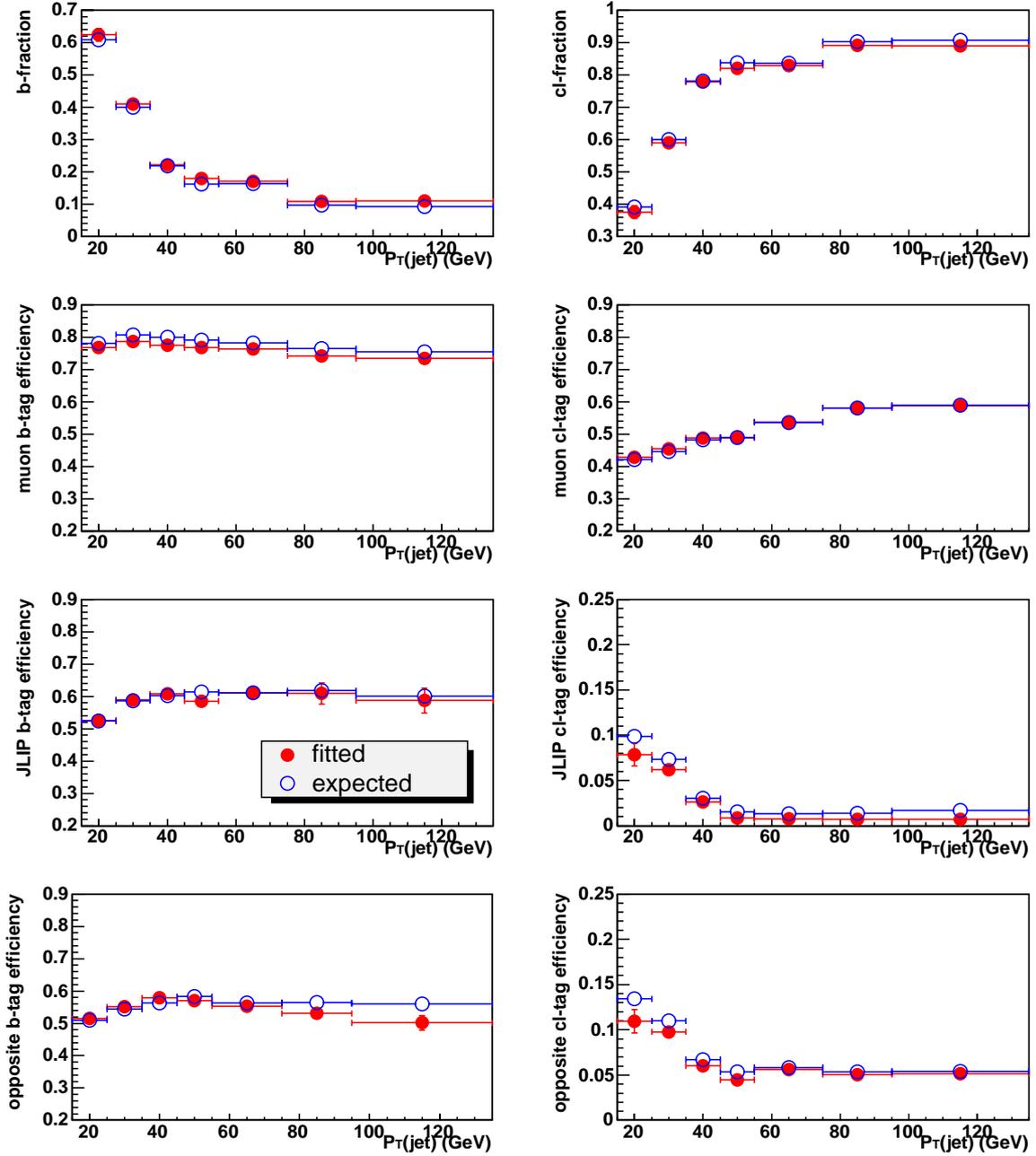


FIG. 3.16: Comparaison entre le résultat de *System8* (disques rouges) et les valeurs attendues (cercles bleu) dans un mélange d'événements simulés pour *JLIP Medium*. De gauche à droite et de haut en bas :  $f_b$ ,  $f_{cl}$ ,  $\varepsilon_b^{(M)}$ ,  $\varepsilon_{cl}^{(M)}$ ,  $\varepsilon_b^{(J)}$ ,  $\varepsilon_{cl}^{(J)}$ ,  $\varepsilon_b^{(B)}$ ,  $\varepsilon_{cl}^{(B)}$ . L'efficacité de *JLIP* déterminée par *System8* est en bon accord avec celle attendue.

Source d'erreur	incertitude relative sur $\varepsilon_b^{(J)}$ (%)
$\kappa_b^{(J,M)}$	$\pm 0.9$
$\kappa_b^{(B,J)}$	$\pm 0.4$
$\kappa_{cl}^{(J,M)}$	$< 0.01$
$\kappa_{cl}^{(B,J)}$	$\pm 0.3$
$p_{Trel} > 0.5 - 0.9$	$\pm 0.9$
Total	$\pm 1.4$
Statistique	de $\pm 1$ à $\pm 9$

TAB. 3.3: Erreurs systématiques relatives sur l'efficacité d'identification des jets  $b$  extraites avec *System8*. L'erreur statistique dépend du point de fonctionnement et des intervalles en  $p_T$  ou en  $|\eta|$ .

L'incertitude statistique sur les fractions  $Q^{(x)}$  est estimée en séparant les 8 lots de jets  $N^{(x)}$  passant chaque critère de sélection en 8 lots statistiquement indépendants :

$$Q^{(x)} = N^{(x)} / N^{(0)}, \text{ où } N^{(0)} \text{ est le nombre total de jets,}$$

$$\begin{aligned}
w_1 &= N^{(J,M,B)} \\
w_2 &= N^{(B,J)} - w_1 \\
w_3 &= N^{(B,M)} - w_1 \\
w_4 &= N^{(B)} - w_3 - w_2 - w_1 \\
w_5 &= N^{(J,M)} - w_1 \\
w_6 &= N^{(J)} - N^{(J,M)} - w_2 \\
w_7 &= N^{(M)} - N^{(J,M)} - w_3 \\
w_8 &= N^{(0)} - w_7 - w_6 - w_5 - w_4 - w_3 - w_2 - w_1
\end{aligned} \tag{3.22}$$

Chaque  $w_i$  est modifié séparément de  $\pm\sqrt{w_i}$ . Le système est résolu pour chaque variation et l'écart par rapport à la solution centrale du système est calculé. Les écarts positifs (négatifs) sont sommés quadratiquement pour obtenir l'erreur systématique supérieure (inférieure) due à la statistique du lot BID.

Les erreurs sur les facteurs correctifs prenant en compte les corrélations sont estimées en résolvant le système pour chaque variation de  $\pm 1$  écart standard de chaque facteur. L'erreur liée au regroupement des jets  $c$  et légers est estimée en faisant varier  $\kappa_{cl}^{(B,J)}$  de 0.2 à 1.8. Pour des variations plus grandes le système est instable et ne possède plus de solution. Une erreur sur la stabilité du système en fonction de la valeur de la coupure sur le  $p_{Trel}$  du muon est calculée en faisant varier cette coupure de 0.5 à 0.9 GeV/c. L'approximation d'une même efficacité de la coupure en  $p_{Trel}$  pour des jets  $c$  et légers est valide dans cette gamme d'impulsion.

Le tableau 3.3 regroupe les incertitudes sur l'efficacité d'identification des jets  $b$  liées à chacune de ces sources. L'erreur dominante provient de la statistique du lot **BID**. Plus la coupure sur  $\mathcal{P}_{jet}^+$  est sévère, plus le nombre de jets identifiés diminue et plus l'erreur statistique augmente. Cette incertitude varie suivant l'intervalle d'impulsion transverse ou de pseudorapidité considéré. Elle est particulièrement élevée dans les intervalles peu peuplés : à grande impulsion transverse ( $p_T > 65$  GeV/c) et à grande pseudorapidité ( $2 < |\eta| < 2.5$ ).

## 3.5.2 Efficacité d'identification des saveurs lourdes

### 3.5.2.1 Efficacité d'identification dans la simulation

La mesure de l'efficacité d'identification dans la simulation est aisée puisque la saveur de chaque jet est connue. Cette efficacité est simplement le rapport du nombre de jets  $b$  ( $c$ ) étiquetés au nombre des jets  $b$  ( $c$ ) *taggables*. Ces efficacités pour *JLIP Medium* sont données sur la figure 3.17. L'efficacité d'identification d'un jet  $b$  contenant un muon ( $\sim 0.61$ ) est supérieure de 50% à celle déterminée dans les données avec *System8* ( $\sim 0.42$ ). Les rapports d'efficacité entre différents processus simulés sont illustrés sur la figure 3.18. Les variations entre processus physiques sont inférieures à  $\pm 1\%$  pour l'efficacité d'identification des jets  $b$  contenant un muon et à  $\pm 3\%$  pour des jets  $b$  et  $c$  inclusifs. Ces variations sont prises en compte dans le calcul des erreurs systématiques.

### 3.5.2.2 Efficacité d'identification dans les données réelles

Les résultats obtenus par la méthode *System8* sont utilisés pour déterminer l'efficacité d'identifier les jets de saveurs lourdes dans les données réelles. Les efficacités de *JLIP* calculées avec *System8* ( $\varepsilon_b^{(J)}$ ) ne sont valables que pour des jets contenant un muon. Pour obtenir l'efficacité inclusive d'identifier un jet  $b$ , le résultat de *System8* est corrigé à l'aide d'événements simulés. Les efficacités d'identification d'un jet  $b$  quelconque ( $\varepsilon_{b \rightarrow incl.}^{MC}$ ) et d'un jet  $b$  contenant un muon ( $\varepsilon_{b \rightarrow \mu}^{MC}$ ) sont déterminées dans les lots Monte-Carlo  $Z \rightarrow b\bar{b}$  et  $Z \rightarrow b\bar{b} \rightarrow \mu X$ . Ces efficacités sont calculées dans les mêmes intervalles d'impulsion transverse et de pseudorapidité que pour *System8*. L'efficacité de *JLIP* pour les jets  $b$  devient finalement :

$$\varepsilon_b^{JLIP} = \frac{\varepsilon_b^{(J)} \cdot \varepsilon_{b \rightarrow incl.}^{MC}}{\varepsilon_{b \rightarrow \mu}^{MC}} = SF_b \cdot \varepsilon_{b \rightarrow incl.}^{MC} \quad (3.23)$$

$SF_b$  est le rapport d'efficacité d'identification des jets  $b$  entre données et simulation. Le facteur  $SF_b$  et l'efficacité de *JLIP* en fonction de l'impulsion transverse et de la pseudorapidité du jet  $b$  sont donnés sur les figures 3.19 et 3.20 respectivement.

Les jets  $c$  sont aussi des jets de saveurs lourdes mais il n'est pas toujours possible de les séparer des jets légers en raison de caractéristiques voisines. Il n'existe pas actuellement de méthode pour déterminer l'efficacité des jets  $c$  directement dans les données. L'identification d'un jet  $c$  est, comme l'identification des jets  $b$ , liée au temps de vol

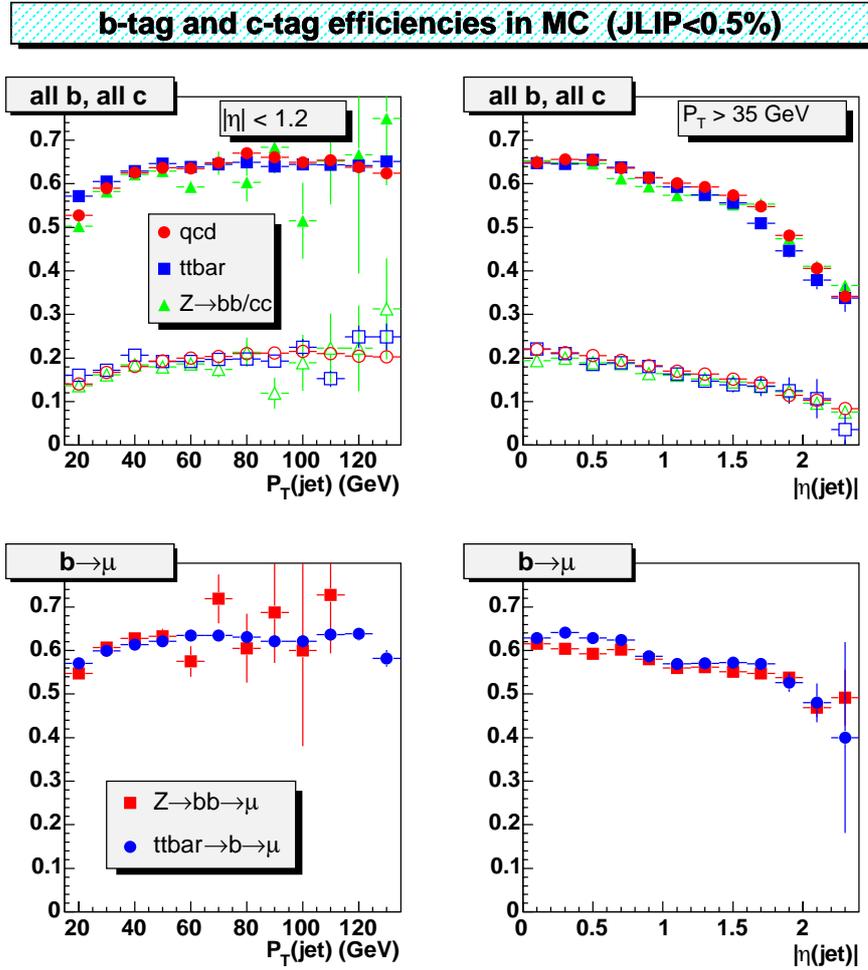


FIG. 3.17: Efficacités de *JLIP Medium* dans la simulation en fonction de l'impulsion transverse (à gauche) et de la pseudorapidité des jets (à droite).

En haut : Efficacité pour l'ensemble des jets  $b$  ( $c$ ) *taggables* en cercles pleins (vides) dans différents processus simulés : événements multijet (disques rouges),  $t\bar{t}$  (carrés bleus) et  $Z \rightarrow b\bar{b}/Z \rightarrow c\bar{c}$  (triangles verts).

En bas : Efficacité d'identification des jets  $b$  contenant un muon dans les processus  $Z \rightarrow b\bar{b}$  (carrés rouges) et  $t\bar{t}$  (disques bleus).

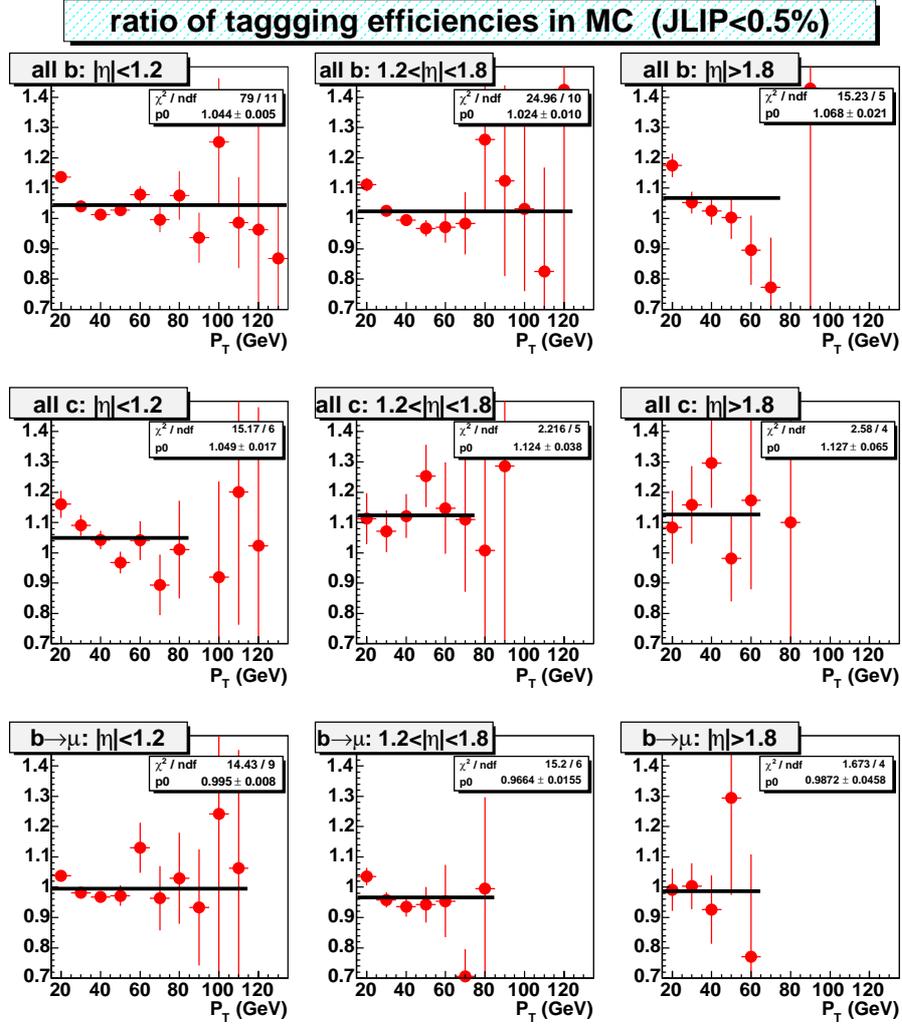


FIG. 3.18: Rapports d'efficacité d'identification des jets  $b$  et  $c$  entre différents processus simulés pour *JLIP* medium en fonction de  $p_T$ . En haut : jets  $b$  ( $t\bar{t}$ ,  $Z \rightarrow b\bar{b}$ ) ; au centre : jets  $c$  ( $Z \rightarrow c\bar{c}$ , dijets  $c\bar{c}$ ) ; en bas : jets  $b$  contenant un muon ( $t\bar{t} \rightarrow b\bar{b} \rightarrow \mu X$ ,  $Z \rightarrow b\bar{b} \rightarrow \mu X$ ).

Les rapports sont présentés dans trois intervalles de pseudorapidité ; de gauche à droite :  $|\eta| < 1.2$ ,  $1.2 < |\eta| < 1.8$ ,  $|\eta| > 1.8$ .

des hadrons  $c$  (contrairement à l'identification des jets légers qui est en grande partie due à des effets de résolution et de reconstruction des traces). Il donc est vraisemblable que le facteur correctif entre données et simulation est similaire pour les jets  $b$  et  $c$ , soit  $SF_c = SF_b$ . L'efficacité d'identification des jets  $c$  est ainsi estimée à partir de l'efficacité d'identification des jets  $b$ . L'efficacité d'identifier un jet  $c$  est alors :

$$\varepsilon_c^{JLIP} = \frac{\varepsilon_b^{(J)} \cdot \varepsilon_{c \rightarrow incl.}^{MC}}{\varepsilon_{b \rightarrow \mu}^{MC}} = SF_b \cdot \varepsilon_{c \rightarrow incl.}^{MC} \quad (3.24)$$

Cette efficacité est présentée sur la figure 3.21.

Le tableau 3.4 résume les erreurs systématiques sur chaque composante intervenant dans le calcul de ces efficacités ( $\varepsilon_b^{(J)}$ ,  $\varepsilon_{b \rightarrow incl.}^{MC}$ ,  $\varepsilon_{c \rightarrow incl.}^{MC}$ ,  $\varepsilon_{b \rightarrow \mu}^{MC}$ ), ainsi que les erreurs systématiques totales sur  $SF_b$ ,  $\varepsilon_b^{JLIP}$ ,  $\varepsilon_c^{JLIP}$ . Les erreurs systématiques sur les efficacités finales sont obtenues en sommant quadratiquement les erreurs relatives de chaque terme intervenant dans leur calcul. Les erreurs liées à la statistique des lots de données, y compris les erreurs statistiques de *System8* décrites précédemment, seront traitées au § 3.5.4.2.

	Contribue à			$\mathcal{P}_{jet}^+$					
	$SF_b$	$\varepsilon_b^{JLIP}$	$\varepsilon_c^{JLIP}$	< 0.1%	< 0.3%	< 0.5%	< 1.0%	< 2.0%	< 4.0%
$\varepsilon_b^{(J)}$	✓	✓	✓	1.4%	1.4%	1.4%	1.4%	1.4%	1.4%
$\varepsilon_{b \rightarrow incl.}^{MC}$		✓		2.8%	2.4%	2.1%	1.6%	1.3%	1.1%
$\varepsilon_{c \rightarrow incl.}^{MC}$			✓	2.8%	3.0%	3.3%	2.9%	2.9%	2.9%
$\varepsilon_{b \rightarrow \mu}^{MC}$	✓	✓	✓	1.0%	0.8%	0.7%	0.7%	0.7%	0.7%
$SF_b$				1.7%	1.6%	1.6%	1.6%	1.6%	1.6%
$\varepsilon_b^{JLIP}$				3.3%	2.9%	2.6%	2.2%	2.0%	1.9%
$\varepsilon_c^{JLIP}$				3.3%	3.4%	3.7%	3.3%	3.3%	3.3%

TAB. 3.4: Erreurs systématiques relatives partielles et totales sur le facteur correctif entre données et simulation ( $SF_b$ ) et les efficacités d'identification des jets  $b$  ( $\varepsilon_b^{JLIP}$ ) et des jets  $c$  ( $\varepsilon_c^{JLIP}$ ) pour chacun des 6 points de fonctionnement.

### 3.5.3 Mesure du taux de mauvaise identification

Pour caractériser pleinement un point de fonctionnement, il faut déterminer en plus de l'efficacité d'identification des jets  $b$ , le taux de mauvaise identification, c'est à dire l'efficacité de l'algorithme sur des jets légers. Dans un cas idéal, la distribution de la probabilité  $JLIP$  positive est uniforme pour les jets légers. L'identification d'un jet résultant d'une coupure supérieure sur cette probabilité  $\mathcal{P}_{jet}^+ < C$ , il en découle une efficacité d'identification pour les jets légers voisine de  $C$ . De plus, la présence de particules à grand temps de vol ( $V^0$ ) dans les jets légers, l'existence de traces mal reconstruites

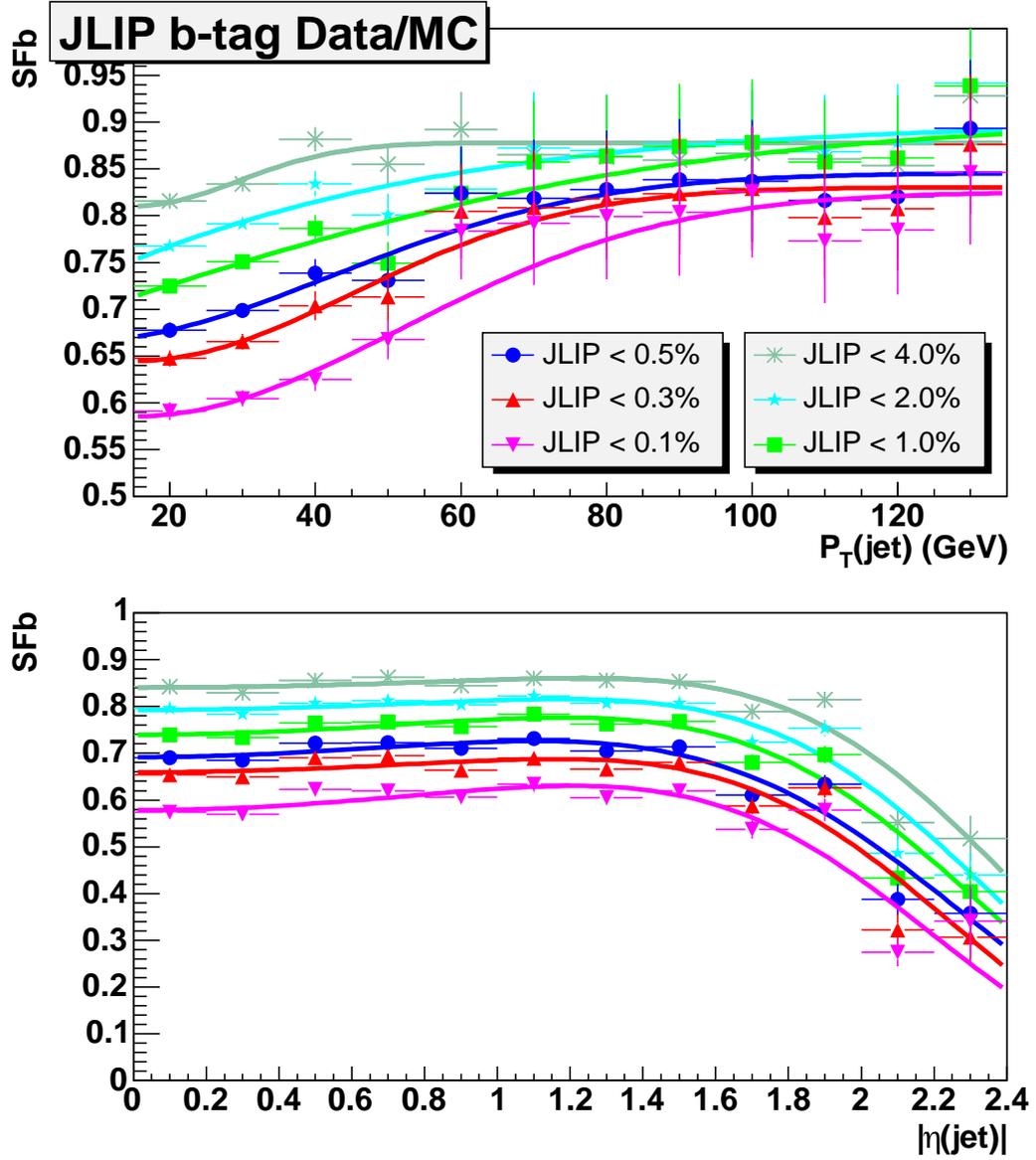


FIG. 3.19: Facteur correctif  $SF_b$  entre données et simulation pour des jets  $b$  en fonction de l'impulsion transverse (en haut) et de la pseudorapidité du jet pour les 6 points de fonctionnement.

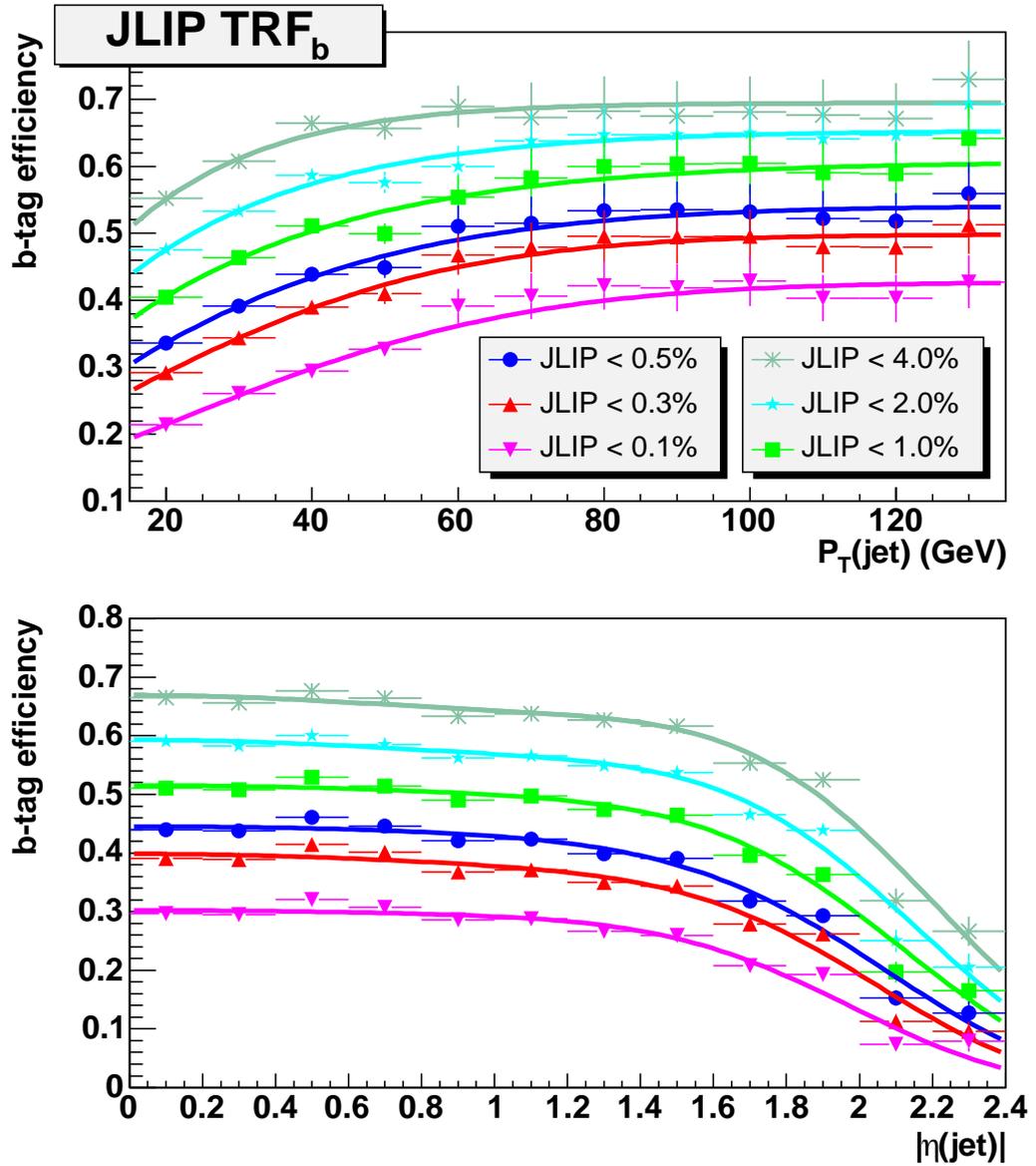


FIG. 3.20: Efficacité d'identification des jets  $b$  avec  $JLIP$   $\varepsilon_b^{JLIP}$  en fonction de l'impulsion transverse (en haut) et de la pseudorapidité du jet pour les 6 points de fonctionnement.

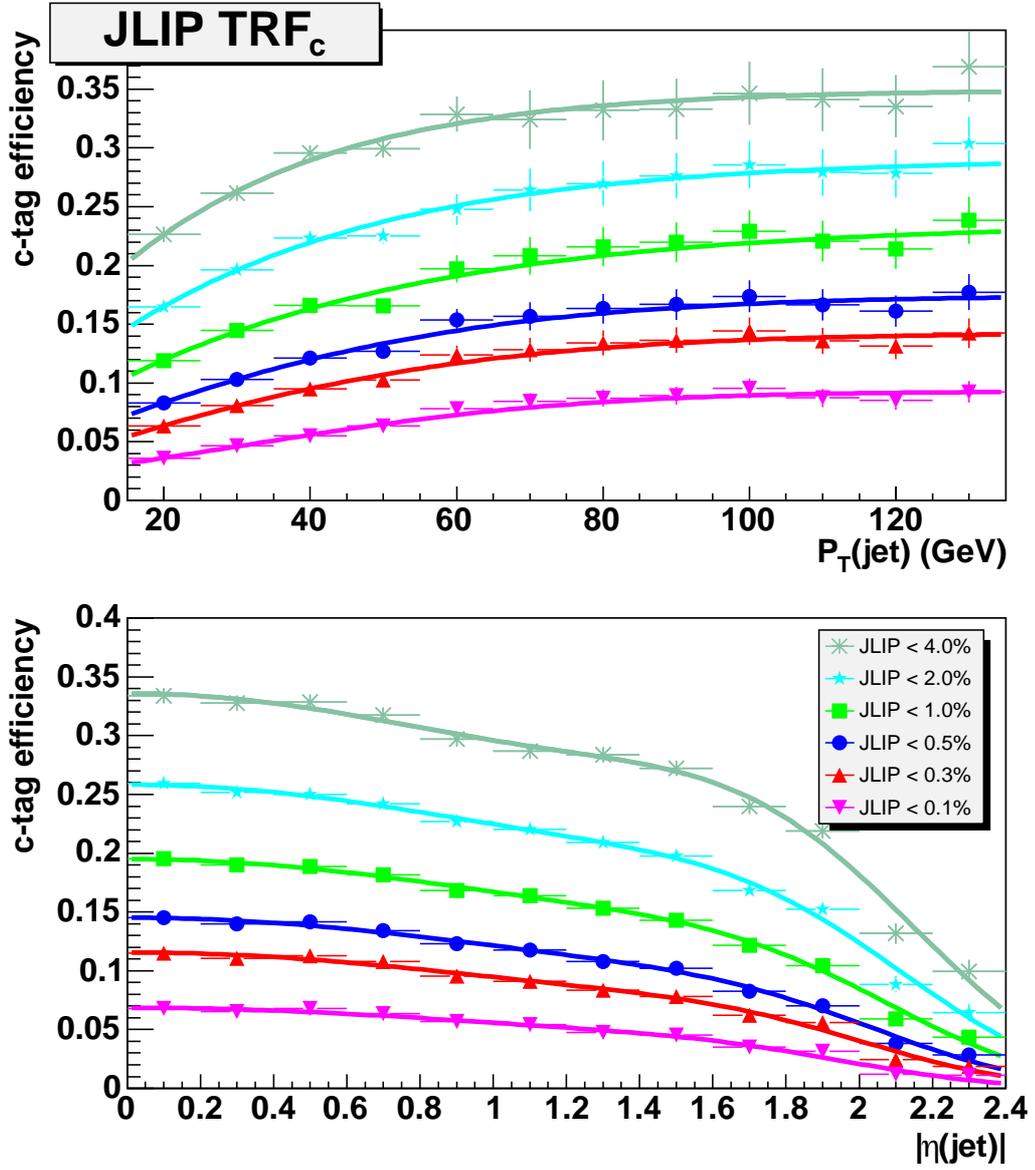


FIG. 3.21: Efficacité d'identification des jets  $c$  avec  $JLIP \varepsilon_c^{JLIP}$  en fonction de l'impulsion transverse (en haut) et de la pseudorapidité du jet pour les 6 points de fonctionnement.

et une petite contribution des saveurs lourdes aux données multijet utilisées pour dériver les fonctions de résolution augmentent ce taux de mauvaise identification. *System8* ne peut pas être utilisé pour mesurer ce taux, puisque les jets *c* et légers sont traités globalement. Une autre méthode décrite ci-après doit donc être utilisée.

### 3.5.3.1 Efficacité d'étiquetage négatif

La probabilité de temps de vol calculée avec les traces de signifiacnce négative  $\mathcal{P}_{jet}^-$  a été introduite au § 3.2.2. Pour des jets ne contenant que des traces issues du vertex primaire, la distribution en signifiacnce signée est symétrique, donc les distributions  $\mathcal{P}_{jet}^+$  et  $\mathcal{P}_{jet}^-$  sont identiques. Un jet est étiqueté négativement si  $\mathcal{P}_{jet}^- < C$ . L'efficacité d'étiquetage négatif et l'efficacité d'étiquetage positif sont identiques pour des traces provenant toutes du vertex primaire. Ainsi, l'efficacité d'étiquetage négatif dans les données multijet QCD et EM1TRK  $\varepsilon_{data}^-$  est utilisée pour estimer l'efficacité d'étiquetage positif des jets légers  $\varepsilon_l^{JLIP}$ . Cette efficacité  $\varepsilon_{data}^-$  est corrigée de deux effets :

- Les données multijet contiennent également des jets de saveurs lourdes qui peuvent produire des traces de grande signifiacnce négative mal signée. Ceci affecte la probabilité d'étiquetage négatif, qui n'est plus parfaitement uniforme, mais pique légèrement à 0 comme l'illustre la figure 3.11.
- Il reste des  $V^0$  dans les jets malgré la procédure de réjection décrite au § 3.4.3.

Ces deux effets sont estimés dans les événements multijet simulés. Pour chacun, un facteur correctif est dérivé :

- Le facteur  $SF_{hf} = \frac{\varepsilon_{MC,légers}^-}{\varepsilon_{MC}^-}$  est le rapport des efficacités d'étiqueter négativement un jet quelconque ( $\varepsilon_{MC}^-$ ) et un jet léger ( $\varepsilon_{MC,léger}^-$ ) dans la simulation.  $SF_{hf}$  corrige le premier effet lié à la présence de saveurs lourdes dans les données multijet.
- Le facteur  $SF_{ll} = \frac{\varepsilon_{MC,légers}^+}{\varepsilon_{MC,légers}^-}$  est le rapport des efficacités d'étiquetage positif ( $\varepsilon_{MC,légers}^+$ ) et négatif ( $\varepsilon_{MC,légers}^-$ ) d'un jet léger dans la simulation. Il corrige le second effet résultant de la différence entre étiquetage positif et négatif pour les jets légers.

Finalement, le taux de mauvaise identification est déterminé par :

$$\varepsilon_l^{JLIP} = SF_{hf} \cdot SF_{ll} \cdot \varepsilon_{data}^- \quad (3.25)$$

Le taux de mauvaise identification et les facteurs correctifs pour *JLIP Medium* sont présentés sur la figure 3.22

### 3.5.3.2 Différences entre les lots QCD et EM1TRK

Les deux lots de données QCD et EM1TRK sont utilisés pour dériver l'efficacité d'étiquetage négatif. La figure 3.23 montre une différence d'efficacité entre les deux lots allant de 12 % (*UltraTight*) à 3 % (*ExtraLoose*) selon le point de fonctionnement. Cette différence est prise en compte dans la détermination de l'erreur systématique sur le taux de mauvaise identification. Elle résulte vraisemblablement d'une différence de

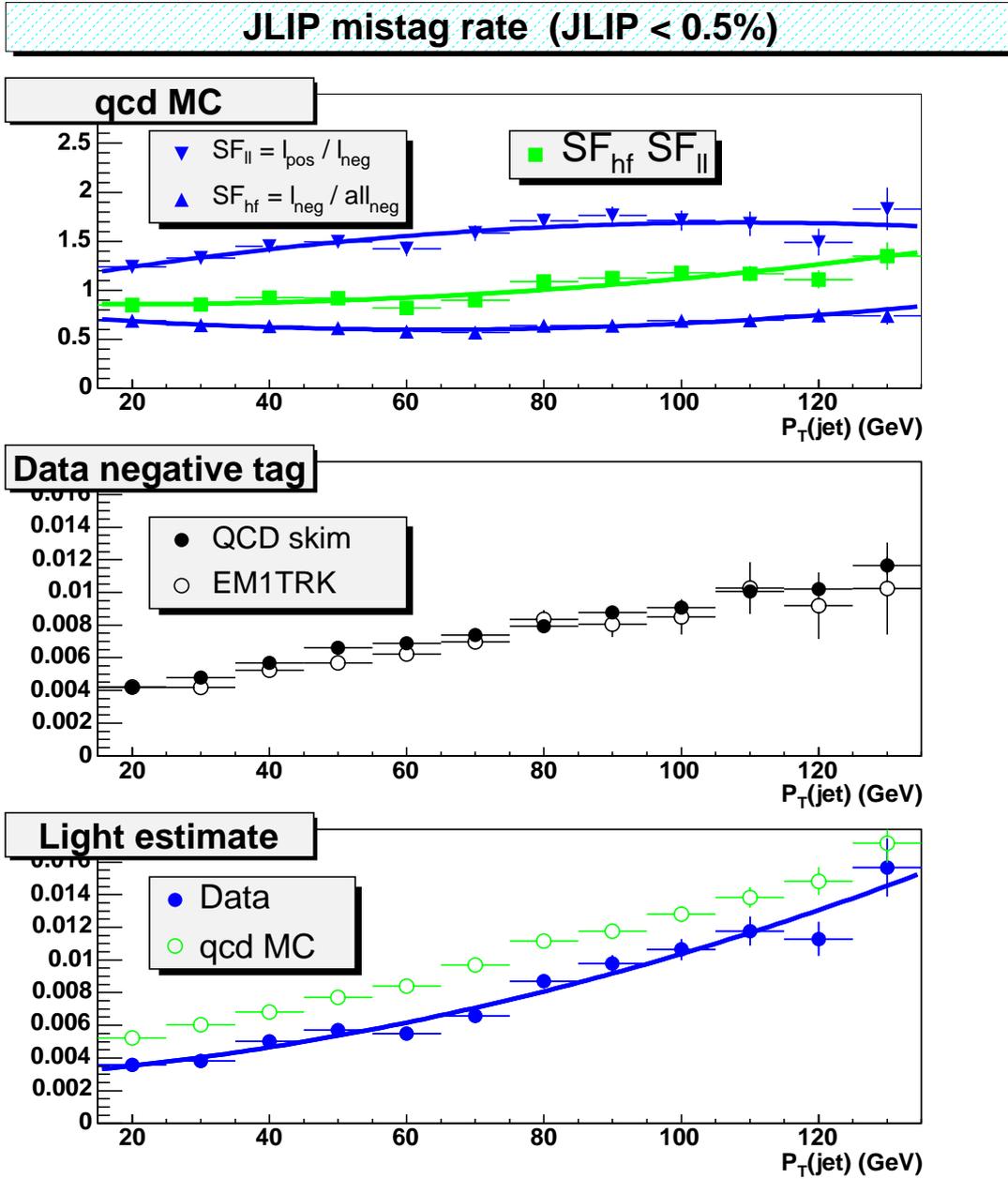


FIG. 3.22: En haut : facteurs correctifs  $SF_{hf}$  et  $SF_{||}$  et leur produit pour *JLIP Medium* en fonction de l'impulsion transverse du jet. Au centre : efficacité d'étiquetage négatif dans les données QCD (cercles) et EM1TRK (disques). En bas : taux de mauvaise identification pour *JLIP Medium* dans les données (disques bleus) et la simulation (cercles verts).

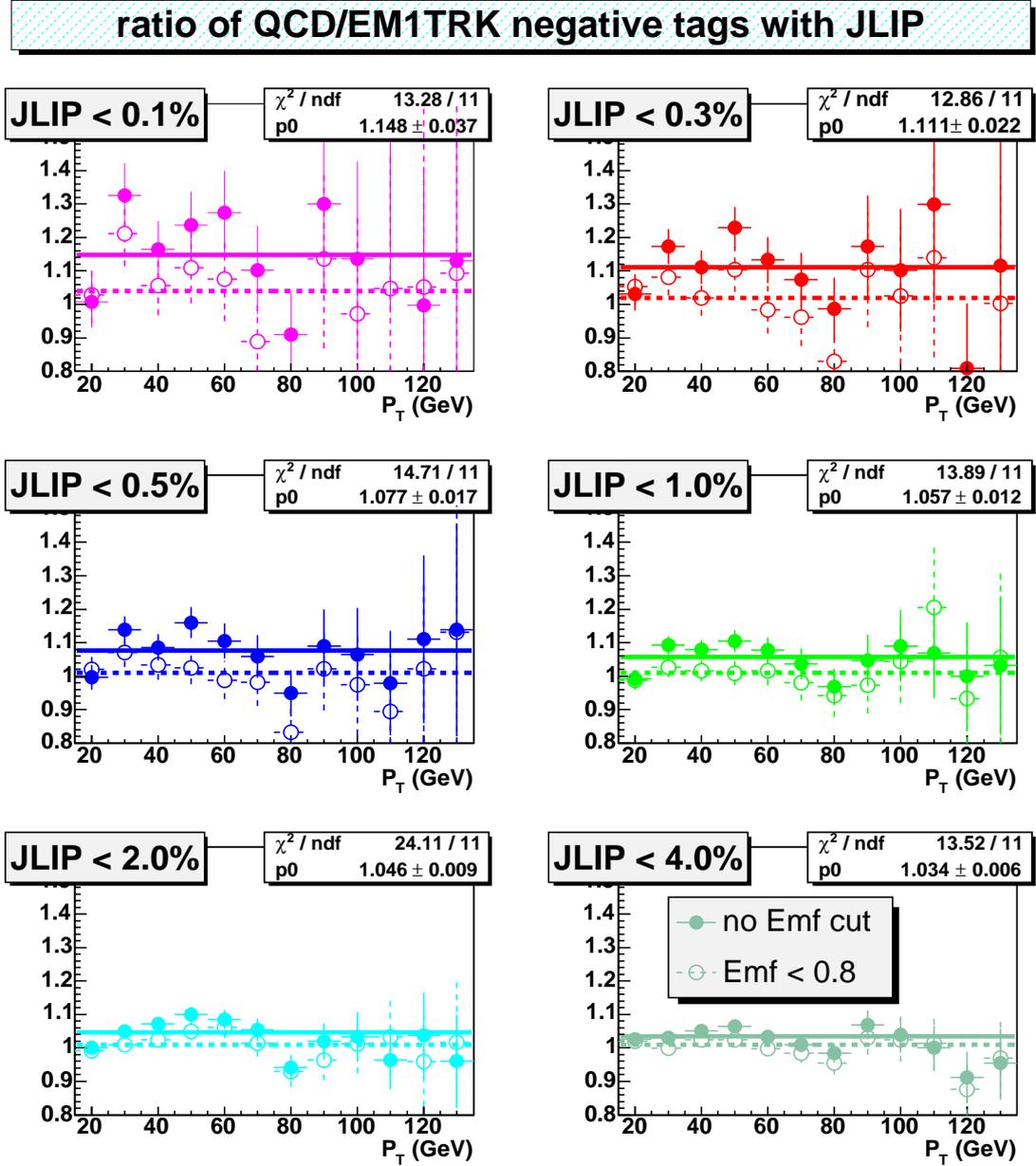


FIG. 3.23: Rapport des efficacités d'étiquetage négatif entre QCD et EM1TRK en fonction de l'impulsion transverse pour les 6 points de fonctionnement de *JLIP*. Les disques et les droites en traits pleins correspondent à la totalité des jets. Les cercles et les lignes pointillées correspondent aux jets avec  $Emf < 0.8$ .

topologie entre les deux lots d'événements illustrée par la figure 3.24. Deux hypothèses ont été avancées pour expliquer les différences d'efficacité :

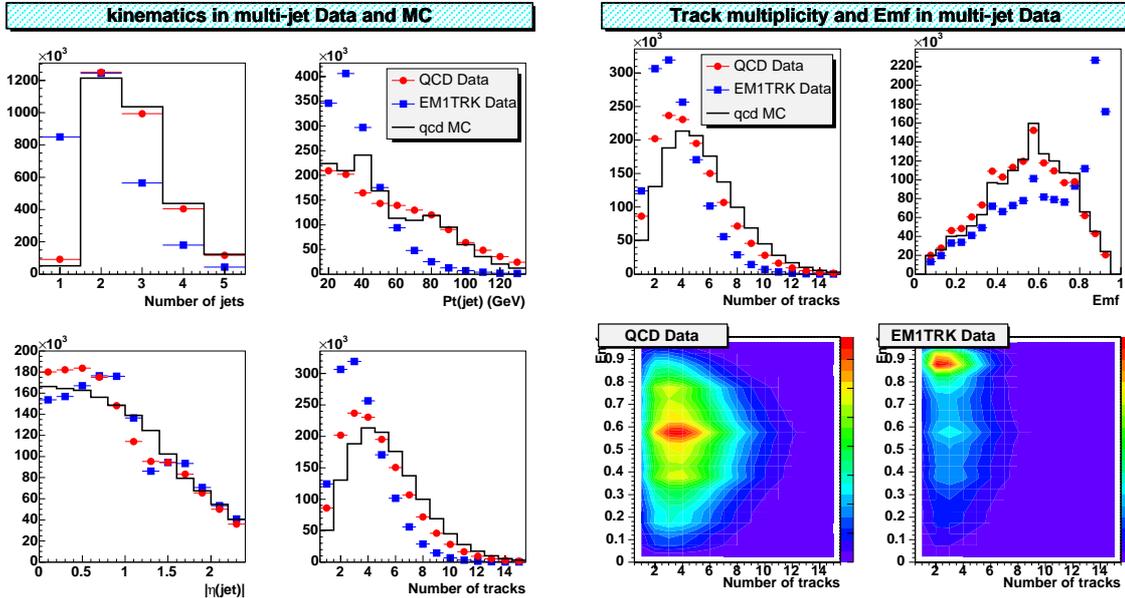


FIG. 3.24: Comparaisons de distributions entre les données QCD et EM1TRK et les événements multijet simulés.

A gauche : la multiplicité de jets dans l'événement, l'impulsion transverse, la pseudorapidité et la multiplicité de traces des jets.

En haut à droite : la multiplicité de traces et la fraction électromagnétique.

En bas à droite : corrélation entre la multiplicité de traces et la fraction électromagnétique des jets dans les données QCD et EM1TRK.

Tous les histogrammes sont normalisés au même nombre d'entrées.

- La pente du spectre en impulsion transverse est très différente. Un même  $p_T$  reconstruit peut correspondre à des  $p_T$  au niveau partonique différents du fait de la résolution du détecteur. Cela peut biaiser la comparaison entre QCD et EM1TRK. Il a été démontré pour l'algorithme *SVT* [75], qu'en ne prenant que les événements du lot QCD passant le critère de déclenchement JT25 (un jet de 25 GeV/c au niveau 3), les efficacités d'étiquetage négatif sont en bon accord. Cette étude n'a pas été refaite avec *JLIP*. Si cette explication est correcte, elle implique que les paramétrisations en fonction du  $p_T$  du taux de mauvaise identification peuvent dépendre du processus physique.
- Le lot EM1TRK demande un objet électromagnétique. La distribution en fraction électromagnétique ( $Emf$ ) des jets présente un pic pour les grandes valeurs de  $Emf$ . Ces jets sont principalement soit des électrons reconstruits comme des jets,

soit des jets contenant plus de  $\pi^0$  que des jets ordinaires et donc de fait moins de traces chargées. Un bon accord est obtenu en appliquant une coupure  $Emf < 0.8$  dans les deux lots de données. Ceci est illustré par les lignes pointillées de la figure 3.23.

La question de la validité de l'une ou l'autre des ces hypothèses reste en suspens. Le compromis retenu pour la mesure du taux de mauvaise identification consiste à utiliser la moyenne des efficacités négatives entre les deux lots. La différence entre les deux valeurs constitue l'erreur systématique, soit :

$$\varepsilon_{data}^- = \frac{\varepsilon_{EM1TRK}^- + \varepsilon_{QCD}^-}{2} \pm \frac{|\varepsilon_{EM1TRK}^- - \varepsilon_{QCD}^-|}{2} \quad (3.26)$$

Le taux de mauvaise identification final, pour chacun des 6 points de fonctionnement de *JLIP* est donné sur la figure 3.25

### 3.5.3.3 Estimation des erreurs systématiques

A l'erreur provenant des différences d'efficacité entre les lots QCD et EM1TRK, s'ajoute en quadrature une erreur liée à la différence du contenu en saveurs lourdes entre simulation et données réelles. Les fractions en jets  $c$  (4.6 %) et en jets  $b$  (2.6 %) dans le lot multijet simulés sont variées séparément de  $\pm 20$  %. Les écarts observés sont sommés linéairement pour estimer cette erreur. Le tableau 3.5 résume les erreurs systématiques pour chaque point de fonctionnement.

$\mathcal{P}_{jet}^+$	< 0.001	< 0.003	< 0.005	< 0.01	< 0.02	< 0.04
QCD/EM1TRK	$\pm 5.9\%$	$\pm 5.1\%$	$\pm 3.6\%$	$\pm 2.6\%$	$\pm 2.2\%$	$\pm 1.7\%$
fraction de jets $c$	$\pm 2.3\%$	$\pm 2.2\%$	$\pm 2.1\%$	$\pm 1.9\%$	$\pm 1.5\%$	$\pm 1.2\%$
fraction de jets $b$	$\pm 7.5\%$	$\pm 5.5\%$	$\pm 4.6\%$	$\pm 3.5\%$	$\pm 2.5\%$	$\pm 1.7\%$
Total	$\pm 11.4\%$	$\pm 9.2\%$	$\pm 7.6\%$	$\pm 6.0\%$	$\pm 4.6\%$	$\pm 3.4\%$

TAB. 3.5: Erreurs systématiques relatives sur le taux de mauvaise identification.

L'erreur liée à la statistique des lots Monte-Carlo utilisés pour dériver les facteurs correctifs est prise en compte lors du calcul de l'erreur finale sur les paramétrisations des efficacités au § 3.5.4.2.

## 3.5.4 Paramétrisation des efficacités

### 3.5.4.1 Factorisation des efficacités

Pour pouvoir pondérer les événements simulés, il faut disposer de paramétrisations (*TRF*) de toutes ces efficacités. Ces *TRF* sont déterminées en fonction de l'impulsion

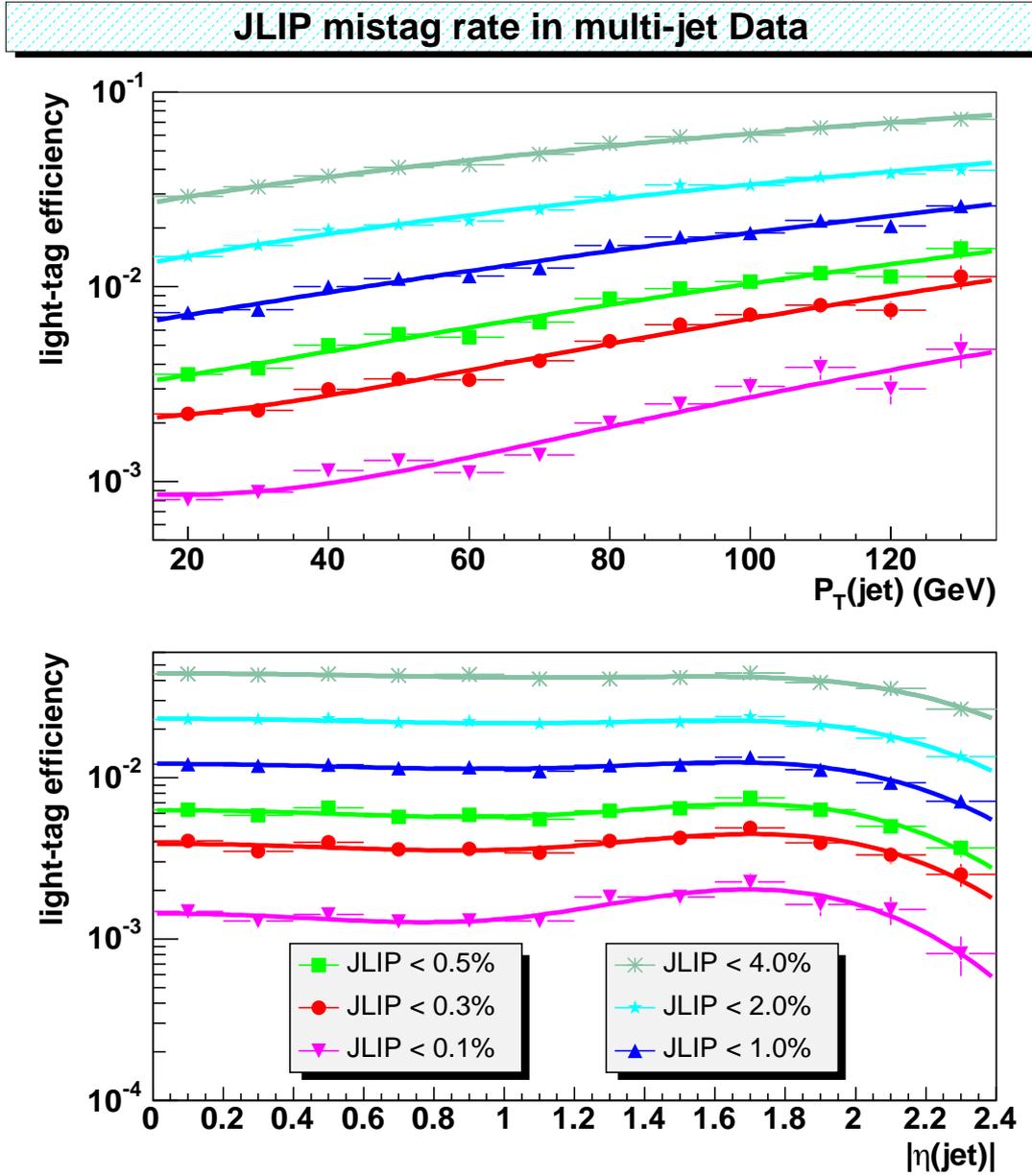


FIG. 3.25: Taux de mauvaise identification de *JLIP*  $\varepsilon_i^{JLIP}$  en fonction de l'impulsion transverse (en haut) et de la pseudorapidity du jet pour les 6 points de fonctionnement.

transverse et de la pseudorapidité du jet. En supposant ces deux paramètres peu corrélés, les efficacités peuvent être factorisées en deux termes indépendants :

$$TRF_{b,c,l}(p_T, |\eta|) = \varepsilon_{b,c,l}^{p_T}(p_T) \cdot \varepsilon_{b,c,l}^{|\eta|}(|\eta|) \quad (3.27)$$

Chacun de ces termes correspond aux efficacités déterminées séparément en fonction de chacune des variables présentées sur les figures 3.20 (jets  $b$ ), 3.21 (jets  $c$ ) et 3.25 (jets légers). Les paramétrisations à une dimension sont obtenues en ajustant les distributions d'efficacité à l'aide de fonctions reproduisant au mieux les formes observées. Il n'y a pas de justification physique au choix des formes fonctionnelles à l'exception de leur comportement asymptotique : les paramétrisations en fonction de  $p_T$  doivent tendre vers un plateau à grand  $p_T$  et celles en fonction de  $|\eta|$  doivent tendre vers 0 à grand  $|\eta|$ . Cette sécurité permet d'extrapoler les  $TRF$  en dehors des domaines étudiés, notamment en  $p_T$ . Les fonctions retenues sont :

$$\varepsilon_{b,c,l}^{p_T}(p_T) = A_0 \cdot (1 + A_1 p_T^2 + A_2 p_T^4) \cdot e^{-\frac{p_T^4}{A_3}}, \quad (3.28)$$

$$\varepsilon_{b,c,l}^{|\eta|}(|\eta|) = B_0 \cdot (1 - B_1 e^{-\frac{|\eta|^2}{B_2}}), \quad (3.29)$$

où  $A_{i=0..3}$  et  $B_{i=0..2}$  sont les paramètres ajustés. La méthode de construction de ces  $TRF$  est testée dans la simulation et dans les données. Les efficacités d'étiquetage des jets  $b$  sont déterminées dans les lots simulés. Une  $TRF$  est dérivée pour chaque efficacité et deux classes d'histogrammes sont réalisées sur les mêmes événements utilisés pour dériver la  $TRF$  :

- Les distributions en  $p_T$  et en  $|\eta|$  des jets  $b$  étiquetés.
- Les distributions en  $p_T$  et en  $|\eta|$  de tous les jets  $b$  *taggables* pondérés de la valeur de la  $TRF$ .

La même procédure est appliquée pour l'efficacité d'identification des jets  $c$  simulés et l'efficacité d'étiquetage négatif dans les données multijet. Si les paramétrisations sont correctes, les deux histogrammes doivent être similaires. Ceci permet de tester la normalisation des  $TRF$ , la qualité des ajustements et l'hypothèse de factorisation des  $TRF$ . La figure 3.26 montre ces histogrammes de contrôle. L'accord observé est satisfaisant mais n'est pas parfait, en particulier pour les jets  $c$ . Dans ce cas particulier ce sont les ajustements qui ne sont pas optimaux, les fonctions d'ajustement choisies pour les efficacités finales n'étant pas nécessairement adéquates pour les efficacités simulées.

### 3.5.4.2 Erreurs systématiques

Aux erreurs systématiques déjà déterminées dans les tableaux 3.3 et 3.5 s'ajoute un terme lié à l'hypothèse de factorisation et un autre terme lié à la statistique des différents lots de données utilisés.

Pour étudier plus précisément l'hypothèse de factorisation, les histogrammes de contrôle

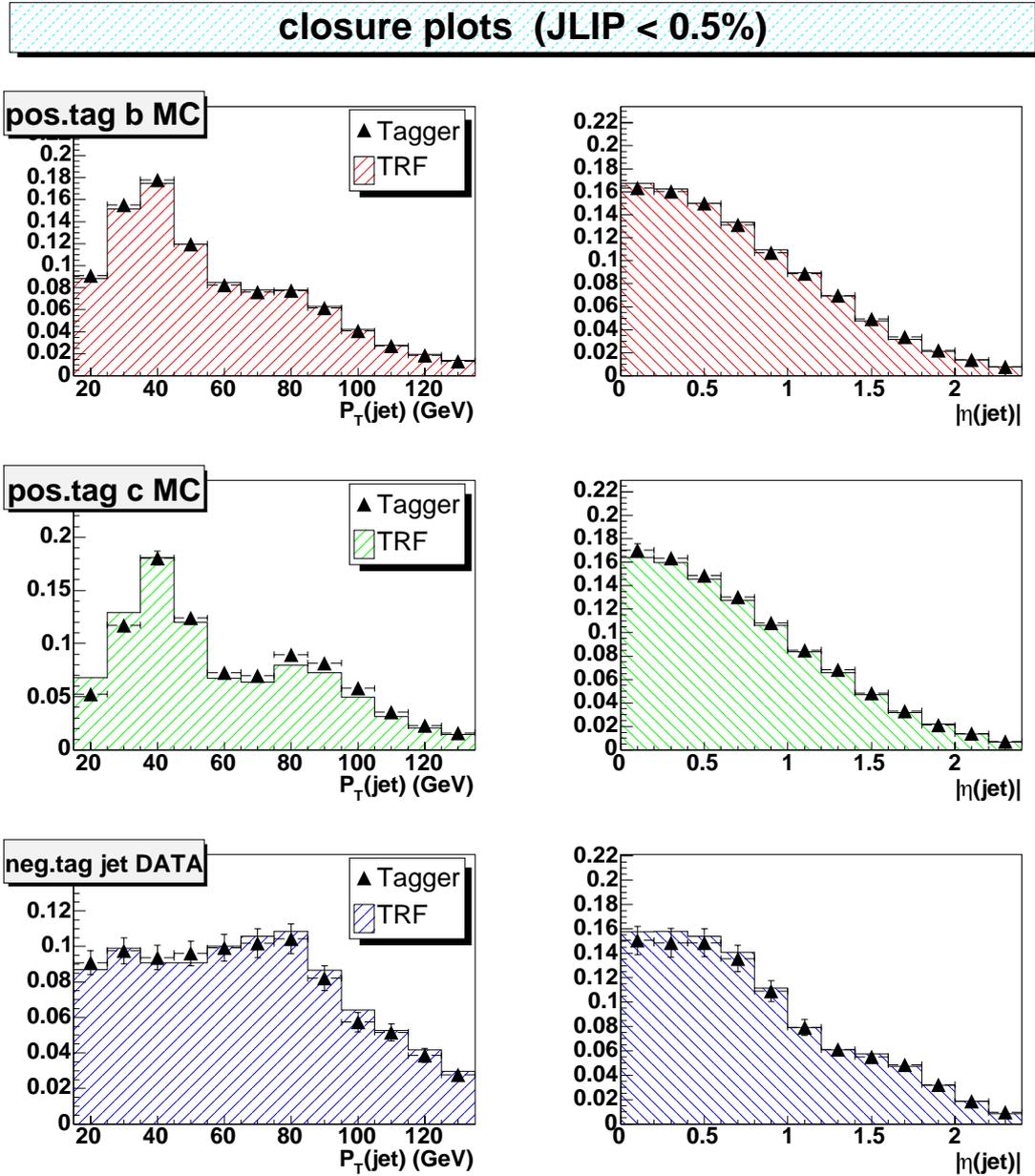


FIG. 3.26: Distributions en impulsion transverse (à gauche) et en pseudorapidité (à droite) obtenues par application de *JLIP Medium* (triangles noirs) et de la *TRF* correspondante sur tout les jets *taggables* (histogrammes hachurés) pour des jets  $b$  simulés (haut), des jets  $c$  simulés (centre) et des jets négativement étiquetés (bas).

de la figure 3.26 sont recréés en fonction du  $p_T$  du jet dans trois intervalles de pseudo-rapacité ( $|\eta| < 0.8$ ,  $0.8 < |\eta| < 1.4$  et  $|\eta| > 1.4$ ). Ces histogrammes et leur rapport sont présentés sur la figure 3.27. Ces rapports sont voisins de l'unité. La déviation par rapport à l'unité de la moyenne pondérée de ces rapports sert à introduire une nouvelle erreur systématique de "factorisation". L'erreur de factorisation des jets  $b$  et  $c$  est identique car les deux  $TRF$  sont dérivées de la même mesure d'efficacité d'étiquetage des jets  $b$ . Les erreurs de factorisation pour chaque point de fonctionnement et chaque  $TRF$  sont données dans le tableau 3.6.

$\mathcal{P}_{jet}^+$	< 0.001	< 0.003	< 0.005	< 0.01	< 0.02	< 0.04
$TRF_b$ et $TRF_c$	$\pm 0.5\%$	$\pm 0.4\%$				
$TRF_l$	$\pm 2.2\%$	$\pm 1.5\%$	$\pm 2.1\%$	$\pm 1.0\%$	$\pm 1.1\%$	$\pm 0.8\%$

TAB. 3.6: Erreurs systématiques relatives dues à l'hypothèse de factorisation.

Toutes les erreurs systématiques évoquées jusqu'ici ne dépendent pas des intervalles en  $p_T$  et  $\eta$  considérés. Ce n'est plus le cas pour les incertitudes d'origine statistique. Les ajustements sont effectués sur des histogrammes obtenus par le produit successif des distributions des différents termes décrits dans les paragraphes § 3.5.2 et § 3.5.3. Lors de chaque produit d'histogrammes, l'erreur statistique associée à chaque cellule de chaque histogramme est propagée. L'erreur statistique sur une  $TRF$  est alors estimée en déplaçant tous les points des distributions d'efficacité en  $p_T$  et en  $|\eta|$  de +1 écart standard. Les ajustements sont réeffectués pour ces deux nouvelles distributions en utilisant les mêmes formes fonctionnelles et une nouvelle  $TRF$  correspondant au décalage de +1 écart standard est ainsi construite. La même opération est répétée pour déterminer la  $TRF$  à -1 écart standard. Ces  $TRF$  servent à estimer l'incertitude statistique, qui est alors asymétrique. L'écart avec la valeur centrale donne l'erreur statistique en fonction de  $p_T$  et de  $|\eta|$ , à laquelle il faut ajouter en quadrature un terme constant, qui regroupe toutes les contributions déjà citées. Le tableau 3.7 résume tous ces termes constants pour chaque  $TRF$ . L'erreur totale sur chaque  $TRF$  en fonction de  $p_T$  et  $\eta$ , pour chaque point de fonctionnement, est donnée sur la figure 3.28.

$\mathcal{P}_{jet}^+$	< 0.001	< 0.003	< 0.005	< 0.01	< 0.02	< 0.04
$TRF_b$	$\pm 3.3\%$	$\pm 2.9\%$	$\pm 2.6\%$	$\pm 2.3\%$	$\pm 2.1\%$	$\pm 2.0\%$
$TRF_c$	$\pm 3.3\%$	$\pm 3.4\%$	$\pm 3.7\%$	$\pm 3.3\%$	$\pm 3.3\%$	$\pm 3.3\%$
$TRF_l$	$\pm 11.6\%$	$\pm 9.4\%$	$\pm 7.9\%$	$\pm 6.1\%$	$\pm 4.7\%$	$\pm 3.5\%$

TAB. 3.7: Erreurs systématiques relatives totales, indépendantes de l'impulsion transverse ou de la pseudorapacité sur les  $TRF$ .

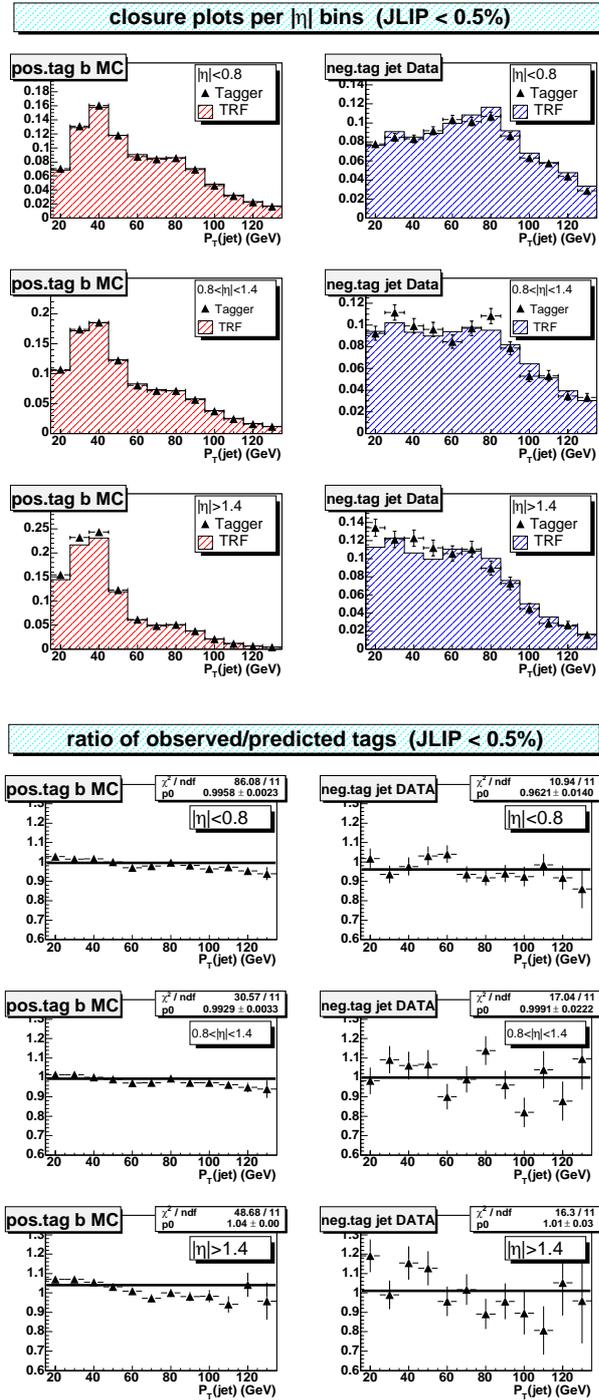


FIG. 3.27: En haut : histogrammes de contrôle des *TRF* pour les jets *b* simulés (à gauche) et les données négativement étiquetées (à droite) en fonction de l'impulsion transverse dans trois gammes de pseudorapidité.

En bas : rapports des histogrammes précédents. Les droites indiquent les valeurs moyennes.

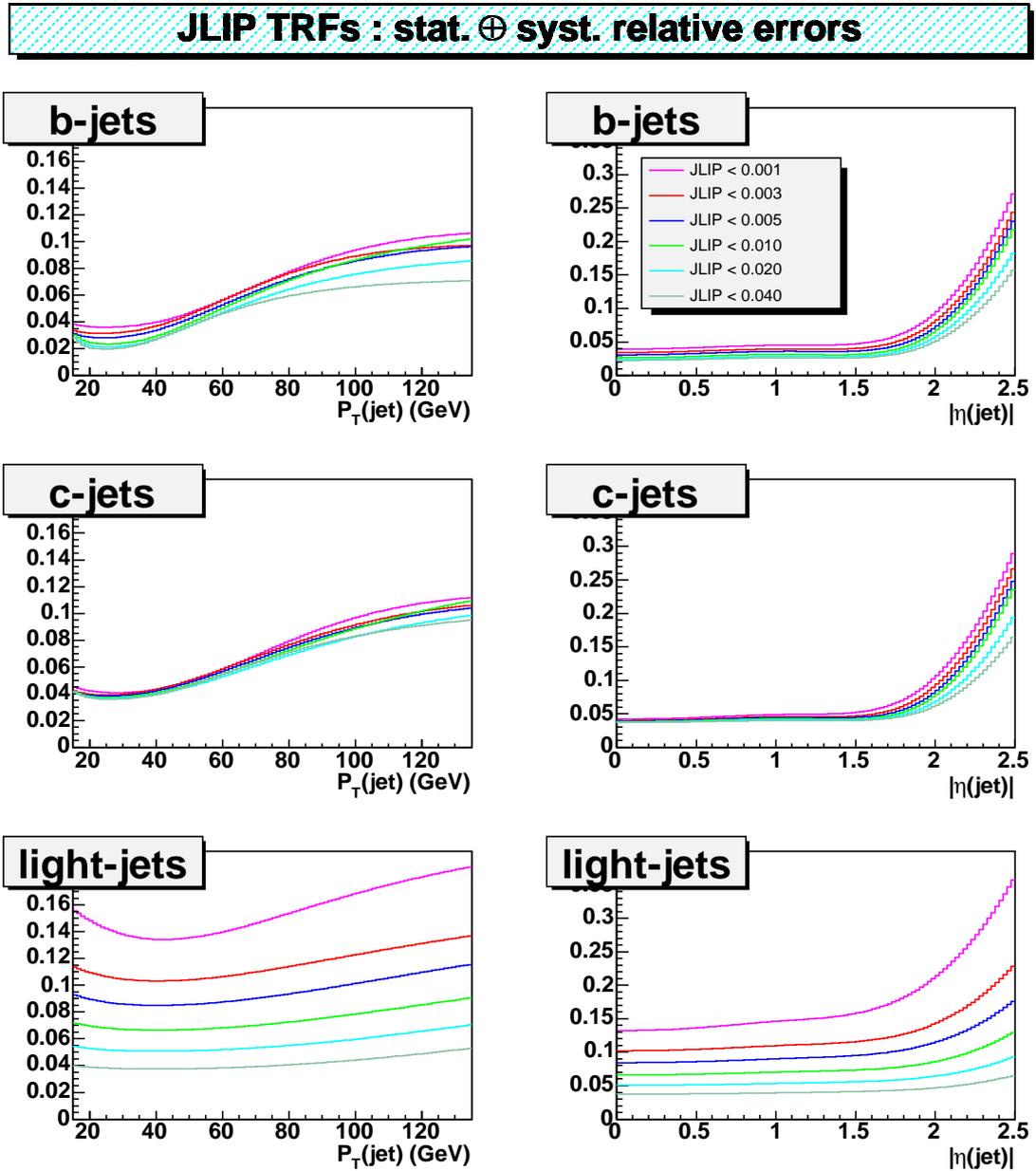


FIG. 3.28: Erreur systématique et statistique totale sur les  $TRF$  pour les jets  $b$  (haut),  $c$  (centre) et légers (bas) en fonction de  $p_T$  et  $|\eta|$  pour les 6 points de fonctionnement : *ExtraLoose* (gris), *SuperLoose* (cyan), *Loose* (vert), *Medium* (bleu), *Tight* (rouge) et *UltraTight* (rose).

## 3.6 Les autres algorithmes d'identification des jets $b$ à $D\emptyset$

Outre *JLIP*, il existe d'autres algorithmes d'identification des jets de quark  $b$  basés sur le temps de vol des hadrons beaux en usage à  $D\emptyset$  :

- **SVT** (*S*econdary *V*ertex *T*agger) [71] : Les jets de traces définis pour le critère de *taggabilité* sont utilisés pour rechercher des vertex, en utilisant un algorithme de Kalman similaire à celui du vertex primaire (cf. § 2.2.2.4). Dans un premier temps des vertex à deux traces sont créés puis des traces supplémentaires sont ajoutées afin de réduire le  $\chi^2$  du vertex. Finalement, un jet est étiqueté si un vertex secondaire associé au jet possède une grande signification en distance de vol (distance entre le vertex primaire et le vertex secondaire, divisée par son erreur). Trois points de fonctionnement sont définis pour *SVT* : *Tight*, *Medium*, *Loose*. Ces points de fonctionnement correspondent aux mêmes taux de mauvaise identification que les points de fonctionnement de *JLIP* correspondants. Les efficacités d'identification de *SVT* et *JLIP* sont similaires.
- **CSIP** (*C*ounted *S*igned *I*mpact *P*arameter) [76] : Cet algorithme basique n'est plus utilisé seul dans  $D\emptyset$ . Un jet est considéré comme étiqueté s'il contient deux traces de signification  $S_{IP} > 3$  ou trois traces de signification  $S_{IP} > 2$ .
- **NN** (*N*eural *N*etwork) [77] : Cet algorithme est nouveau dans  $D\emptyset$  et n'a pas encore été entièrement validé. Il consiste en la combinaison des informations fournies par *JLIP*, *SVT* et *CSIP* par un réseau de neurones. Il offre un gain d'efficacité d'environ 30% comparé à *JLIP* seul comme l'illustre la figure 3.29. Le réseau de neurones est entraîné sur des événements simulés. À terme ce réseau de neurones devrait être l'algorithme final d'identification des jets  $b$  de la collaboration  $D\emptyset$ .

La procédure de certification décrite au § 3.5 a été appliquée à chacun de ces algorithmes. Les mêmes méthodes sont utilisées pour déterminer les efficacités et les taux de mauvaise identification.

## 3.7 Conclusion

Un algorithme d'identification des jets  $b$ , *JLIP*, a été présenté de même que les détails de sa calibration pour l'expérience  $D\emptyset$ . Le processus de certification de l'algorithme, conduisant à la production des paramétrisations des efficacités d'identification des jets  $b$  et  $c$  et du taux de mauvaise identification a été décrit. Ces paramétrisations sont essentielles pour utiliser *JLIP* dans une analyse de physique, comme nous le montrerons au chapitre 4. Les performances de l'algorithme sont résumées sur la figure 3.30.

La complexité de cette certification provient principalement des grandes différences observées entre données et simulation concernant la résolution et la multiplicité des traces dans les jets. La nécessité d'estimer les efficacités dans les données réelles, nous a conduit

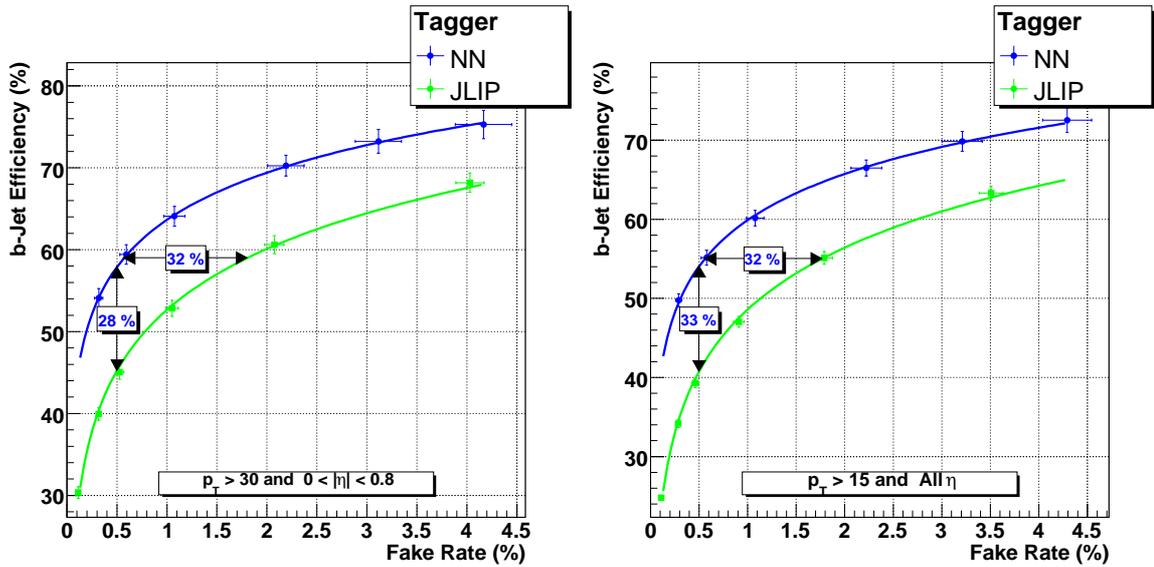


FIG. 3.29: Courbes de performance des algorithmes *JLIP* (vert) et *NN* (bleu). Les efficacités sont estimées dans le lot BID et le taux de mauvaise identification dans les lots QCD et EM1TRK. Deux domaines d'impulsion et de pseudorapidité sont représentés :  $p_T > 30$  GeV/c,  $|\eta| < 0.8$  (à gauche) et  $p_T > 15$  GeV/c,  $|\eta| < 2.5$  (à droite)

à développer des méthodes originales (*System8*, étiquetage négatif). Néanmoins, ces procédures sont sources d'erreurs systématiques supplémentaires, qui se répercutent dans les analyses de physique.

Il est important de remarquer que toutes les efficacités sont déterminées relativement aux jets *taggables*. L'utilisation des *TRF* dans une analyse nécessite également de paramétrer l'efficacité de *taggabilité* dans les données. Comme la *taggabilité* inclut les effets d'efficacité du détecteur, cette *TRF* peut dépendre du lot de données utilisé. Elle est donc spécifique à chaque analyse.

*JLIP* a été utilisé avec succès dans une dizaine d'analyses à  $D\bar{O}$ , notamment l'étude de la production de paires  $t\bar{t}$  [78] de la production par interaction faible du quark top (voir chapitre 4), la recherche du boson de Higgs et de la production associée  $Wb\bar{b}$  [79] et la recherche de particules supersymétriques [80, 81].

Les améliorations récentes du programme de simulation du détecteur, laissent envisager un meilleur accord entre données et simulation pour les prochaines certifications des algorithmes d'identification des jets *b*. Il serait alors possible d'identifier les jets *b* directement dans la simulation, et de corriger les déviations résiduelles par des facteurs correctifs. Les analyses s'en retrouveraient simplifiées et les incertitudes systématiques réduites.

Les algorithmes d'identification des jets *b* *JLIP* et *SVT* ont probablement atteint leur limite d'efficacité. Pour améliorer les performances d'identification, il faut désormais combiner les différents algorithmes. Ceci est réalisé au moyen d'un réseau de neurones.

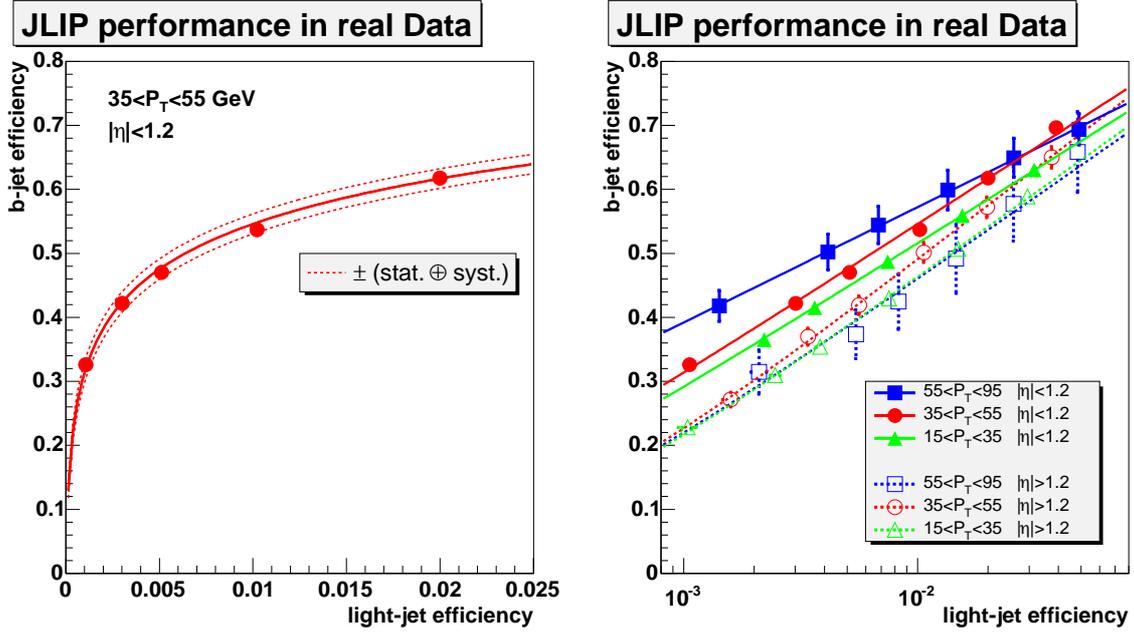


FIG. 3.30: Efficacité d'identification des jets  $b$  en fonction du taux de mauvaise identification. A gauche : Performances pour  $35 < p_T < 55$  GeV/c et  $|\eta| < 1.2$  avec les barres d'erreurs (pointillés)

A droite : Performances pour trois intervalles en  $p_T$  (triangles verts :  $15 < p_T < 35$  GeV/c, disques rouges :  $35 < p_T < 55$  GeV/c, carrés bleus :  $55 < p_T < 95$  GeV/c) et deux intervalles en  $|\eta|$  (traits pleins :  $|\eta| < 1.2$ , pointillés :  $|\eta| > 1.2$ )

Cet algorithme combiné permet un gain d'environ 30% sur l'efficacité d'identification des jets  $b$  pour un même taux de mauvaise identification. Il sera utilisé dans les prochaines analyses de  $D\bar{O}$ .



Second Witch —

*Fillet of a fenny snake,  
In the cauldron boil and bake;  
Eye of newt and toe of frog,  
Wool of bat and tongue of dog,  
Adder's fork and blind-worm's sting,  
Lizard's leg and owlet's wing,  
For a charm of powerful trouble,  
Like a hell-broth boil and bubble.*

William Shakespeare, “*Macbeth*”, Act IV, Scene 1.

# 4

## Production du quark top par voie électrofaible à DØ

### 4.1 Introduction

Dans le Modèle Standard, le quark top peut être produit par interaction électrofaible. Ce mode de production a été étudié au Run I et au Run II de Tevatron par les expériences DØ [82, 83] et CDF [84, 85]. Ces analyses n’avaient pas la sensibilité nécessaire pour mettre en évidence ce processus et ont posé des limites à 95 % de niveau de confiance sur les sections efficaces de production :

$$\sigma_s < 6.4 \text{ pb}, \sigma_t < 5 \text{ pb} \quad [83] \quad (4.1)$$

Les diagrammes de Feynman à l’ordre de Born liés à la production électrofaible du quark top ont été décrits au § 1.3.3.2 de même que les sections efficaces théoriques (cf. tableau 1.4). Les deux diagrammes dominants au Tevatron sont rappelés sur la figure 4.1. Au Tevatron, la production associée  $Wt$  est fortement défavorisée et notre étude se limite donc aux voies  $s$  et  $t$ .

En présupposant qu’il n’existe que trois familles de quarks, l’élément de matrice CKM  $|V_{tb}|$ , peut être déterminé indirectement par les contraintes d’unitarité :

$$|V_{tb}|^2 + |V_{cb}|^2 + |V_{ub}|^2 = 1 \quad (4.2)$$

$$0.9990 < |V_{tb}| < 0.9992 \quad [5] \quad (4.3)$$

La mesure du rapport d’embranchement  $t \rightarrow Wb$  donne aussi accès à  $|V_{tb}|$ . La mesure de ce rapport d’embranchement dans des événements  $t\bar{t}$  a été réalisée à DØ conjointement

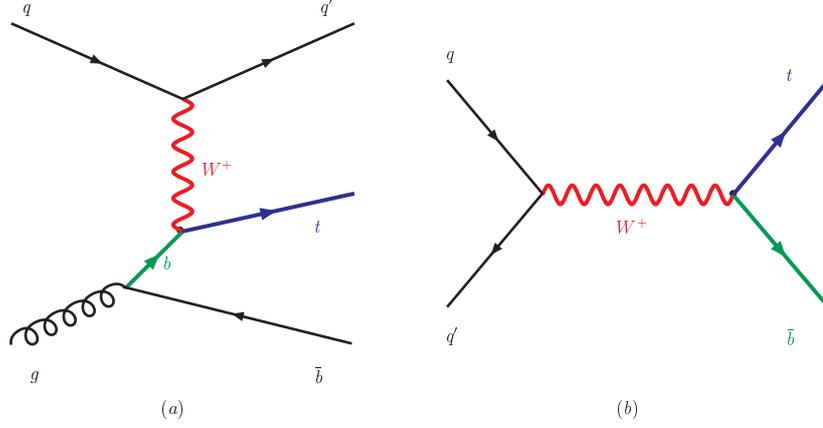


FIG. 4.1: Processus dominants pour la production de *single top* au Tevatron dans la voie *t* (a) et dans la voie *s* (b).

à la mesure de section efficace de production  $t\bar{t}$  [86] :

$$R_{tb} = \frac{\Gamma_{t \rightarrow Wb}}{\Gamma_{t \rightarrow Wq}} = \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2} = 1.03_{-0.17}^{+0.19} \text{ (stat+syst)} \quad (4.4)$$

$$\text{pour } \sigma_{t\bar{t}} = 7.9_{-1.5}^{+1.7} \text{ (stat+syst)} \pm 0.5 \text{ (luminosité) pb} \quad (4.5)$$

Cette dernière mesure ne peut être sensible à un quatrième quark lourd de type *down*  $b'$  que pour  $m_{b'} < m_t - m_W$ .

La section efficace de production de *single top* est directement reliée au couplage par courant chargé du quark top et permet donc une mesure directe de  $|V_{tb}|$ . Il n'y a plus de limite de masse sur une quatrième famille de quarks. La mesure séparée des sections efficaces des voies *s* et *t* peut également donner des informations concernant différents modèles de nouvelle physique. La voie *s* sera sensible à l'existence de bosons chargés lourds (Higgs chargé, boson de jauge  $W'$ , technicouleur...). La voie *t* pourra être sensible à des couplages non standard du quark top (*FCNC*) ou à une nouvelle famille de quarks. Certains de ces modèles seront discutés dans la section 4.10.2.

## 4.2 Stratégie de l'analyse

### 4.2.1 Signature des événements *single top*

La voie *t* correspond à des états finals  $tq\bar{b}$ ,  $\bar{t}q\bar{b}$  (le  $b$  provenant de la conversion d'un gluon), ainsi que  $tq$  et  $\bar{t}q$ . Ce processus sera appelé *tqb*. Les états finals en voie *s* sont  $t\bar{b}$  et  $\bar{t}b$ . Ce processus sera référencé par la dénomination *tb*.

Dans tous les cas, le quark top se désintègre en  $t \rightarrow Wb$ . Seuls les modes leptoniques

pour le  $W$  sont considérés ( $W \rightarrow \mu\nu_\mu$  et  $W \rightarrow e\nu_e$ ) car la présence d'un électron ou d'un muon isolé, de grande impulsion transverse permet à la fois de définir des critères de déclenchement efficaces et de réduire le bruit de fond multijet. La désintégration du  $W$  en lepton  $\tau$  est incluse dans ces modes quand le  $\tau$  se désintègre en leptons ( $W \rightarrow \tau\nu_\tau \rightarrow \ell\nu_\ell\nu_\tau$ , avec  $\ell = \mu$  ou  $e$ ). Les états finals recherchés sont donc  $\ell\nu_\ell bb$  et  $\ell\nu_\ell bqb$ . Le lepton chargé est isolé, c'est-à-dire qu'il est séparé des jets. La figure 4.2 montre les distributions en impulsion transverse et en pseudorapidité de chaque particule dans les états finals  $\ell\nu_\ell bb$  et  $\ell\nu_\ell bqb$ . Ces distributions cinématiques permettent d'identifier la signature des événements *single top* dans le détecteur.

Dans la voie  $s$ , quarks et lepton chargé sont centraux et de grande impulsion transverse. Seuls les jets de  $p_T > 15$  GeV/c sont utilisés dans l'analyse, car la correction en énergie des jets n'est pas valide en deçà de cette valeur (cf. § 2.2.3.6). Il semble donc possible de reconstruire les deux jets présents dans cette voie. De plus, d'après la figure 3.20 l'identification des jets  $b$  est meilleure à grande impulsion transverse ( $p_T > 60$  GeV/c) et dans la région centrale du détecteur ( $|\eta| < 1.5$ ). Il est donc envisageable d'identifier les deux jets  $b$ . Enfin le neutrino, qui ne figure pas dans les distributions de la figure 4.2, a des distributions en  $p_T$  et  $\eta$  similaires à celles du lepton chargé, les deux particules provenant de la désintégration à deux corps du  $W$ . Le neutrino n'est pas détecté et apparaît sous forme d'énergie transverse manquante. La signature dans le détecteur de ces événements est donc un lepton de grand  $p_T$ , de l'énergie transverse manquante et deux jets de quarks  $b$ . D'autres jets légers issus de la radiation de gluons dans l'état initial ou final ou d'interaction multiple peuvent s'y rajouter.

Pour la voie  $t$ , le quark  $b$  issu du top de même que le lepton chargé présentent des caractéristiques cinématiques similaires à celles de la voie  $s$  et donc sont identifiables. En revanche le second jet  $b$  (issu généralement de la conversion d'un gluon) est beaucoup plus mou que dans la voie  $s$  et possède une distribution en pseudorapidité plus large. Il n'est pas souvent reconstruit et a une faible probabilité d'être identifié comme un jet  $b$ . Le jet léger est suffisamment énergétique pour être reconstruit. C'est un jet dirigé vers l'avant du détecteur. Sa distribution en pseudorapidité est asymétrique, présentant un excès à  $\eta > 0$  pour une production de top et à  $\eta < 0$  pour un anti-top. La charge du lepton est directement corrélée à la charge du (anti-)quark top, et la variable  $Q_\ell \times \eta_{jet\ léger}$  peut permettre de caractériser le processus. La signature dans le détecteur de ces événements est un lepton de grand  $p_T$ , de l'énergie transverse manquante, un jet  $b$  central, un jet léger et éventuellement un second jet  $b$  vers l'avant du détecteur.

### 4.2.2 Principaux bruits de fond

Les bruits de fond principaux proviennent de processus ayant soit la même signature que les événements *single top*, soit des signatures voisines avec un ou plusieurs objets mal identifiés (jets ou leptons mal reconstruits, mauvaise identification de jets  $b$ ). Les processus comportant des états finals composés d'un boson  $W$  et de jets sont :

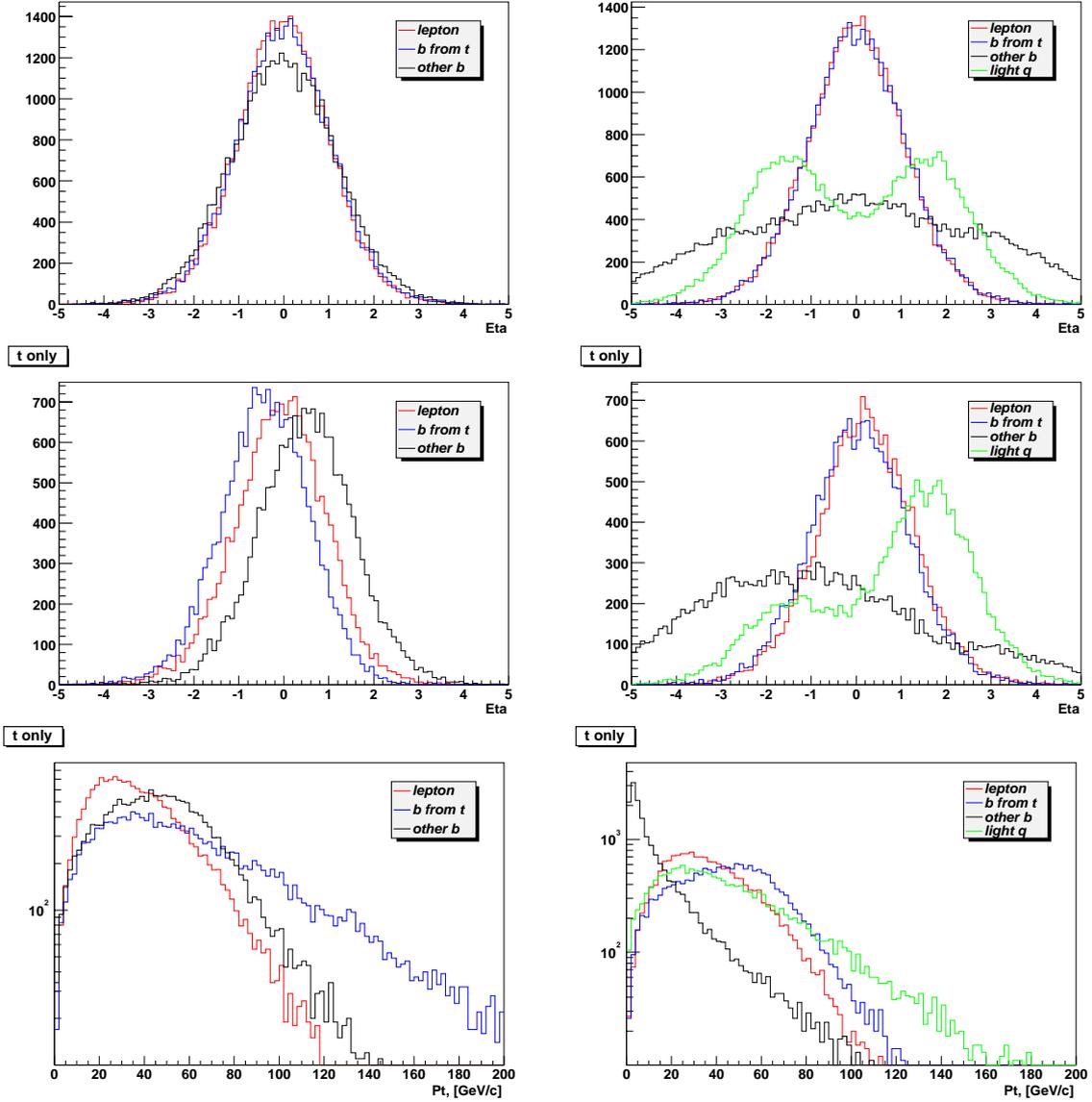


FIG. 4.2: Distributions partoniques calculées avec CompHEP [87] (générateur à éléments de matrice à l'ordre dominant). Pour la voie  $t$  une version modifiée de CompHEP est utilisée pour traiter correctement les processus  $(2 \rightarrow 2)$  et  $(2 \rightarrow 3)$ .

En haut : Distributions en pseudorapidité pour  $t\bar{b}/t\bar{b}$  (gauche) et  $tq\bar{b}/t\bar{q}b$  (droite).

Au centre : Distributions en pseudorapidité pour  $t\bar{b}$  (gauche) et  $tq\bar{b}$  (droite) uniquement.

En bas : Distributions en impulsion transverse pour  $t\bar{b}$  (gauche) et  $tq\bar{b}$  (droite).

La ligne rouge représente le lepton chargé, la ligne bleue le quark  $b$  issu du top, la ligne noire le second quark  $b$  et la ligne verte le quark léger (voie  $t$  uniquement).

- La production associée d'un boson  $W$  et de jets (figures 4.3a et 4.3b). Elle constitue le bruit de fond dominant. La section efficace  $Wjj$  (jets légers) est 100 fois plus élevée que la section efficace de production *single top*. Les événements  $Wbb$  ne contribuent qu'à 1.5% des événements  $W$ +jets. La présence de jets de saveurs lourdes fait du processus  $Wbb$  un bruit de fond irréductible. Après identification d'un jet  $b$  dans l'événement, les deux processus  $Wjj$  et  $Wbb$  contribuent dans des proportions voisines au bruit de fond total. Les sections efficaces des processus  $W$ +jets dépendent sensiblement des coupures imposées lors de la génération. Le choix des diagrammes à inclure dans la simulation à l'ordre dominant ( $LO$  : *Leading Order*), les problèmes de double comptage ainsi que les  $PDF$  contribuent aux grandes incertitudes sur les sections efficaces (de l'ordre de 20 à 30%).
- La production de paires de quarks top, dans les modes de désintégration lepton+jets (figure 4.3d) et dileptons. La section efficace de production est deux à trois fois supérieure à celle du quark top isolé mais les événements possèdent une plus grande multiplicité de jets (lepton+jets) ou de leptons (dilepton).
- La production d'événements diboson  $WZ$  et  $WW$ , dans les canaux de décroissance  $\ell\nu_\ell jj$  (figure 4.3c). Ces processus ont des sections efficaces voisines du *single top* et contiennent rarement des jets  $b$ . La contribution résiduelle après étiquetage d'un jet  $b$  est quasiment négligeable.

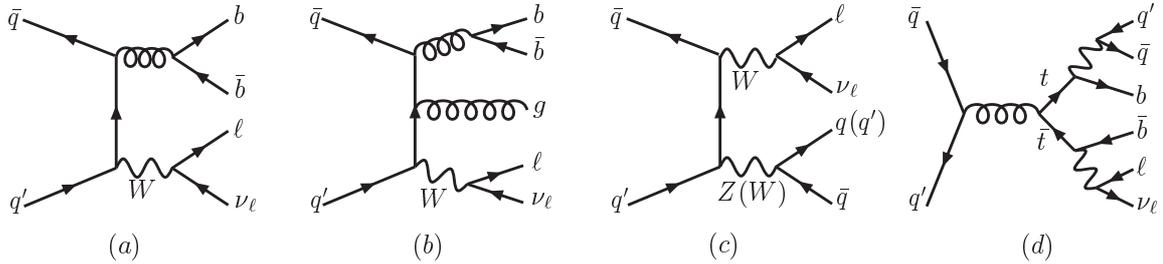


FIG. 4.3: Principaux bruits de fond au signal *single top*.

- (a) et (b) : Exemples de production associée  $W$ +jets. (a) :  $Wbb$ . (b) :  $Wbbj$   
 (c) : Production de di-bosons  $WZ, WW \rightarrow \ell\nu_\ell qq$   
 (d) : Production de paire  $t\bar{t}$  dans le canal lepton+jets.

A ces bruits de fond physiques il faut ajouter les événements multijet issus de processus QCD, où un lepton isolé est reconstruit par erreur :

- Un électron isolé peut être reconstruit à la place d'un jet très électromagnétique.
- Un muon de grand  $p_T$  peut provenir d'un jet (décroissance de hadrons beaux, charmés, kaons et pions) et apparaître isolé soit parce que le jet n'est pas reconstruit, soit parce que le muon sort du cône du jet.

Ces événements multijet avec un faux lepton isolé sont rares mais la très grande section efficace multijet les rend non négligeables relativement au signal recherché. Les sections efficaces des différents bruits de fond sont résumées dans le tableau 4.1.

Processus	Section efficace (pb)	Mode de désintégration	Rapport d'embranchement
$t\bar{t}$	$6.77 \pm 1.22$	$\ell^+ \ell^- \nu_\ell \bar{\nu}_\ell b\bar{b}$	$0.111 \pm 0.002$
$t\bar{t}$	$6.77 \pm 1.22$	$\ell \nu_\ell j b\bar{b}$	$0.444 \pm 0.009$
$W + jj$	$287 \pm 57$	$\ell \nu_\ell jj$	$0.111 \pm 0.002$
$W + b\bar{b}$	$3.35 \pm 0.67$	$\ell \nu_\ell b\bar{b}$	$0.111 \pm 0.002$
$WW$	$8.91 \pm 0.08$	$\ell \nu_\ell jj$	$0.393 \pm 0.008$
$WZ$	$2.84 \pm 0.01$	$\ell \nu_\ell jj$	$0.393 \pm 0.008$

TAB. 4.1: Sections efficaces théoriques et rapports d'embranchement pour les principaux bruits de fonds. La section efficace  $t\bar{t}$  est calculée à l'ordre  $NNLO - NNNLL$  [17] et l'erreur systématique inclut l'incertitude sur la masse du quark top [18]. Les autres sections efficaces sont calculées à l'ordre  $NLO$  avec MCFM [88]. Les sections efficaces  $W + jets$  sont sensibles aux coupures partoniques utilisées lors de la génération. Elles sont données pour  $p_T(parton) > 8 \text{ GeV}/c$  et  $|\eta(parton)| < 3.0$ . L'erreur relative associée est estimée à 20%.

### 4.2.3 Structure de l'analyse

Le but de cette analyse est d'améliorer les limites existantes sur les sections efficaces de production *single top* dans les voies  $s$  et  $t$ , voire de mettre en évidence un signal et de déterminer les sections efficaces de production. Pour améliorer la sensibilité de l'analyse différents canaux orthogonaux sont traités séparément puis combinés.

L'analyse se déroule en quatre étapes :

- 1<sup>re</sup> étape : **La sélection des événements** en fonction de la signature recherchée. Les critères de sélection diffèrent pour l'identification des leptons (électron ou muon) et ces deux canaux d'analyse sont traités séparément.
- 2<sup>me</sup> étape : **L'identification des jets de quark  $b$** . Cette étape est différenciée de la sélection précédente car l'utilisation des *TRF* nécessite la mise en oeuvre de techniques particulières dans le traitement des événements simulés. Dans chaque canal d'analyse, les événements sont séparés en fonction du nombre de jets  $b$  identifiés (exactement un ou au moins deux). Ce découpage permet d'obtenir une meilleure séparation des voies  $s$  et  $t$ , puisque le second jet  $b$  a une moindre probabilité de reconstruction et d'identification dans la voie  $t$  que dans la voie  $s$ .
- 3<sup>me</sup> étape : **La construction de variables discriminantes**. Pour chaque canal d'analyse, quatre fonctions de vraisemblance combinant diverses variables discriminantes sont élaborées. Chacune de ces fonctions constitue une variable discriminante permettant de séparer un signal (voie  $s$  ou  $t$ ) d'un des principaux bruits de fond ( $t\bar{t}$  ou  $W + jets$ ). Les distributions de ces nouvelles variables discriminantes sont ensuite déterminées pour le signal et les fonds ainsi que pour les données réelles.

- *4<sup>me</sup> étape* : **Le calcul des limites**. Les distributions des fonctions de vraisemblance sont utilisées pour déterminer des limites à 95% de niveau de confiance sur les sections efficaces de chaque processus, en combinant les différents canaux (électron, muon, simple et double étiquetage). Les limites sont déterminées par une méthode bayésienne.

La structure de l'analyse est résumée par la figure 4.4.

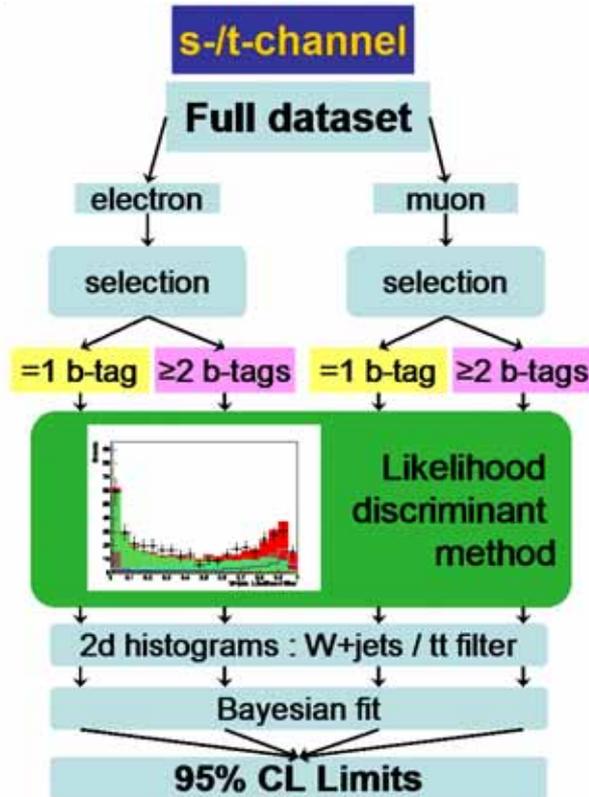


FIG. 4.4: Structure générale de l'analyse avec les étapes successives de sélection, d'étiquetage des jets  $b$ , d'analyse discriminante et d'extraction des limites sur les sections efficaces.

### 4.3 Données et événements simulés utilisés pour l'analyse

Les données réelles utilisées dans cette analyse ont été collectées entre août 2002 et août 2004. Ces données ont été traitées avec la version `lpanema` du programme `TopAnalyze` [73]. Les événements enregistrés lorsque le fonctionnement d'un ou de plusieurs sous-détecteurs est douteux sont rejetés de l'analyse en utilisant les listes de runs valides officielles. Dans le canal muon, seuls les runs avec une qualité "raisonnable" pour le

système à muons sont utilisés. Ces données correspondent à une luminosité intégrée de  $366 \pm 24 \text{ pb}^{-1}$  pour le canal d'analyse électron et  $363 \pm 24 \text{ pb}^{-1}$  pour le canal muon. Plusieurs lots d'événements simulés sont utilisés pour estimer l'acceptance du signal et des bruits de fond. Les méthodes mises en œuvre pour ces estimations seront détaillées au § 4.7. Les événements simulés ont été processés avec la version `lpanema` du programme TopAnalyze.

Ces événements sont produits avec des générateurs à éléments de matrice à ordre fixe (*FOME*, *Fixed Order Matrix Element*). La masse du quark top est fixée à  $175 \text{ GeV}/c^2$ . Les quadri-vecteurs des partons finals sont ensuite introduits dans le générateur PYTHIA afin de simuler les phases de fragmentation et d'hadronisation ainsi que les processus de radiations initiales et finales et les interactions multiples. La version 6.2 de PYTHIA est utilisée avec la version CTEQ5L des fonctions de distribution de partons. Les paramètres d'interactions multiples sont ajustés sur les données de CDF. Les désintégrations des hadrons beaux et des leptons  $\tau$  sont assurées par les outils EVTGEN [89] et TAUOLA [90]. Le signal (voies  $s$  et  $t$ ) a été produit avec une version modifiée du générateur CompHEP, SingleTop [87]. Quatre lots correspondant aux quatre états finals  $e\nu_e bb$ ,  $\mu\nu_\mu bb$  (voie  $s$ ),  $e\nu_e bqb$  et  $\mu\nu_\mu bqb$  (voie  $t$ ) ont été produits. Ces états finals incluent les désintégrations leptoniques du lepton  $\tau$ . Quarks top et anti-top sont produits en quantité égale.

Les bruits de fond  $t\bar{t}$  (lepton+jets et dilepton), dibosons ( $WW$  et  $WZ$ ) et  $W$ +jets ( $Wbb$  et  $Wjj$ ) sont produits avec le générateur ALPGEN [91].

Pour les  $W$ +jets, deux lots ont été produits,  $Wjj$  et  $Wbb$ . Seuls deux partons sont produits dans l'état final, au niveau d'ALPGEN. Les multiplicités de partons supérieures sont assurées par les radiations initiales et finales dans PYTHIA. Le lot  $Wjj$  contient toutes les saveurs de quark, à l'exception des quarks beaux. Les coupures suivantes sont appliquées aux partons générés :  $p_T(\text{parton}) > 8 \text{ GeV}/c$ ,  $|\eta(\text{parton})| < 3.0$ ,  $\Delta R(\text{parton}_1, \text{parton}_2) > 0.4$ .

Le tableau 4.2 décrit les principales caractéristiques des lots Monte Carlo utilisés dans cette analyse.

## 4.4 Sélection initiale

La sélection initiale permet de supprimer une partie importante du bruit de fond multijet ainsi que les événements mal reconstruits tout en conservant une quantité analysable du signal recherché. Cette série de coupures est commune à toutes les analyses *single top* effectuées dans  $D\bar{O}$ . Les critères de reconstruction des objets physiques ont été décrits au § 2.2.

### 4.4.1 Critères de déclenchement

Les événements doivent avoir passé un critère de déclenchement incluant un électron ou un muon et un jet. Les déclenchements dans le canal électron demandent deux jets,

Processus	Générateur <i>FOME</i>	Pdf	Nombre d'événements
Signal <i>single top</i>			
$tb \rightarrow e\nu_e bb$	SingleTop	CTEQ6M	32500
$tb \rightarrow \mu\nu_\mu bb$	SingleTop	CTEQ6M	31000
$tqb \rightarrow e\nu_e bqb$	SingleTop	CTEQ6M	33000
$tqb \rightarrow \mu\nu_\mu bqb$	SingleTop	CTEQ6M	31500
Bruits de fond			
$t\bar{t} \rightarrow \ell\bar{\nu}_\ell jjbb$	ALPGEN v1.3	CTEQ6.1M	191300
$t\bar{t} \rightarrow \ell^+ \ell^- \nu_\ell \bar{\nu}_\ell b\bar{b}$	ALPGEN v1.3	CTEQ6.1M	97750
$W + jj \rightarrow e\nu_e jj$	ALPGEN v1.3	CTEQ5L	99500
$W + jj \rightarrow \mu\nu_\mu jj$	ALPGEN v1.3	CTEQ5L	99000
$W + bb \rightarrow e\nu_e bb$	ALPGEN v1.3	CTEQ5L	189500
$W + bb \rightarrow \mu\nu_\mu bb$	ALPGEN v1.3	CTEQ5L	188000
$WW \rightarrow \ell\nu_\ell jj$	ALPGEN v1.2 PYTHIA v6.2	CTEQ5L	23000
$WZ \rightarrow \ell\nu_\ell jj$	ALPGEN v1.2 PYTHIA v6.2	CTEQ5L	23000

TAB. 4.2: Caractéristiques des lots Monte Carlo utilisés pour l'analyse

les électrons étant également reconstruits comme des jets au niveau du système de déclenchement. Le déclenchement n'est pas décrit dans la simulation. L'efficacité de chaque critère de déclenchement est déterminée dans les données réelles puis utilisée pour pondérer les événements simulés. Ces pondérations seront discutées au § 4.6. Au cours du temps, la liste des critères de déclenchement disponibles évolue, en fonction de la mise en service de nouveaux systèmes de déclenchement et de l'augmentation de la luminosité instantanée. Les versions v8 à v12 de la liste de déclenchement ont été utilisées pendant la période totale de prise de données correspondant à cette analyse. Le tableau 4.3 présente les critères de déclenchement utilisés dans cette analyse pour différentes périodes de prise de données ainsi que la luminosité intégrée associée. Le détail de chaque critère de déclenchement est donné ci-après :

- EM15\_2JT15
  - Niveau 1 : au moins une tour électromagnétique du calorimètre avec  $E_T > 10$  GeV/c et deux tours complètes avec  $E_T > 5$  GeV.
  - Niveau 2 : au moins un objet électromagnétique avec  $E_T > 10$  GeV/c et  $Emf > .85$ ; au moins deux candidats jets avec  $E_T > 10$  GeV/c.
  - Niveau 3 : au moins un candidat électron avec  $E_T > 15$  GeV/c et deux candidats jets avec  $E_T > 15$  GeV/c.
- E1\_SHT15\_2JT20
  - Niveau 1 : au moins une tour électromagnétique du calorimètre avec  $E_T > 11$  GeV/c.
  - pas de critère particulier au niveau 2.

Nom	Liste	Période de prise de données	Luminosité (pb <sup>-1</sup> )
canal électron			
EM15_2JT15	v8.2-v12.0	Août 2002 - Juillet 2003	110
E1_SHT15_2J20	v12.0-v13.0	Juillet 2003 - Juin 2004	210
E1_SHT15_2J_J25	v13.0-v13.3	Juin 2004 - Août 2004	46
total			366
canal muon			
MU_JT20_L2M0	v8.2-v12.0	Août 2002 - Juillet 2003	114
MU_JT25_L2M0	v12.0-v13.0	Juillet 2003 - Juin 2004	210
MUJ2_JT25	v13.0-v13.2	Juin 2004 - Août 2004	26
MU_J2_JT25_LM3	v13.2-v13.3	Août 2004	13
total			363

TAB. 4.3: Critères de déclenchement utilisés pour les canaux électron et muon.

- Niveau 3 : au moins un candidat électron avec  $E_T > 15$  GeV/c avec une sélection sur la forme de la gerbe électromagnétique et deux candidats jets avec  $E_T > 20$  GeV/c.
- E1\_SHT15\_2J25, identique au précédent avec un jet satisfaisant  $E_T > 25$  GeV/c au niveau 3.
- MU\_JT20\_L2M0
  - Niveau 1 : au moins un scintillateur du système à muons et une tour du calorimètre avec  $E_T > 3$  GeV/c.
  - Niveau 2 : au moins un candidat muon et un candidat jet de  $E_T > 10$  GeV/c.
  - Niveau 3 : au moins un candidat jet avec  $E_T > 20$  GeV/c.
- MU\_JT25\_L2M0, identique au précédent avec une coupure  $E_T > 25$  GeV/c sur le jet au niveau 3.
- MUJ2\_JT25 et MU\_J2\_JT25\_LM3 (même critère, avec changement de dénomination entre les listes v13 et v13.2)
  - Niveau 1 : au moins un scintillateur ou une chambre à fils du système à muons et une tour du calorimètre avec  $E_T > 5$  GeV/c.
  - Niveau 2 : au moins un candidat muon et un candidat jet de  $E_T > 8$  GeV/c.
  - Niveau 3 : au moins un candidat jet avec  $E_T > 25$  GeV/c.

#### 4.4.2 Vertex primaire

De manière analogue à la certification de *JLIP*, les événements utilisés dans l'analyse doivent avoir un vertex primaire correctement reconstruit. Pour cela le vertex doit être dans l'acceptance du *SMT* et avoir au moins trois traces associées.

- $N_{\text{traces}} \geq 3$

$$- |z_{\text{vertex}}| < 60 \text{ cm}$$

Ces coupures sont illustrées sur la figure 4.5.

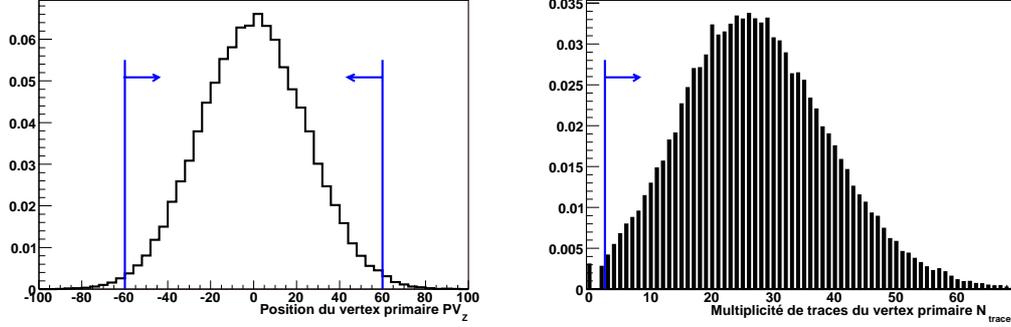


FIG. 4.5: Position  $z$  (à gauche) et multiplicité de traces (à droite) du vertex primaire dans des événements simulés *single top* voie  $t$ . Les flèches bleues indiquent les coupures de sélection.

### 4.4.3 Leptons isolés

Le signal recherché ainsi que les principaux bruits de fond physiques possèdent un lepton isolé de grande impulsion transverse issu de la désintégration d'un boson  $W$ . Deux niveaux inclusifs de qualité d'identification et d'isolation sont définis pour les leptons utilisés dans l'analyse, *loose* et *tight*. Chaque événement doit posséder un unique lepton isolé (électron ou muon) de qualité *tight*. Tout événement contenant en plus d'autres leptons de qualité *loose* est rejeté afin d'assurer l'orthogonalité de la sélection entre les canaux d'analyse électron et muon. Cette réjection basée sur les leptons *loose* supplémentaires permet de réduire les bruits de fond avec deux leptons dans l'état final dont l'un est mal reconstruit ( $t\bar{t}$ ,  $WW$ ,  $WZ$ ).

#### 4.4.3.1 Electrons

Electrons et photons sont reconstruits en formant des amas à partir des cellules du calorimètre électromagnétique. Un électron doit avoir une trace associée à l'amas calorimétrique.

Les électrons *loose* doivent vérifier les coupures standard d'identification d'un objet électromagnétique détaillé au § 2.2.3.4. Les photons sont également inclus dans cette sélection. Ces coupures visent notamment à séparer les électrons des jets. Elles portent sur :

- $Emf > 0.9$ ,
- $iso < 0.15$ , représentée sur la figure 4.6,
- $\chi^2$  de la matrice H à sept variables (voir 2.2.3.4)  $\chi_{Hmx7}^2 < 50$ ,

- l'impulsion transverse  $p_T > 15$  GeV/c,
- la pseudorapidité géométrique. Seuls les électrons dans la partie centrale du calorimètre sont utilisés,  $\eta_{det} < |1.1|$ .

Le niveau de qualité *tight* demande en plus que l'électron soit associé à une trace issue du vertex primaire. Une fonction de vraisemblance discriminante combinant 7 variables,  $\mathcal{L}_{EM}$ , permet également de séparer les électrons des jets très électromagnétiques. Les coupures supplémentaires définissant un électron *tight* sont alors :

- l'association à une trace : le  $\chi^2$  d'association électron-trace doit être défini,
- le paramètre d'impact de la trace selon l'axe du faisceau  $|IP_z| < 1$  cm,
- la fonction de vraisemblance,  $\mathcal{L}_{EM} > 0.75$ , illustrée sur la figure 4.6.

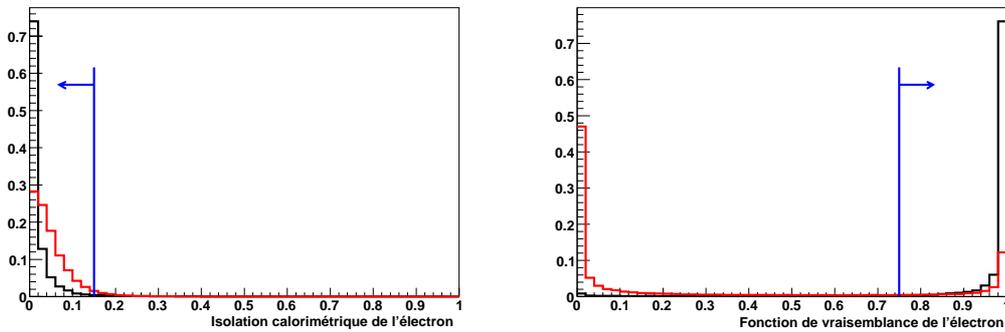


FIG. 4.6: Isolation calorimétrique (à gauche) et fonction de vraisemblance discriminante des électrons dans des événements simulés *single top* voie *t* (noir) et les données avant sélection (rouge). Les flèches bleues indiquent les coupures de sélection.

#### 4.4.3.2 Muons

Les muons sélectionnés doivent vérifier :

- le critère *medium* du groupe *muon-Id* [58],
- $n_{seg} = 3$ , c'est-à-dire que le muon a été reconstruit avec des impacts dans les segments A et BC du détecteur de muons et est associé à une trace reconstruite par les détecteurs de traces centraux (SMT et *CFT*),
- les coupures de réjection des rayons cosmiques :
  - les différences de temps des scintillateurs A et BC doivent satisfaire  $|t_A| < 10$  ns,  $t_{BC} < 10$  ns.
  - la trace associée au muon doit être correctement reconstruite. L'ajustement de la trace doit vérifier  $\chi^2 < 4$ .
  - la trace associée doit provenir du vertex primaire de l'interaction. Ceci est assuré par des contraintes sur la signification en paramètre d'impact transverse  $|S_{xy}| < 3$  et sur le paramètre d'impact selon l'axe du faisceau  $|IP_z| < 1$  cm.

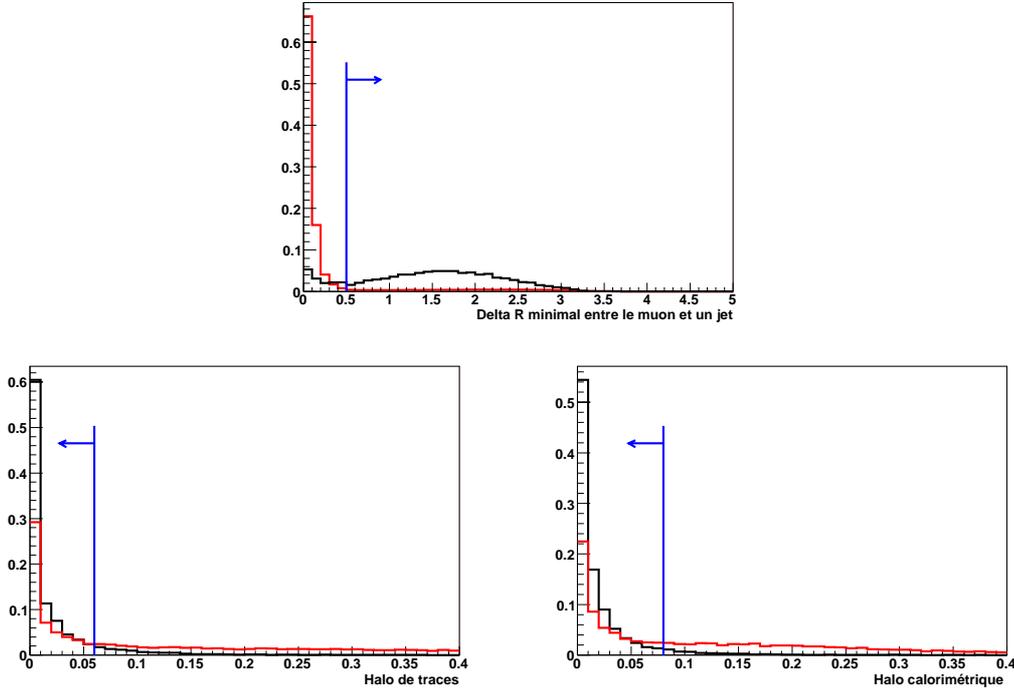


FIG. 4.7: Variables d'isolation des muons dans des événements simulés *single top* voie *t* (noir) et les données avant sélection (rouge). En haut : Séparation angulaire  $\Delta R(\mu, jet)$  pour les muons sélectionnés.

En bas : Halo de traces (à gauche) et halo calorimétrique (à droite) pour les muons sélectionnés après la coupure d'isolation *loose* et la coupure en impulsion transverse  $p_T > 15$  GeV/c. Les flèches bleues indiquent les coupures de sélection.

Les muons doivent également être isolés par rapport aux jets. Les deux niveaux de qualité *loose* et *tight* sont définis en fonction de la sévérité de cette isolation et de l'impulsion transverse du muon. L'isolation *loose* requiert uniquement une séparation angulaire minimale entre le muon et tous les jets de l'événement,  $\Delta R(\mu, jet) > 0.5$

Le niveau de qualité *tight* demande en plus que le muon soit isolé dans le détecteur de traces et dans le calorimètre. Ceci est obtenu en coupant sur les variables suivantes :

- le halo de traces :  $Rattrk = |\sum^{\text{traces}} p_T|/p_T(\mu) < 0.06$  pour les traces dans un cône de rayon  $R=0.5$  autour de la trace du muon,
- le halo calorimétrique :  $Rat11 = |\sum^{\text{cellules}} E_T|/p_T(\mu) < 0.08$  pour les cellules du calorimètre vérifiant  $0.1 < \Delta R(\text{cellule}, \mu) < 0.4$ .
- l'impulsion transverse  $p_T > 15$  GeV/c.

La figure 4.7 illustre les coupures d'isolation du muon. La coupure sur l'impulsion transverse du muon est placée au niveau *tight* afin de rejeter les événements contenant des muons supplémentaires de basse impulsion transverse.

#### 4.4.4 Jets

Les jets utilisés ont été reconstruits avec l'algorithme de cône de rayon  $R = 0.5$ . Les coupures standard § 2.2.3.5 de qualité des jets et la correction en énergie sont appliquées.

Pour des événements *single top*, on attend entre 2 et 3 jets, dont un jet  $b$  central, de grande impulsion transverse issu de la désintégration quark top. Les jets supplémentaires, notamment en voie  $t$  peuvent être vers l'avant du détecteur. Le jet de plus grande impulsion transverse doit satisfaire :

- $p_T > 25$  GeV/c,
- $\eta_{det} < 2.5$ .

Les jets supplémentaires doivent vérifier :

- $p_T > 15$  GeV/c,
- $\eta_{det} < 3.4$ .

Ces coupures sont illustrées sur la figure 4.8. Seuls les événements contenant de 2 à 4 jets passant ces critères sont retenus. Les événements à 4 jets contiennent principalement des événements  $t\bar{t}$  mais également une partie du signal, où un gluon de grand  $p_T$  a été rayonné. Enfin les événements contenant plus de deux jets de bruit (jets reconstruits non associés à une tour calorimétrique utilisée au niveau 1 du déclenchement [92]) sont rejetés de l'analyse.

#### 4.4.5 Energie transverse manquante

La désintégration leptonique du boson  $W$  produit également un neutrino de grande impulsion transverse. Ce neutrino n'interagit pas dans le détecteur et apparaît sous forme d'énergie transverse manquante,  $\cancel{E}_T$ . Les coupures suivantes sont appliquées :

- $\cancel{E}_T^{\text{JES}} > 15$  GeV,
- $15 < \cancel{E}_T < 200$  GeV/c

L'observable  $\cancel{E}_T^{\text{JES}}$  est l'énergie transverse manquante dans le calorimètre, tenant compte de la calibration des objets électromagnétiques et des jets. L'énergie transverse manquante entièrement corrigée  $\cancel{E}_T$  tient également compte des muons éventuellement présents dans l'événement. L'utilisation combinée de ces deux variables permet de réduire la sensibilité de l'analyse à des muons mal reconstruits. La coupure supérieure à 200 GeV/c rejette également des événements mal reconstruits. Les coupures sur  $\cancel{E}_T^{\text{JES}}$  et  $\cancel{E}_T$  sont représentées sur la figure 4.9.

#### 4.4.6 Coupures triangulaires

L'énergie transverse d'un lepton ou d'un jet mal reconstruit est généralement erronée. Si l'écart entre l'énergie reconstruite et l'énergie réelle est grand, il se traduit par l'apparition artificielle d'énergie transverse manquante colinéaire à l'objet mal reconstruit, dans le même sens ou en sens opposé. De tels événements apparaissent particulièrement

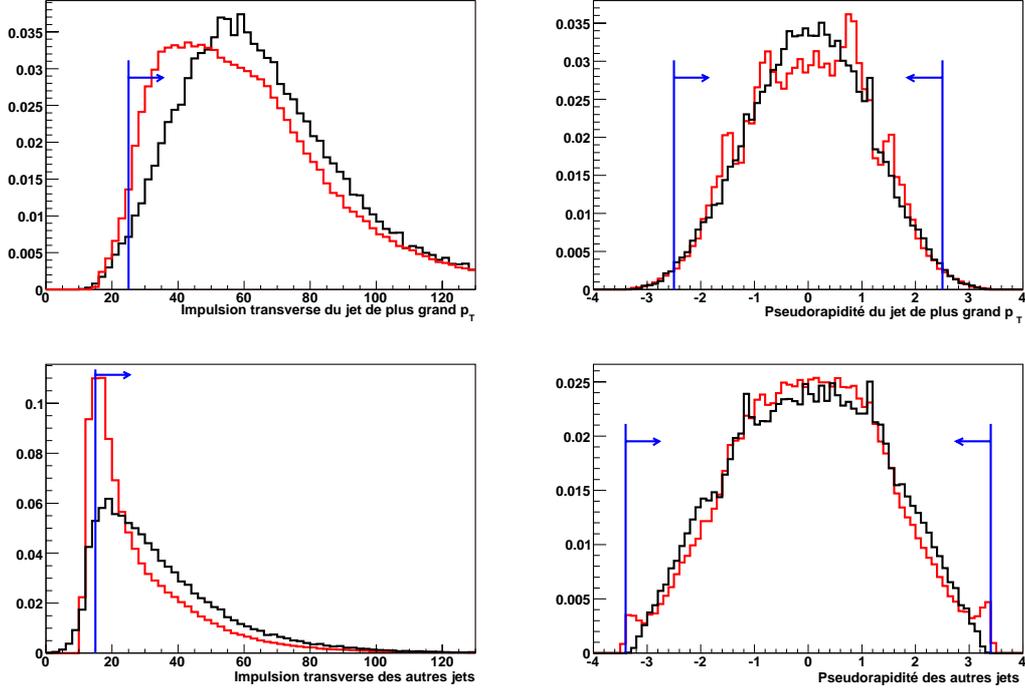


FIG. 4.8: Impulsion transverse (à gauche) et pseudorapidité géométrique (à droite) du jet de plus grand  $p_T$  (en haut) et des jets supplémentaires (en bas) dans des événements simulés *single top* voie  $t$  (noir) et les données avant sélection (rouge). Les flèches bleues indiquent les coupures de sélection.

dans le fond multijet où la sélection d'un lepton isolé est corrélée à une mauvaise reconstruction (jet reconstruit comme électron et donc mal calibré, muon provenant d'un jet mal reconstruit). Pour s'affranchir de ces événements, des coupures sont imposées sur la différence d'angle azimutal entre  $\cancel{E}_T$  et les objets reconstruits. Les valeurs de ces coupures varient avec  $\cancel{E}_T$ , d'où leur nom de coupures triangulaires. Ces coupures sont optimisées sur les données et la simulation. Pour le canal d'analyse électron, les coupures triangulaires sont :

- Dans le plan  $(|\Delta\phi(\text{electron}, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| > 1 - \frac{1}{80}\cancel{E}_T$ ,
- Dans le plan  $(|\Delta\phi(\text{jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| < 1.5 + \frac{\pi-1.5}{35}\cancel{E}_T$ ,
- Dans le plan  $(|\Delta\phi(\text{second jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| < 1.5 + \frac{\pi-1.5}{35}\cancel{E}_T$ .

Pour le canal d'analyse muon, les coupures triangulaires sont :

- Dans le plan  $(|\Delta\phi(\text{muon}, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| > 1.5 - \frac{1}{35}\cancel{E}_T$ ,
- Dans le plan  $(|\Delta\phi(\text{muon}, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| < 2 + \frac{\pi-2}{50}\cancel{E}_T$ ,
- Dans le plan  $(|\Delta\phi(\text{jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T) : |\Delta\phi| < 1.5 + \frac{\pi-1.5}{35}\cancel{E}_T$ .

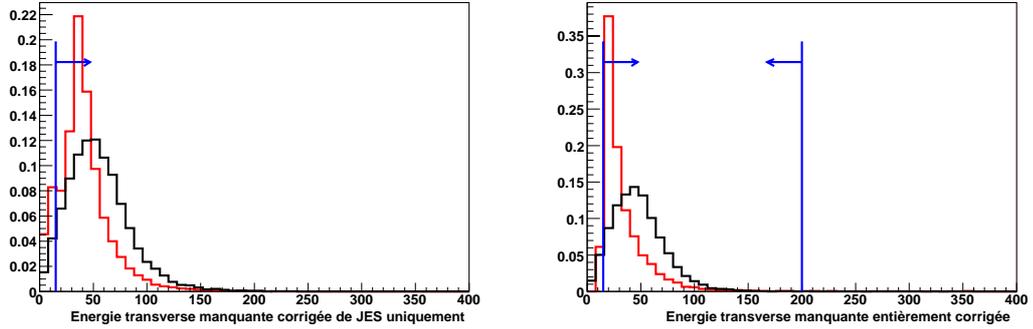


FIG. 4.9: Energie transverse manquante  $E_T^{\text{JES}}$  (à gauche) et  $E_T$  (à droite) dans les événements simulés *single top* voie *t* (noir) et les données avant sélection (rouge) contenant au moins deux jets et un électron isolé. Les flèches bleues indiquent les coupures de sélection.

Sélection	signal voie <i>s</i>		signal voie <i>t</i>	
	électron	muon	électron	muon
Vertex	0.98	0.98	0.98	0.98
Jets	0.72	0.73	0.72	0.72
Lepton	0.33	0.27	0.33	0.27
Energie manquante	0.30	0.24	0.30	0.24
Coupures triangulaires	0.27	0.22	0.27	0.22

TAB. 4.4: Efficacités des coupures de sélection cumulées pour chaque catégorie d'objets pour le signal simulé.

L'ensemble de ces coupures est illustré sur les figures 4.10 et 4.11 pour le signal simulé et des données multijet.

Les efficacités de ces sélections sur le signal simulé sont données dans le tableau 4.4.

## 4.5 Etiquetage des jets de quarks beaux

La dernière étape de la sélection est cruciale pour réduire les bruits de fond ne contenant que des jets légers ( $W$ +jets légers, dibosons, multijet). Chaque événement doit posséder au moins un jet *b* identifié avec *JLIP Tight* ( $\mathcal{P}_{jet}^+ < 0.003$ ). Pour des jets centraux ( $|\eta| < 1.2$ ) et d'impulsion transverse comprise entre 35 et 55 GeV/*c*, ce point de fonctionnement correspond à une efficacité d'identification des *b* d'environ 0.41 pour un taux de mauvaise identification de 0.003 (c.f. figure 3.30). Ces plages de pseudorapacité et d'impulsion correspondent approximativement à celles attendues pour un jet *b* provenant de la désintégration d'un quark top (cf. figure 4.2).

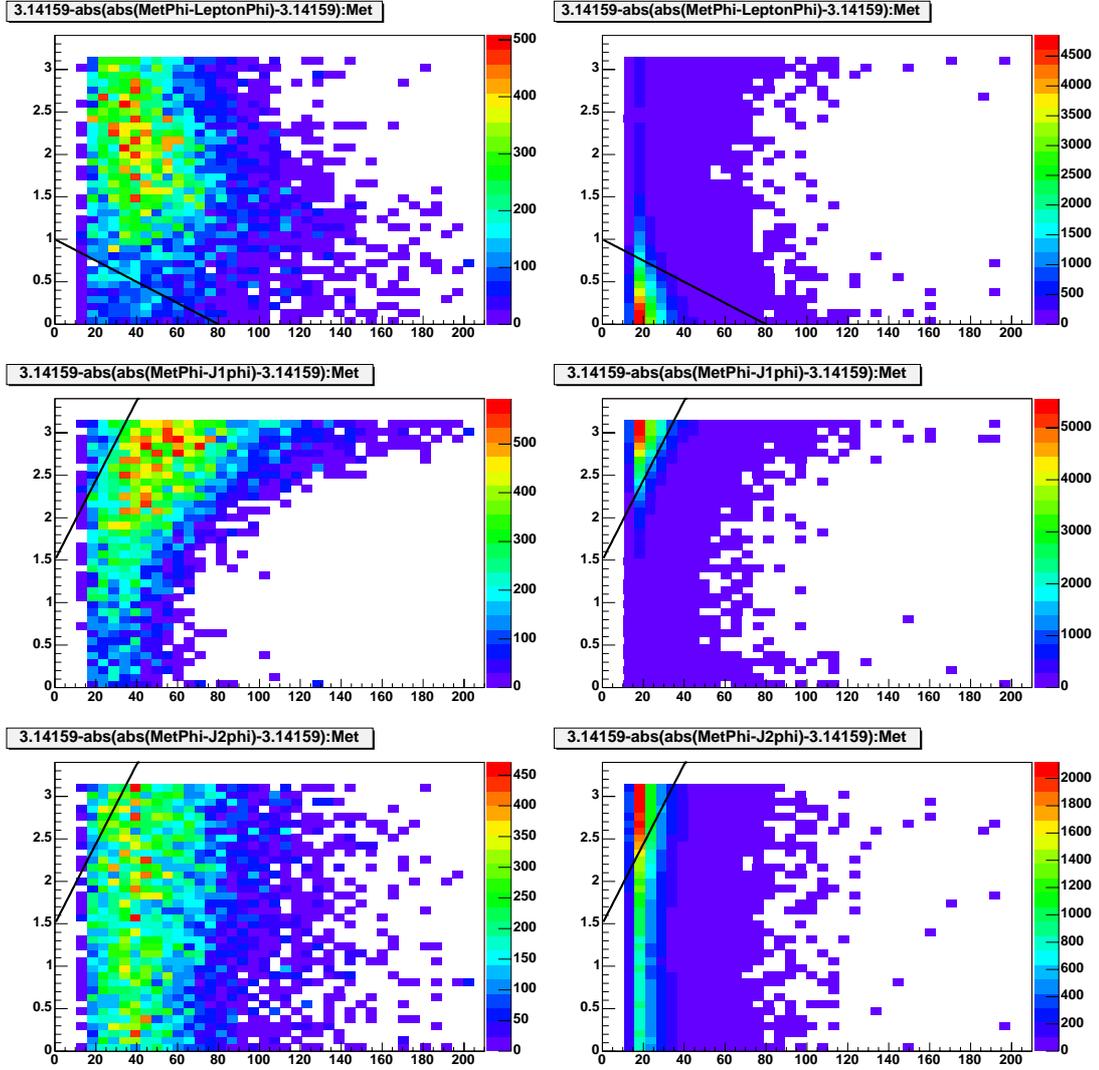


FIG. 4.10: Coupures triangulaires appliquées dans le canal électron dans le signal *single top* voie *t* simulé (à gauche) et des données multijet passant les coupures de sélection avec un électron *loose* (à droite).

En haut : plan  $(|\Delta\phi(\text{électron}, \cancel{E}_T)|, \cancel{E}_T)$

Au centre : plan  $(|\Delta\phi(\text{jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T)$

En bas : plan  $(|\Delta\phi(\text{second jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T)$ .

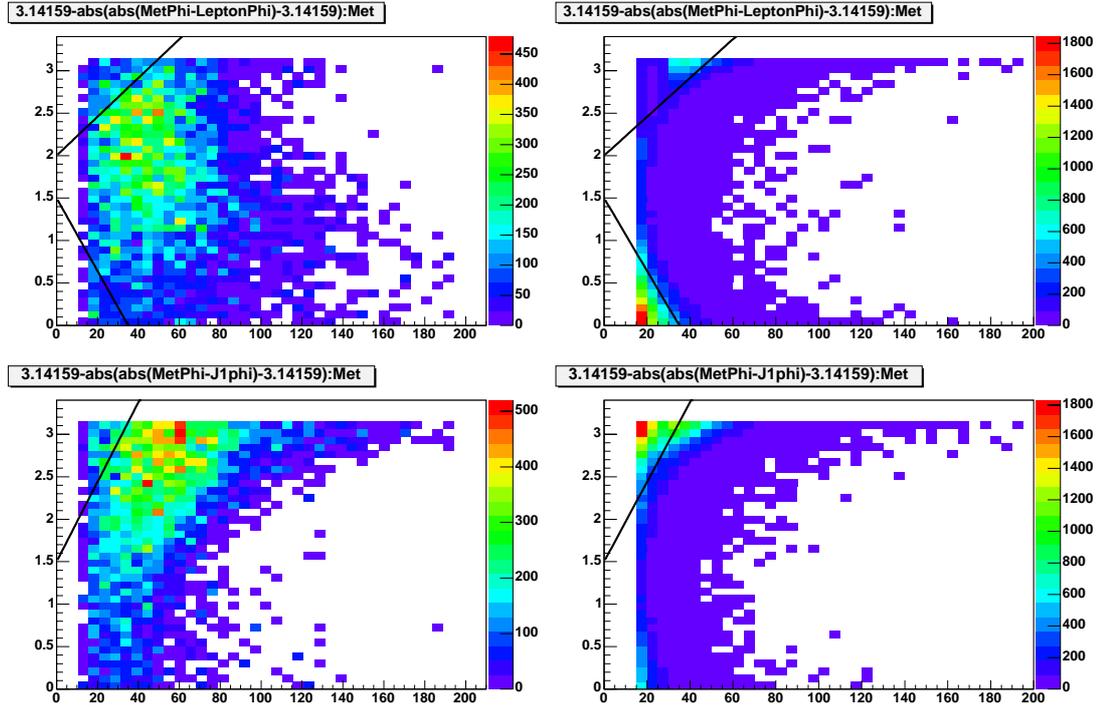


FIG. 4.11: Coupures triangulaires appliquées dans le canal muon dans le signal *single top* voie  $t$  simulé (à gauche) et des données multijet passant les coupures de sélections avec un muon *loose* (à droite).

En haut : plan  $(|\Delta\phi(\text{muon}, \cancel{E}_T)|, \cancel{E}_T)$

En bas : plan  $(|\Delta\phi(\text{jet de plus grand } p_T, \cancel{E}_T)|, \cancel{E}_T)$ .

Contrairement aux autres critères de sélection (déclenchement exclu), l'étiquetage des jets  $b$  n'est pas appliqué dans la simulation. Son effet est reproduit de manière statistique à l'aide d'une méthode de pondération par  $TRF$ .

Comme décrit au § 4.2.3, les événements de chaque canal d'analyse (électron et muon) sont séparés en fonction du nombre de jets  $b$  étiquetés en événements simplement ou doublement étiquetés. Les deux lots ainsi formés sont orthogonaux, permettant de les combiner à la fin de l'analyse. Chaque sous-lot ainsi défini correspond à un schéma d'étiquetage. Pour notre analyse deux modes de séparation ont été étudiés :

- **les schémas standard** : un événement simplement étiqueté possède un unique jet étiqueté *Tight* (schéma **STag**) ; un événement doublement étiqueté possède au moins 2 jets étiquetés *Tight* (schéma **DTag**). Ces schémas d'étiquetage des événements sont régulièrement utilisés dans les analyses de  $D\bar{O}$ , notamment pour l'étude de la production de paires  $t\bar{t}$ .
- **les schémas asymétriques**, qui utilisent un point de fonctionnement moins sévère ( $JLIP\ Loose, \mathcal{P}_{jet}^+ < 0.01$ ) pour l'étiquetage du second jet : un événement

simplement étiqueté doit avoir un jet *Tight* et aucun jet *Loose* (schéma **STagVeto**); un événement doublement étiqueté doit avoir un jet *Tight* et au moins un jet *Loose* supplémentaire (schéma **DTagAsym**). Le veto sur les jets *Loose* dans la définition de l'étiquetage simple assure l'orthogonalité des deux lots **STagVeto** et **DTagAsym**. Les schémas asymétriques permettent d'augmenter l'acceptance du lot doublement étiqueté et autorisent une meilleure séparation entre les voies *s* et *t*. L'utilisation d'un point de fonctionnement moins sévère permet également de diminuer l'erreur systématique liée aux *TRF*.

## 4.6 Pondération de la simulation

A diverses reprises, il a été indiqué que la qualité de la simulation ne permet pas d'utiliser directement les lots d'événements simulés pour estimer les bruits de fond. Les résolutions en impulsion des électrons, des muons et des jets sont dégradées dans la simulation pour s'accorder aux données. L'énergie transverse manquante est alors recalculée en conséquence. Il est alors nécessaire de pondérer les événements simulés afin de reproduire les efficacités de sélection des données. Ces pondérations sont de deux types. Si le critère de sélection peut être appliqué dans la simulation et que la différence d'efficacité entre les données et la simulation n'est pas excessive, le critère de sélection est effectivement appliqué et l'événement est pondéré d'un facteur correctif constant. Dans les cas plus extrême où l'écart d'efficacité est trop important, comme c'est le cas pour le déclenchement et l'identification des jets *b*, le critère de sélection n'est pas appliqué. Tous les événements sont pondérés par l'efficacité de la sélection estimée dans les données réelles. Ces pondérations dépendent de la topologie de l'événement, ce qui permet de corriger la normalisation et la forme des distributions des variables cinématiques.

### 4.6.1 Facteurs correctifs entre données réelles et simulation

Ces facteurs correctifs sont appliqués à tous les événements simulés après application des coupures de sélection correspondantes. Ils ont été déterminés par le groupe de travail "top" [93] de l'expérience  $D\bar{O}$  et sont utilisés dans les analyses de physique du quark top [94].

#### 4.6.1.1 Vertex primaire

L'efficacité de sélection du vertex primaire est mesurée dans les données réelles et la simulation pour les canaux électron et muon. Les coupures de sélection sont appliquées, à l'exception des coupures impliquant le vertex primaire (paramètre d'impact du lepton, la signification transverse du muon, fonction de vraisemblance de l'électron). Dans la simulation, l'efficacité a été étudiée pour différents processus. Les facteurs correctifs entre données réelles et simulées ainsi obtenus sont très proches de l'unité [95] :

- $\varepsilon_{vertex} = 1.004 \pm 0.018$  dans le canal électron,
- $\varepsilon_{vertex} = 0.991 \pm 0.004$  dans le canal muon.

L'erreur associée tient compte de la statistique disponible et des différences observées entre les différents processus simulés.

#### 4.6.1.2 Electrons

Le facteur de correction pour la sélection des électrons tient compte des efficacités de reconstruction des objets électromagnétiques  $\varepsilon_{Id}^{Data,MC}$ , de l'association avec une trace  $\varepsilon_{Trackmatch}^{Data,MC}$ , des coupures sur le paramètre d'impact  $\varepsilon_{z,PV-EM}^{Data,MC}$  et la fonction de vraisemblance  $\varepsilon_{Likelihood}^{Data,MC}$ . Le calcul de ces efficacités est explicité dans [96]. Ces efficacités sont mesurées sur des événements réels contenant un boson  $Z \rightarrow e^+e^-$  en étudiant la masse invariante di-électron. Les nombres de bosons  $Z$  dans le pic de masse avant et après application d'une coupure de sélection à l'un des électrons permettent de déterminer l'efficacité. Le facteur correctif global pour chaque électron simulé est ainsi :

$$\begin{aligned} \varepsilon_{e-ID} &= \frac{\varepsilon_{Id}^{Data}}{\varepsilon_{Id}^{MC}} \times \frac{\varepsilon_{Trackmatch}^{Data}}{\varepsilon_{Trackmatch}^{MC}} \times \frac{\varepsilon_{z,PV-EM}^{Data}}{\varepsilon_{z,PV-EM}^{MC}} \times \frac{\varepsilon_{Likelihood}^{Data}}{\varepsilon_{Likelihood}^{MC}} \\ &= 0.996 \times 0.9838 \times 0.9969 \times 0.9102 \\ &= 0.889 \pm 0.032. \end{aligned}$$

#### 4.6.1.3 Muons

Le facteur de correction pour la sélection des muons tient compte des efficacités de reconstruction d'un muon *medium* avec  $nseg = 3$   $\varepsilon_{Id}^{Data,MC}$ , de l'association avec une trace  $\varepsilon_{TrackEff}^{Data,MC}$ , de la qualité de la trace associée  $\varepsilon_{Track\chi^2}^{Data,MC}$ , des coupures sur le paramètre d'impact  $\varepsilon_{z,PV-MU}^{Data,MC}$ , sur la signification transverse  $\varepsilon_{DCA}^{Data,MC}$  et des coupures d'isolation  $\varepsilon_{TightIsol}^{MC}$ . Le calcul de ces efficacités est explicité dans [96]. Comme pour les électrons ces efficacités sont estimées à la résonance du boson  $Z$ . Le facteur correctif pour chaque muon simulé est finalement :

$$\begin{aligned} \varepsilon_{\mu-ID} &= \frac{\varepsilon_{Id}^{Data}}{\varepsilon_{Id}^{MC}} \times \frac{\varepsilon_{TrackEff}^{Data}}{\varepsilon_{TrackEff}^{MC}} \times \frac{\varepsilon_{Track\chi^2}^{Data}}{\varepsilon_{Track\chi^2}^{MC}} \times \frac{\varepsilon_{DCA}^{Data}}{\varepsilon_{DCA}^{MC}} \times \frac{\varepsilon_{TightIsol}^{Data}}{\varepsilon_{TightIsol}^{MC}} \times \frac{\varepsilon_{PV-MU}^{Data}}{\varepsilon_{PV-MU}^{MC}} \\ &= 1.00 \times 0.988 \times 0.9832 \times 0.9972 \times 0.9853 \times 0.9984 \\ &= 0.953 \pm 0.050 \end{aligned}$$

### 4.6.2 Efficacité de déclenchement

#### 4.6.2.1 Description générale

Pour corriger la simulation des effets de déclenchement, il faut déterminer la probabilité pour chaque événement simulé d'avoir passé le critère de déclenchement de-

mandé dans les données, connaissant les propriétés des objets reconstruits (jets, électrons, muons). Ces probabilités sont déterminées dans les données réelles, en comparant les objets physiques reconstruits au niveau de l'analyse aux objets reconstruits à chaque niveau du système de déclenchement. Dans les données réelles, un seul critère de déclenchement est requis par événement. Ce critère varie en fonction du canal d'analyse. Il dépend également de la période de prise de données, car la liste et la définition des critères de déclenchement disponibles dans l'expérience varient au cours du temps. La liste complète des critères utilisés dans notre analyse a été donnée au § 4.4.1.

Le système de déclenchement comporte trois niveaux successifs, que nous référencerons par L1, L2 et L3. Un critère de déclenchement est composé à chaque niveau d'une condition. Cette condition se décompose en termes portant sur la sélection d'un type d'objet particulier (jet, électron ou muon). La probabilité pour un événement de passer un critère de déclenchement est calculée à partir des efficacités individuelles de chaque terme. La référence [97] donne une description détaillée de la méthode de pondération probabiliste par événement. Nous n'en donnons ici qu'un aperçu général. La probabilité qu'un événement passe l'ensemble des conditions de déclenchement, peut se décomposer en terme de probabilités conditionnelles :

$$P(L1L2L3) = P(L1).P(L2|L1).P(L3|L1L2) \quad (4.6)$$

La condition de déclenchement à chaque niveau peut concerner des types d'objets différents. Les probabilités individuelles de chaque terme sont considérées comme décorrélées. Ainsi, chacune des probabilités conditionnelles de l'équation 4.6 peut s'écrire :

$$P(\text{objet1}, \text{objet2}) = P(\text{objet1}).P(\text{objet2}) \quad (4.7)$$

Enfin, les probabilités par classe d'objets sont déterminées à partir de l'efficacité pour un objet particulier de passer le critère considéré. Dans le cas de notre sélection, il n'y a qu'un seul lepton chargé (électron ou muon) par événement et la probabilité est directement égale à l'efficacité de déclenchement de cet objet. Le cas des jets est plus complexe car il y a plusieurs jets par événement. Le critère de déclenchement ne porte que sur un ou deux des jets sélectionnés. Les probabilités d'avoir au moins un jet ( $P_{\geq 1}$ ), exactement un jet ( $P_{=1}$ ) et au moins deux jets ( $P_{\geq 2}$ ) passant la condition de déclenchement parmi  $N$  jets sont alors :

$$P_{\geq 1} = 1 - P_{=0} = 1 - \prod_{i=0}^N (1 - P_{\text{jet}_i}) \quad (4.8)$$

$$P_{=1} = \sum_{i=0}^N \left( P_{\text{jet}_i} \prod_{j \neq i} (1 - P_{\text{jet}_j}) \right) \quad (4.9)$$

$$P_{\geq 2} = P_{\geq 1} - P_{=1} = 1 - \prod_{i=0}^N (1 - P_{\text{jet}_i}) - \sum_{i=0}^N \left( P_{\text{jet}_i} \prod_{j \neq i} (1 - P_{\text{jet}_j}) \right), \quad (4.10)$$

où  $P_{jet_i}$  est la probabilité du  $i^{\text{ème}}$  jet.

Tous les termes intervenant dans les critères de déclenchement considérés dans cette analyse sélectionnent les objets par leur impulsion transverse. Les efficacités sont donc mesurées en fonction du  $p_T$  de l'objet reconstruit. Différents domaines de pseudorapacité  $\eta$  ou d'angle azimutal  $\phi$  sont également considérés pour prendre en compte la géométrie du détecteur (partie centrale et bouchons du calorimètre, pied du détecteur de muons). Etant donné le nombre de termes à caractériser, ce travail est réparti entre les différents membres du groupe de travail "top" de  $D\bar{O}$ . Au cours de ce travail de thèse nous nous sommes concentrés sur le calcul des efficacités des termes L1, L2 et L3 portant sur les jets dans les listes v8 à v12 (critères de déclenchement EM15\_2JT15 et MU\_JT20\_L2M0). Cette étude a été réalisée en collaboration avec le groupe  $D\bar{O}$  du LPNHE/Paris 6 et 7. Le programme *top\_trigger* [98] regroupe l'ensemble des paramétrisations des efficacités de déclenchement pour chaque terme aux niveaux 1, 2 et 3 et assure le calcul des poids d'événements par canal d'analyse. Quand le critère de déclenchement change au cours du temps, ce poids est une moyenne des poids de chaque critère utilisé pondérée par la luminosité intégrée correspondant à chaque critère.

#### 4.6.2.2 Efficacité de déclenchement des jets

Les termes de déclenchement sur les jets intervenant dans les critères EM15\_2JT15 et MU\_JT20\_L2M0 sont :

- Niveau 1 : CJT(1,5) (EM15\_2JT15) et CJT(2,5) (MU\_JT20\_L2M0). CJT(n,x) correspond à  $n$  tours du calorimètre avec  $E_T > x$  GeV/c. La probabilité pour un événement de passer CJT(1,5) ou CJT(2,5) est dérivée de la probabilité pour un jet de passer CJT(1,5) par les équations 4.8 et 4.10
- Niveau 2 : JET(2,10) (EM15\_2JT15). JET(n,x) correspond à  $n$  jets reconstruits au niveau 2 avec  $E_T > x$  GeV/c. La probabilité pour un événement de passer JET(2,10) est dérivée de la probabilité pour un jet de passer JET(1,10) par les équations 4.8 et 4.9
- Niveau 3 : JET(2,15,3) (EM15\_2JT15) et JET(1,20,3) (MU\_JT20\_L2M0). JET(n,x,3) correspond à  $n$  jets reconstruits au niveau 3 avec  $E_T > x$  GeV/c.

Il faut donc déterminer les efficacités pour un jet d'avoir déclenché :

- CJT(1,5),
- JET(1,10) après avoir passé CJT(1,5),
- JET(1,15,3) après avoir passé CJT(1,5) et JET(1,10),
- JET(1,20,3) après avoir passé CJT(1,5).

Pour déterminer ces efficacités, un lot de données particulier JETTRIG est sélectionné sur un critère de déclenchement électromagnétique. Les jets utilisés doivent satisfaire les critères de qualité standard décrit au § 2.2.3.5. Chaque jet reconstruit est successivement associé à une tour calorimétrique (niveau 1) et à un jet des niveaux 2 et 3. A chaque niveau l'association est réalisée en minimisant la séparation  $\Delta R$  dans l'espace  $(\eta, \phi)$  avec le jet reconstruit. De plus cette séparation doit vérifier  $\Delta R < 0.5$ .

Le choix d'événements électromagnétiques introduit un biais, car un des jets de l'événement est susceptible d'avoir servi au déclenchement. En conséquence ce lot contient un excès de jets de grande fraction électromagnétique. Pour supprimer ce biais, seuls les événements contenant un électron associé à un objet électromagnétique aux niveaux 1 et 2 du déclenchement sont considérés. La figure 4.12 représente la fraction électromagnétique des jets avant et après cette sélection.

À titre de vérification, les efficacités sont également mesurées sur des événements dé-

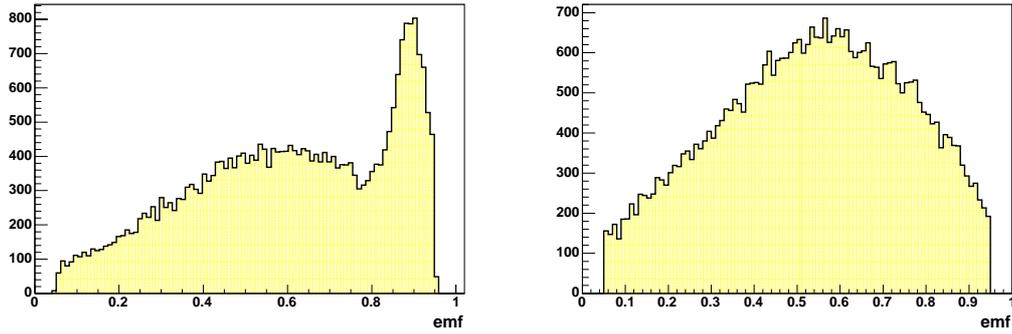


FIG. 4.12: Fraction électromagnétique des jets du lot JETTRIG pour tous les événements (à gauche) et pour les événements avec un électron reconstruit associé à un objet électromagnétique aux niveaux 1 et 2. (à droite). La distribution pour les événements déclenchés sur un véritable électron ne présente plus d'excès pour les grandes fractions électromagnétiques.

clenchés par un muon. Ce lot de données possède un biais plus faible mais plus difficile à quantifier. Le spectre en  $p_T$  des jets peut être affecté si le muon appartient au jet. L'énergie du muon n'est alors pas incluse dans l'énergie du jet. Les efficacités mesurées dans les deux lots de données sont en bon accord.

L'efficacité de CJT(1,5) est déterminée par la fraction de jets associés à une tour du calorimètre au niveau 1 qui satisfont la coupure  $E_T > 5$  GeV/c. Les efficacités aux niveaux 2 et 3 sont déterminées de manière similaire en ne considérant que les jets ayant passé les critères requis aux niveaux précédents. Ces efficacités sont mesurées en fonction de l'impulsion transverse du jet reconstruit, dans trois domaines de pseudorapidité correspondant au calorimètre central ( $|\eta| < 0.8$ ), à la region inter-cryostat (ICR,  $0.8 < |\eta| < 1.5$ ) et aux bouchons du calorimètre ( $|\eta| > 1.5$ ). Les efficacités en fonction de la pseudorapidité sont données sur la figure 4.13. Les lignes rouges séparent les différentes régions considérées. Les paramétrisations obtenues sont données sur les figures 4.14 (CJT(1,5)), 4.6.2.2 (JET(1,10)), 4.16 (JET(1,15,3)) et 4.17 (JET(1,20,3)). La fonction utilisée pour ajuster ces courbes d'efficacité est :

$$f(p_T) = 0.5 \times A_2 \times \left(1 + \text{Erf}\left(\frac{p_T - A_0}{\sqrt{2} \cdot A_1}\right)\right) \quad (4.11)$$

où les  $A_{i=0..2}$  sont les constantes à ajuster.

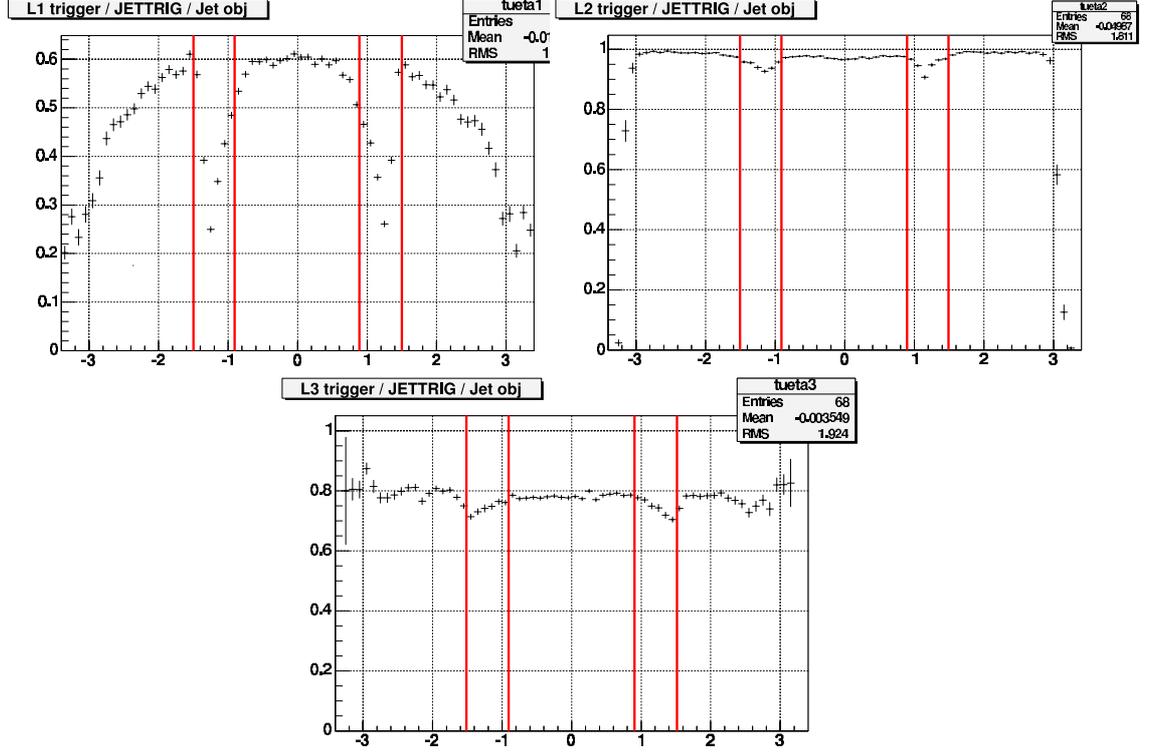


FIG. 4.13: Efficacité moyenne des termes CJT(1,5) (niveau 1), JET(1,10) (niveau 2) et JET(1,15,3) (niveau 3) en fonction de la pseudorapidité du jet reconstruit. Les lignes rouges séparent les trois régions de pseudorapidité utilisées pour les paramétrisations.

### 4.6.3 Utilisation des $TRF$ pour l'identification des jets $b$

#### 4.6.3.1 Probabilité d'étiquetage par événement

La détermination des  $TRF$  pour les différents points de fonctionnement de  $JLIP$  a été détaillée au § 3.5. L'utilisation simultanée de deux points de fonctionnement, implique le calcul du poids par événement. Considérons le cas général de deux points de fonctionnement inclusifs  $A$  et  $B$ , où  $A$  est plus sévère que  $B$ . Inclusif signifie que tout jet satisfaisant  $A$  satisfait également  $B$ . Il existe trois états d'étiquetage possibles pour le jet : (a) étiqueté  $A$ , (b) étiqueté  $B$  mais pas  $A$ , (c) non étiqueté  $B$ . La probabilité de chaque état est donnée par les  $TRF$ . Pour un événement avec  $N$  jets, il existe  $3^N$  permutations possibles de ces états de jet. La probabilité de chacune de ces permutations s'écrit en fonction des probabilités d'étiqueter chaque jet  $A$  ou  $B$ ,  $P_{jet}^A$  et  $P_{jet}^B$  :

$$P_{perm} = \prod_{jet \in (a)} P_{jet}^A \prod_{jet \in (b)} (P_{jet}^B - P_{jet}^A) \prod_{jet \in (c)} (1 - P_{jet}^B) \quad (4.12)$$

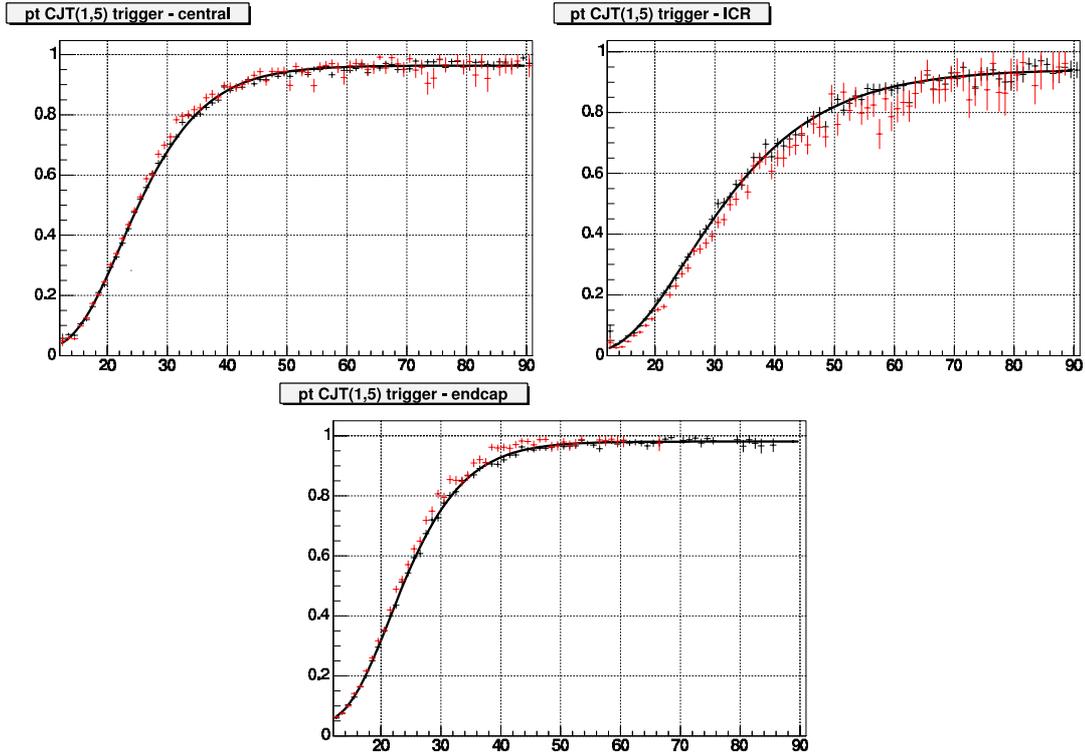


FIG. 4.14: Efficacités du terme de niveau 1 CJT(1,5) en fonction de l'impulsion transverse du jet reconstruit, pour le calorimètre central (à gauche), l'ICR (au centre) et les bouchons (à droite) mesurées avec des événements déclenchés sur un objet électromagnétique (en noir, avec ajustement) et sur un muon (en rouge).

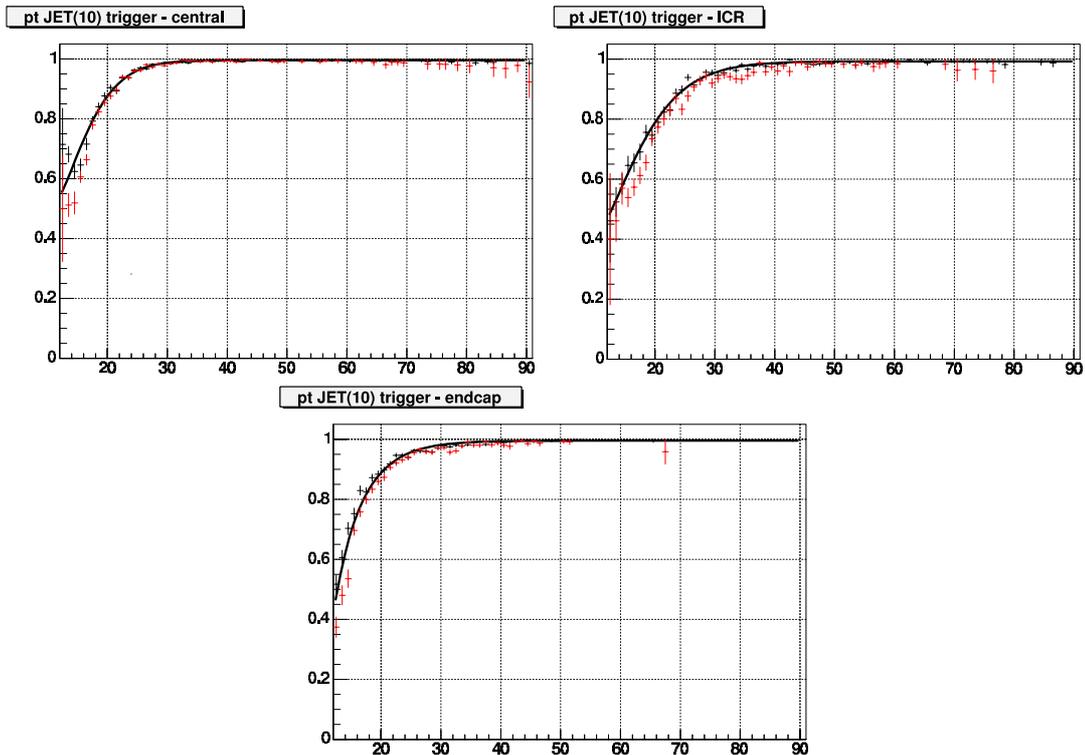


FIG. 4.15: Efficacités du terme de niveau 2 JET(1,10) en fonction de l'impulsion transverse du jet reconstruit, pour le calorimètre central (à gauche), l'ICR (au centre) et les bouchons (à droite) mesurées avec des événements déclenchés sur un objet électromagnétique (en noir, avec ajustement) et sur un muon (en rouge).

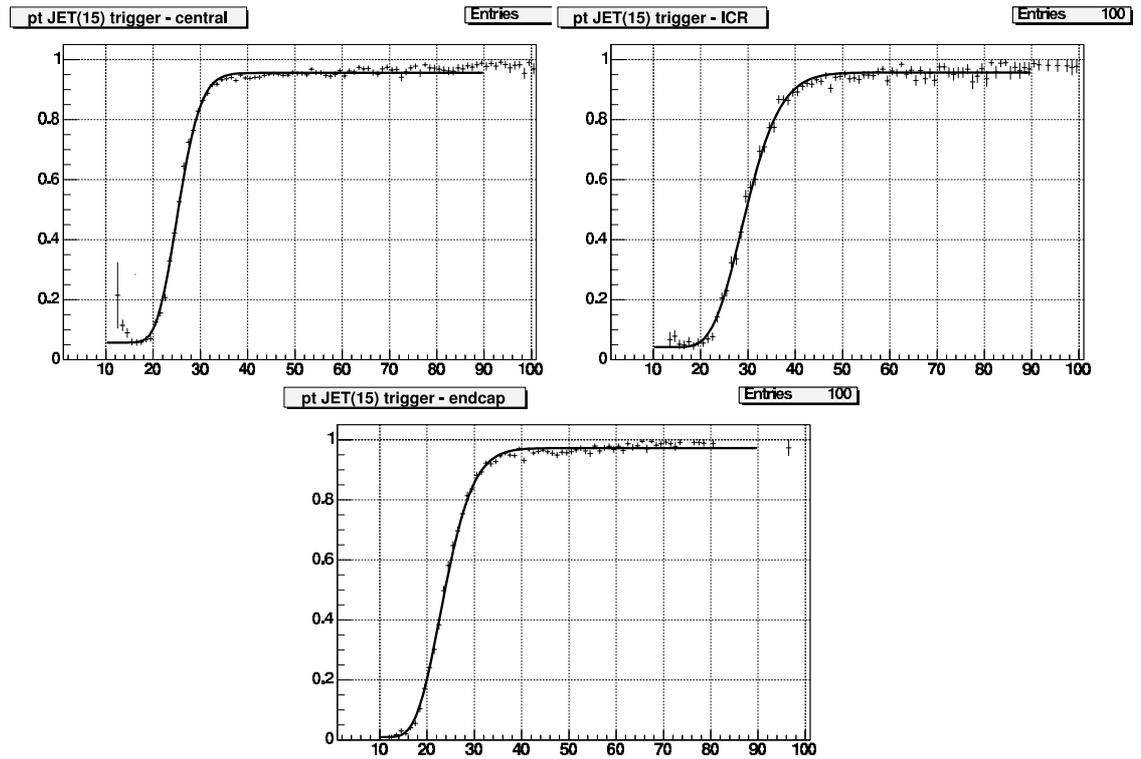


FIG. 4.16: Efficacités du terme de niveau 3 JET(1,15,3) en fonction de l'impulsion transverse du jet reconstruit, pour le calorimètre central (à gauche), l'ICR (au centre) et les bouchons (à droite) mesurées avec des événements déclenchés sur un objet électromagnétique.

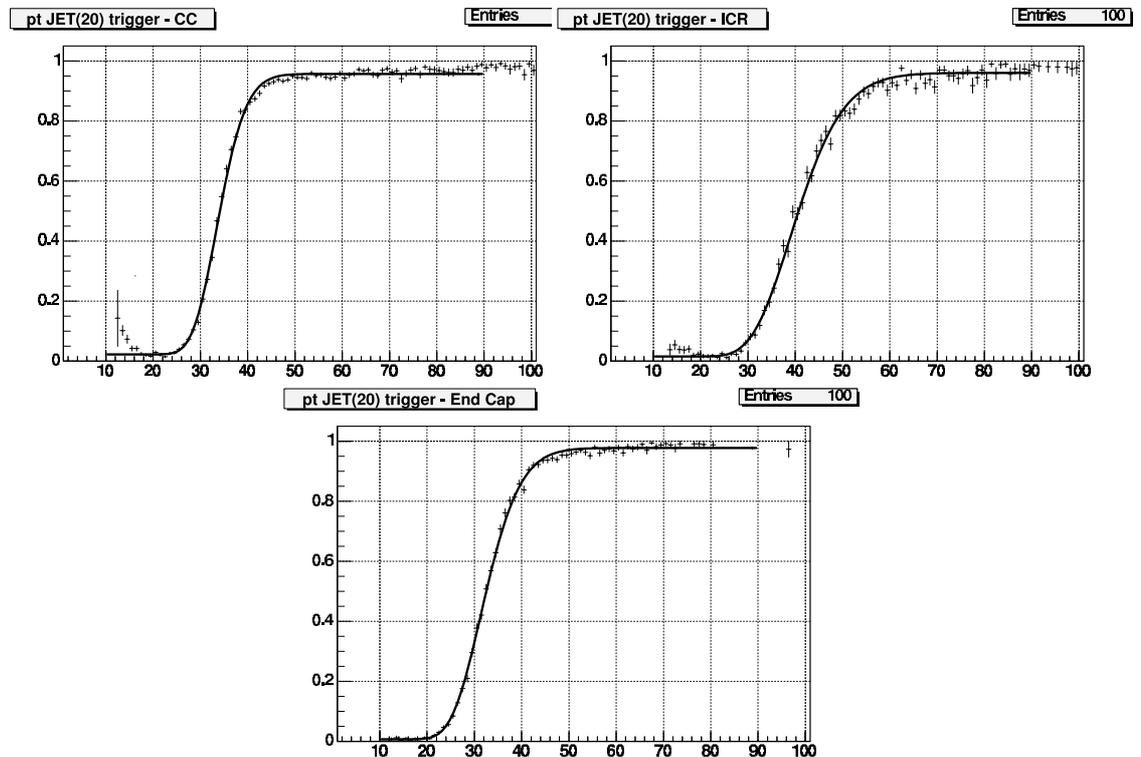


FIG. 4.17: Efficacités du terme de niveau 3 JET(1,20,3) en fonction de l'impulsion transverse du jet reconstruit, pour le calorimètre central (à gauche), l'ICR (au centre) et les bouchons (à droite) mesurées avec des événements déclenchés sur un objet électromagnétique.

Chaque schéma d'étiquetage défini au § 4.5 correspond à un sous-ensemble de ces permutations. La probabilité qu'un événement satisfasse un schéma d'étiquetage  $S$  est alors :

$$P_S = \sum_{perm \in S} P_{perm} \quad (4.13)$$

Dans le cas des schémas STag et DTag où un seul point de fonctionnement est utilisé, les équations 4.12 et 4.13 se simplifient pour retrouver les expressions 4.9 et 4.10.

#### 4.6.3.2 Efficacité de *taggabilité*

Les *TRF* donnent une probabilité d'étiqueter un jet *taggable*. Les probabilités par jet entrant dans le calcul doivent inclure l'efficacité de *taggabilité*. La *taggabilité* est fortement sensible à l'acceptance du détecteur de trace et une dépendance en fonction de la position du vertex primaire est également prise en compte. Les probabilités globales d'étiquetage par jet s'écrivent finalement :

$$P_{jet}^X = P_{taggabilité}(p_T, \eta, PV_z) \times TRF_{saveur=b,c,l}^X(p_T, \eta) \quad (4.14)$$

où  $X$  représente un point de fonctionnement de *JLIP*.

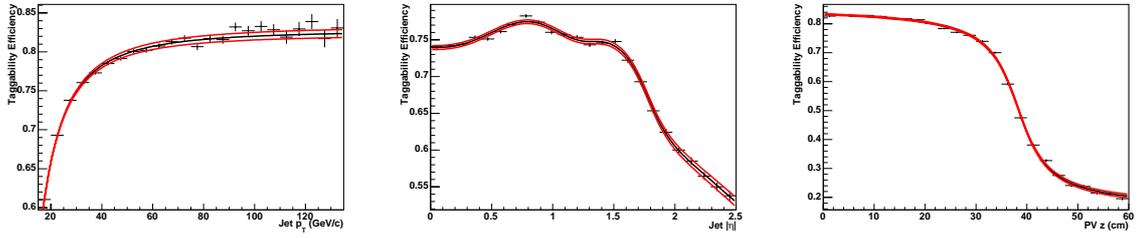


FIG. 4.18: Paramétrisations de la *taggabilité* pour le canal électron en fonction de l'impulsion transverse  $p_T$ , de la pseudorapidité  $\eta$  du jet et la position du vertex primaire  $PV_z$ . Les bandes rouges correspondent aux ajustements à  $\pm 1$  écart standard.

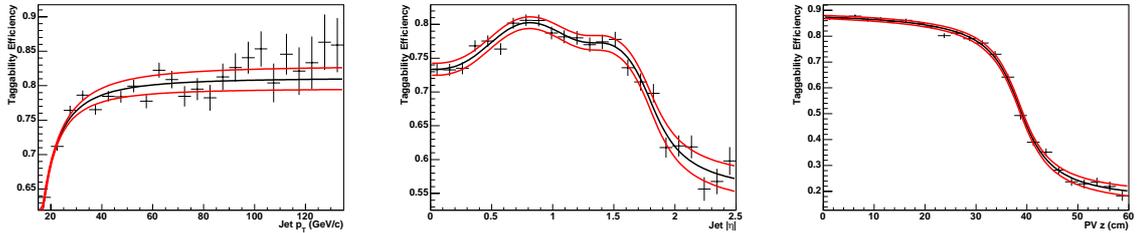


FIG. 4.19: Paramétrisations de la *taggabilité* pour le canal muon en fonction de l'impulsion transverse  $p_T$ , de la pseudorapidité  $\eta$  du jet et la position du vertex primaire  $PV_z$ . Les bandes rouges correspondent aux ajustements à  $\pm 1$  écart standard.

L'efficacité de *taggabilité* peut dépendre du lot de données. Elle est déterminée séparément pour les canaux électron et muon. Pour chaque canal, cette efficacité est déterminée dans les données réelles. Les coupures de sélection décrites au § 4.4 sont appliquées aux données réelles. Pour augmenter le nombre d'événements sélectionnés et pouvoir paramétriser la dépendance en  $p_T$  et  $\eta$  de la *taggabilité*, les coupures sur le jet de plus haut  $p_T$  sont relâchées au même niveau que celles appliquées aux autres jets ( $p_T > 15$  GeV/c,  $|\eta| < 2.5$ ), la *taggabilité* et l'identification des jets  $b$  n'étant valides que pour les jets satisfaisant ces conditions. De plus, la sélection du lepton se limite à un électron ou un muon *loose*. Une paramétrisation à 3 dimensions ( $p_T$ ,  $\eta$ ,  $PV_z$ ) de l'efficacité de *taggabilité* dans chaque canal est déterminée de manière analogue aux *TRF* (voir § 3.5.4). Les figures 4.18 et 4.19 donnent les paramétrisations à une dimension utilisées dans chaque canal d'analyse. Comme pour les *TRF*, deux paramétrisations associées aux erreurs statistiques sont également dérivées. Le lot sélectionnés pour dériver ces *TRF* contiennent respectivement 53648 jets pour le canal muon et 571980 jets pour le canal électron. L'erreur statistique est donc plus grande dans le canal muon. Les formes fonctionnelles sont, ici aussi, choisies afin de s'ajuster au mieux aux données avec un comportement asymptotique approprié (plateaux à grand  $p_T$ ,  $\eta$  et  $PV_z$ ) :

$$f(p_T) = A_0 \times \left( A_1 - \frac{1}{(p_T - A_2)^2} \right) \quad (4.15)$$

$$f(\eta) = B_0 \times \left( B_1 - \text{Arctan}((\eta^2 - B_2^2)^2) + \frac{B_3}{1 + \frac{\eta^4}{B_4}} - B_5 e^{-\frac{|\eta|^3}{B_6}} \right) \quad (4.16)$$

$$f(PV_z) = C_0 \times \left( C_1 - \text{Arctan}\left(\frac{|PV_z| - C_2}{C_3}\right) \right) \quad (4.17)$$

où les  $A_{i=0..2}$ ,  $B_{i=0..6}$  et  $C_{i=0..3}$  sont les constantes d'ajustement.

Les figures 4.20 et 4.21 présentent des histogrammes de contrôles analogues à ceux de la figure 3.26. Ces histogrammes ont été déterminés sur les données sélectionnées pour l'analyse, avant identification des jets  $b$ . Le poids appliqué pour estimer les distributions de la position du vertex primaire et la multiplicité de jets pour des événements contenant exactement un ou au moins deux jets *taggables* sont déterminés avec les équations 4.9 et 4.10. Le bon accord entre pondération et application directe de la *taggabilité* valide ces paramétrisations.

#### 4.6.3.3 Permutations et identification des jets $b$ étiquetés

En pondérant les événements simulés au lieu d'appliquer l'algorithme d'identification, il n'est plus possible de savoir quels jets ont été étiquetés. Cette méthode est parfaitement appropriée pour déterminer le nombre d'événements satisfaisant un schéma d'étiquetage. En revanche, il n'est pas possible a priori d'étudier des variables cinématiques basées sur le ou les jets  $b$  identifiés. Il est néanmoins possible d'estimer de telles variables en considérant séparément chaque permutation de jets étiquetés et non étiquetés. Pour une

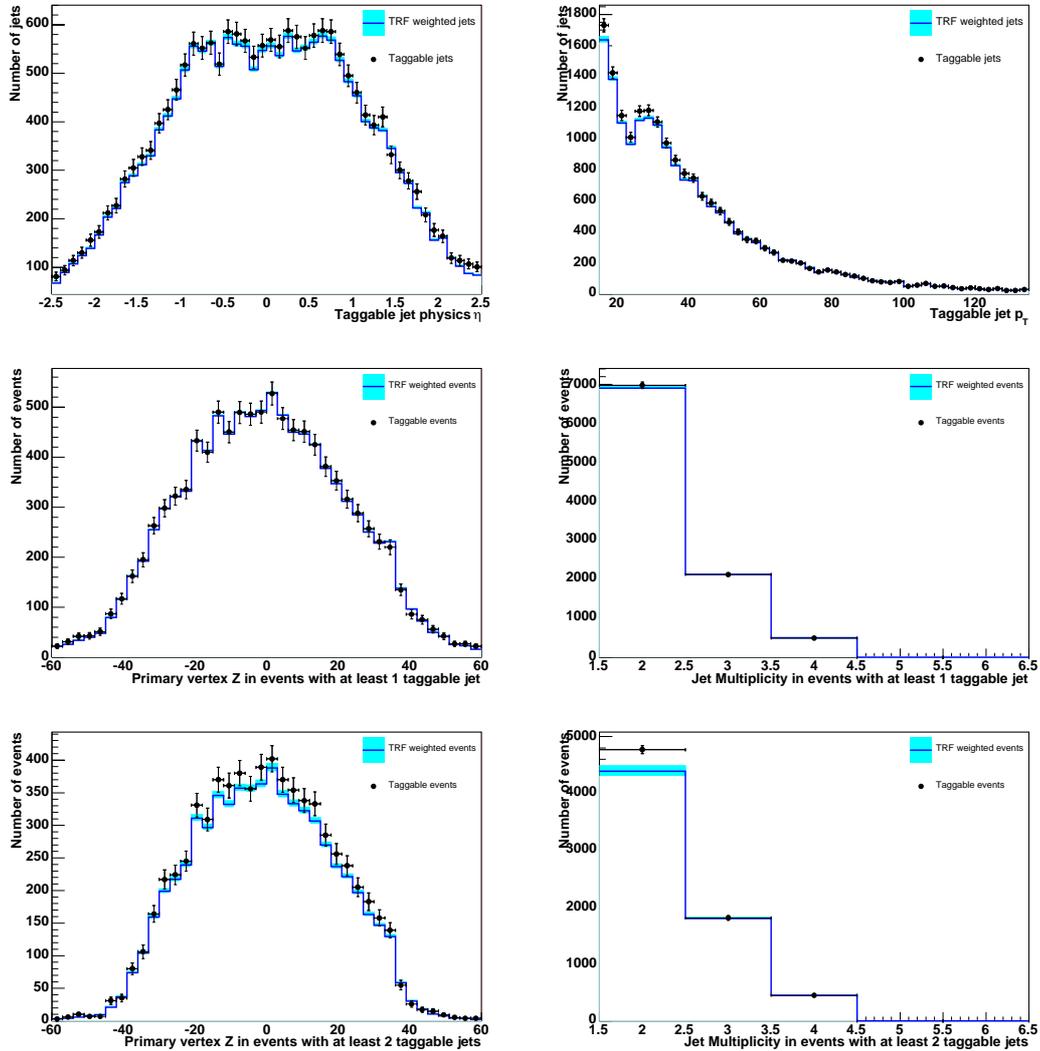


FIG. 4.20: Histogrammes de validation des paramétrisations de la *taggabilité* dans le canal électron.

En haut : impulsion transverse (à gauche) et pseudorapidité (à droite) du jet.

Au centre : position  $z$  du vertex primaire (à gauche) et multiplicité de jets (à droite) dans les événements contenant exactement un jet *taggable*.

En bas : position du vertex primaire (à gauche) et multiplicité de jets (à droite) dans les événements à au moins deux jets *taggables*.

Les points correspondent à l'application directe du critère de *taggabilité*, les lignes aux histogrammes pondérés. Les bandes bleues indiquent l'erreur systématique sur la pondération.

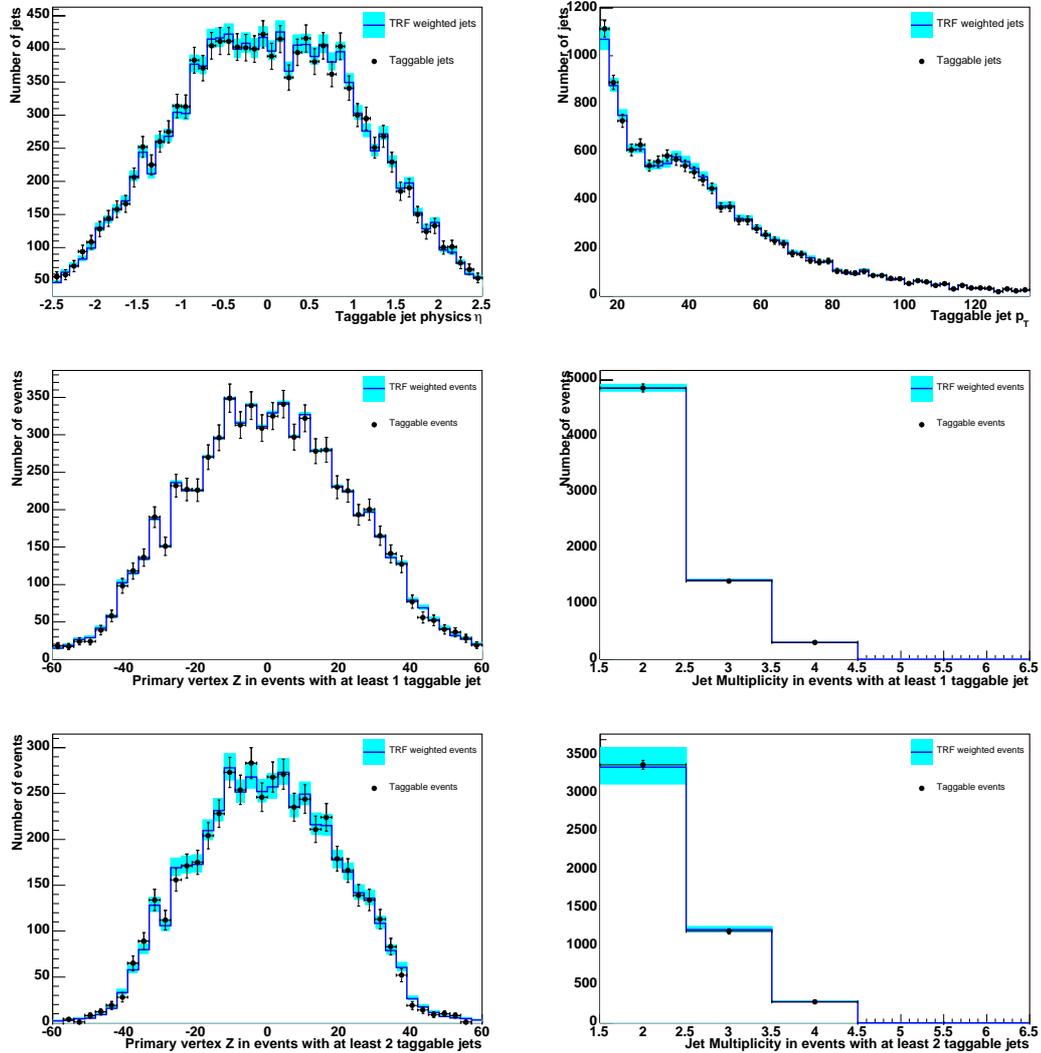


FIG. 4.21: Histogrammes de validation des paramétrisations de la *taggabilité* dans le canal muon.

En haut : impulsion transverse (à gauche) et pseudorapidité (à droite) du jet.

Au centre : position  $z$  du vertex primaire (à gauche) et multiplicité de jets (à droite) dans les événements contenant exactement un jet *taggable*.

En bas : position du vertex primaire (à gauche) et multiplicité de jets (à droite) dans les événements à au moins deux jets *taggables*.

Les points correspondent à l'application directe du critère de *taggabilité*, les lignes aux histogrammes pondérés. Les bandes bleues indiquent l'erreur systématique sur la pondération.

permutation donnée, l'état de chaque jet est connu et les variables cinématiques faisant intervenir l'étiquetage des jets  $b$  sont calculables. Le poids de chaque permutation est donné par l'équation 4.12. L'ensemble des permutations pondérées satisfaisant le schéma d'étiquetage considéré est ainsi substitué à l'événement pondéré de la somme des poids des permutations (équation 4.13). Une variable qui dépend de l'étiquetage des jets  $b$ , au lieu d'avoir une valeur unique pour chaque événement, possède alors une distribution de valeurs, pondérées par la probabilité de la permutation correspondante. Le poids total de l'événement dans l'analyse reste inchangé.

## 4.7 Estimation du signal et des bruits de fond attendus

Il est nécessaire de pouvoir estimer le nombre d'événements de bruit de fond et de signal attendus dans les données après application de toutes les coupures de sélection. En plus du nombre d'événements, il faut pouvoir reproduire les distributions de variables cinématiques. Ces distributions permettent de vérifier la validité de notre modélisation des bruits de fond et serviront dans l'étape suivante de l'analyse à construire des variables discriminantes. Elles sont produites sous formes d'histogrammes. Les méthodes d'estimation des nombres d'événements qui seront présentées dans cette section s'appliquent alors à chaque cellule d'un histogramme. L'estimation de chaque distribution comporte deux aspects, la forme de la distribution et sa normalisation, c'est à dire le nombre total d'événements. Les formes des distributions sont généralement données par les événements simulés, à l'exception du fond multijet. Les distributions simulées sont corrigées par les pondérations de déclenchement et d'identification des jets  $b$ , qui dépendent de la topologie de chaque événement. Ces pondérations affectent à la fois la forme et la normalisation. La normalisation peut se faire soit à partir de la section efficace théorique du processus simulé, soit directement avec les données réelles. Nous allons détailler les différents cas de figure apparaissant dans cette analyse.

### 4.7.1 Acceptances des lots simulés

L'acceptance d'un signal ou d'un bruit de fond est son efficacité totale de sélection. Pour chaque processus simulé, elle est définie par :

$$\mathcal{A}^{\text{MC}} = \frac{\mathcal{B}}{N_{\text{initial}}} \varepsilon_{\text{correction}} \sum_{N_{\text{select}}} P_{\text{déclenchement}} P_{\text{étiquetage}} \quad (4.18)$$

où :

- $\mathcal{B}$  est le rapport d'embranchement du lot Monte Carlo considéré, s'il ne contient pas l'intégralité des modes de désintégration du processus physique simulé.
- $N_{\text{initial}}$  est le nombre total d'événements simulés.

- $N_{\text{select}}$  est le nombre d'événements sélectionnés. Le déclenchement et l'identification des jets  $b$  ne sont pas appliqués dans la simulation.
- $\varepsilon_{\text{correction}}$  est le produit des facteurs correctifs entre données réelles et simulation décrits au § 4.6.1 :
  - $\varepsilon_{\text{correction}} = \varepsilon_{\text{vertex}} \times \varepsilon_{e\text{-ID}} = 0.893 \pm 0.032$  (canal électron).
  - $\varepsilon_{\text{correction}} = \varepsilon_{\text{vertex}} \times \varepsilon_{\mu\text{-ID}} = 0.944 \pm 0.050$  (canal muon).
- $P_{\text{déclenchement}}$  est le poids de déclenchement décrit au § 4.6.2.
- $P_{\text{étiquetage}}$  est le poids d'identification des jets  $b$  correspondant à un des schémas décrits au § 4.6.3.

Les acceptances pour le signal et les fonds simulés, dans les différents schémas d'identification des jets  $b$  sont résumées dans les tableaux 4.5 pour le canal électron et 4.6 pour le canal muon.

## 4.7.2 Normalisation sur les sections efficaces théoriques

Pour les processus simulés dont la section efficace théorique  $\sigma$  est connue avec une bonne précision (signal,  $t\bar{t}$ ,  $WW$ ,  $WZ$ ), le nombre total d'événements produits s'écrit :

$$N^{\text{total}} = \mathcal{L} \times \sigma \quad (4.19)$$

où  $\mathcal{L}$  est la luminosité intégrée. Le nombre d'événements attendus après sélection  $\mathcal{Y}$  est alors simplement :

$$\mathcal{Y} = \mathcal{A} \times N^{\text{total}} = \mathcal{A} \times \mathcal{L} \times \sigma \quad (4.20)$$

## 4.7.3 Normalisation à l'aide des données réelles

Les bruits de fond  $W$ +jets et multijet sont normalisés à l'aide des données avant application des schémas d'identification des jets  $b$ . L'estimation du fond  $W$ +jets repose en partie sur la simulation alors que le fond multijet est entièrement évalué avec des données réelles.

### 4.7.3.1 Mesure du nombre de vrais leptons isolés

Avant application d'un schéma d'étiquetage, les données sélectionnées sont dominées par les fonds  $W$ +jets et multijet. Le bruit de fond multijet est caractérisé par un "faux" lepton isolé, c'est-à-dire que l'objet qui est sélectionné en tant que lepton isolé, bien que passant toutes les coupures de sélection, n'en est pas un. Le nombre de vrais  $N_{\text{vrai}}$  et de faux  $N_{\text{faux}}$  leptons isolés peut être déterminé dans les données en relâchant une coupure de qualité du lepton : la coupure sur la fonction de vraisemblance dans le canal électron et la sélection sur les halos (traces et calorimètre) dans le canal muon (cf. § 4.4.3). Les nombres d'événements sélectionnés avec les coupures standard ( $N_s$ ) et les coupures

Canal électron					
Schéma d'étiquetage	Aucun	S <sub>Tag</sub>	D <sub>Tag</sub>	S <sub>Tag</sub> Veto	D <sub>Tag</sub> Asym
Signal					
<i>tb</i>	2.7	1.2	0.3	1.0	0.5
<i>tqb</i>	2.7	1.0	0.09	1.0	0.1
Bruits de fond					
<i>t<math>\bar{t}</math>l</i> +jets	4.3	1.9	0.5	1.6	0.8
<i>t<math>\bar{t}</math>ll</i>	1.2	0.5	0.2	0.5	0.2
<i>Wbb</i>	5.5	2.2	0.4	2.0	0.7
<i>Wjj</i>	5.0	1.1	<0.01	1.1	<0.01
<i>WW</i>	1.6	0.07	<0.01	0.07	<0.01
<i>WZ</i>	1.6	0.2	0.05	0.2	0.07

TAB. 4.5: Acceptances (%) pour le signal et les fonds simulés dans le canal électron. Les poids d'identification des jets *b* valent 1 quand aucun schéma d'étiquetage n'est appliqué. Les schémas d'étiquetage sont décrits au § 4.5.

Canal muon					
Schéma d'étiquetage	Aucun	S <sub>Tag</sub>	D <sub>Tag</sub>	S <sub>Tag</sub> Veto	D <sub>Tag</sub> Asym
Signal					
<i>tb</i>	2.4	1.1	0.3	0.9	0.5
<i>tqb</i>	2.4	0.9	0.08	0.9	0.1
Bruits de fond					
<i>t<math>\bar{t}</math>l</i> +jets	4.1	1.8	0.6	1.5	0.8
<i>t<math>\bar{t}</math>ll</i>	1.1	0.5	0.1	0.4	0.2
<i>Wbb</i>	3.1	1.4	0.3	1.2	0.4
<i>Wjj</i>	3.1	0.7	<0.01	0.7	<0.01
<i>WW</i>	1.6	0.07	<0.01	0.07	<0.01
<i>WZ</i>	1.7	0.2	0.04	0.2	0.07

TAB. 4.6: Acceptances (%) pour le signal et les fonds simulés dans le canal muon. Les poids d'identification des jets *b* valent 1 quand aucun schéma d'étiquetage n'est appliqué.

relâchées ( $N$ ) peuvent alors s'écrire :

$$N = N_{faux} + N_{vrai} \quad (4.21)$$

$$N^s = N_{vrai}^s + N_{faux}^s = \varepsilon_{faux} \cdot N_{faux} + \varepsilon_{vrai} \cdot N_{vrai} \quad (4.22)$$

où  $\varepsilon_{vrai}$  ( $\varepsilon_{faux}$ ) est l'efficacité de sélection de la coupure de qualité leptonique sur les vrais (faux) leptons isolés. La résolution de ce système de deux équations à deux inconnues donne le nombre de vrais et de faux leptons isolés dans les données avant étiquetage des jets  $b$ . Cette procédure est appelée Méthode de la Matrice :

$$N_{faux} = \frac{N^s - \varepsilon_{faux} \cdot N}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (4.23)$$

$$N_{vrai} = \frac{\varepsilon_{vrai} \cdot N - N^s}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (4.24)$$

$$N_{faux}^s = \varepsilon_{faux} \frac{N^s - \varepsilon_{faux} \cdot N}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (4.25)$$

$$N_{vrai}^s = \varepsilon_{vrai} \frac{\varepsilon_{vrai} \cdot N - N^s}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (4.26)$$

L'efficacité de sélection d'un vrai lepton,  $\varepsilon_{vrai}$ , est déterminée au pic du  $Z$  de manière analogue aux efficacités décrites au § 4.6.1. Ces efficacités dépendent de la multiplicité de jets de l'événement : en effet, plus un événement a une multiplicité de jets élevée, plus l'efficacité de reconstruire un lepton isolé de tout jet diminue. Les efficacités utilisées pour cette analyse sont données dans le tableau 4.7.

L'efficacité de sélection un "faux" lepton,  $\varepsilon_{faux}$ , est estimée sur un lot enrichi en événements multijet. Ceci est réalisé en inversant la coupure sur l'énergie manquante ( $E_T < 10$  GeV/c) : la majorité des événements ainsi sélectionnés possède un lepton reconstruit passant toutes les coupures de sélection mais pas de neutrino. Dans le canal électron, cette efficacité dépend du critère de déclenchement. Cette dépendance est faible dans le canal muon, en revanche, l'efficacité de reconstruire un "faux" muon isolé est sensible à la géométrie du détecteur. Dans la région inter-cryostat (ICR), les jets rencontrent moins de matière et des pions peuvent sortir du calorimètre et traverser les chambres à muons. Les coupures sur les halos sont très efficaces pour rejeter ce type de "faux" muons, ce qui diminue fortement le taux de "faux" muon isolé dans cette région. Une paramétrisation de cette efficacité en fonction de la pseudorapidité du muon présentée sur la figure 4.22, est utilisée. Les taux d'identification d'un faux lepton isolé sont donnés dans le tableau 4.8. A l'exception de la paramétrisation en pseudorapidité, toutes ces efficacités ont été dérivées par le groupe de travail "top" de DØ et sont utilisées dans d'autres analyses.

Pour utiliser des efficacités variables dans le système, la méthode de la matrice est appliquée événement par événement avec  $N_i^s = 0$  ou 1. Le système étant linéaire, les

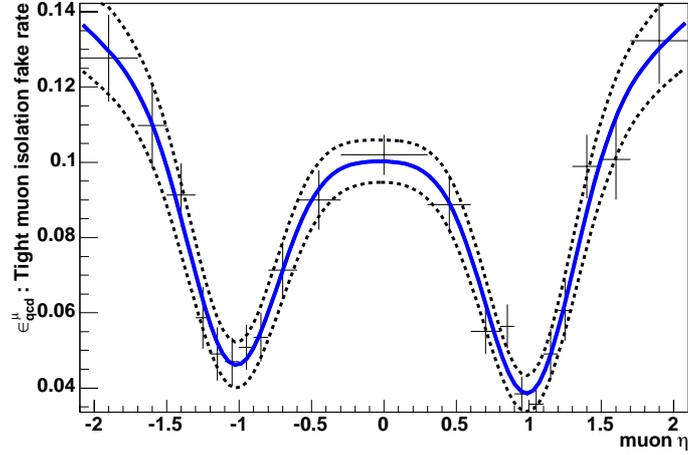


FIG. 4.22: Paramétrisation du taux d'identification d'un faux muon isolé ( $\varepsilon_{faux}$ ) en fonction de la pseudorapidité du muon. Les bandes bleues indiquent l'incertitude statistique sur ce taux. Pour l'analyse une erreur relative moyenne de 4.8 % est utilisée.

Multiplicité de jets	2 jets	3 jets	4 jets
électron	$0.843 \pm 0.004$	$0.836 \pm 0.013$	$0.812 \pm 0.006$
muon	$0.858 \pm 0.013$	$0.857 \pm 0.017$	$0.827 \pm 0.005$

TAB. 4.7: Efficacité d'identification d'un vrai lepton isolé,  $\varepsilon_{vrai}$ .

Liste de déclenchement	v8-11	v12	v13
électron	$0.118 \pm 0.017$	$0.187 \pm 0.016$	$0.217 \pm 0.058$
muon	Paramétrisation en fonction de $\eta$ , erreur relative 4.8%		

TAB. 4.8: Taux d'identification d'un faux lepton isolé,  $\varepsilon_{faux}$ .

solutions par événement de :

$$1 = n_{faux}^i + n_{vrai}^i \quad (4.27)$$

$$N^{s,i} = \varepsilon_{faux}(\eta, liste).n_{faux}^i + \varepsilon_{vrai}(N_{jet}).n_{vrai}^i \quad (4.28)$$

sont sommées pour obtenir le résultat final :

$$N_{faux} = \sum_{i=1}^N n_{faux}^i \quad (4.29)$$

$$N_{vrai} = \sum_{i=1}^N n_{vrai}^i \quad (4.30)$$

$$N_{faux}^s = \sum_{i=1}^N \varepsilon_{faux}(\eta, liste).n_{faux}^i \quad (4.31)$$

$$N_{vrai}^s = \sum_{i=1}^N \varepsilon_{vrai}(N_{jet}).n_{vrai}^i \quad (4.32)$$

Une solution par événement peut être négative mais la solution globale ne doit pas l'être si les efficacités ont été correctement déterminées. Le tableau 4.9 présente les résultats obtenus dans les canaux électron et muon. La figure 4.23 montre les distributions de masse transverse du boson  $W$  ( $m_T^{W2} = (p_x^{lepton} + \cancel{E}_{Tx})^2 + (p_y^{lepton} + \cancel{E}_{Ty})^2 - (p_T^{lepton} + \cancel{E}_T)^2$ ) dans les données avec et sans relâchement des coupures d'isolation. Cette variable est très discriminante entre les fonds  $W$ +jet et multijet et permet de vérifier la validité de la méthode de la matrice.

	$N$	$N^s$	$N_{faux}$	$N_{faux}^s$	$N_{vrai}$	$N_{vrai}^s$
électron	10986	6388	$4158.5 \pm 163.0$	$653.2 \pm 103.5$	$6827.5 \pm 171.0$	$5734.8 \pm 132.5$
muon	7597	3913	$3285.8 \pm 426.2$	$219.0 \pm 30.6$	$4311.2 \pm 427.4$	$3694.0 \pm 70.3$

TAB. 4.9: Nombres d'événements contenant un vrai ou un faux lepton obtenus avec la méthode de la matrice.

#### 4.7.3.2 Le bruit de fond $W$ +jets

Le bruit de fond  $W$ +jet étant le fond dominant, il doit être estimé avec la meilleure précision possible. Les incertitudes théoriques (20 à 30%) sur les sections efficaces, notamment avec des jets de saveurs lourdes ne permettent pas d'avoir la précision nécessaire. De plus, ces sections efficaces sont fortement sensibles aux coupures partoniques utilisées lors de la génération. Pour cette raison on choisit d'estimer ce bruit de fond à partir des données réelles. Deux procédures ont été mises en œuvre au cours de cette

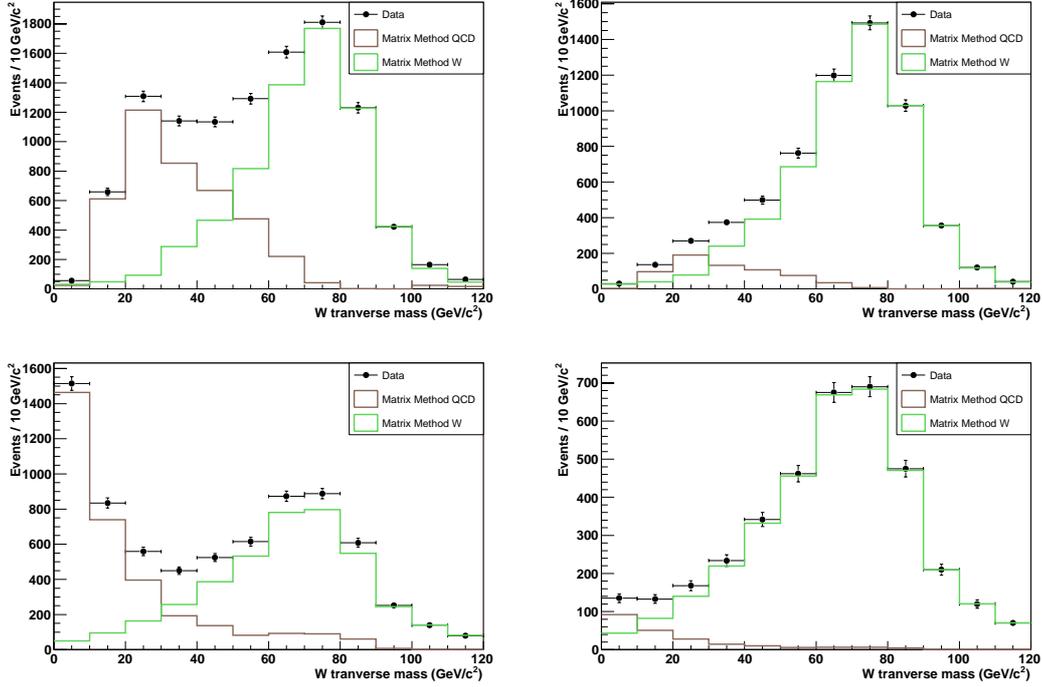


FIG. 4.23: Décomposition de la masse transverse du boson  $W$  en composantes "vrai" (en vert) et "faux" leptons isolés obtenue avec la méthode de la matrice, dans les données sélectionnées sans (à gauche) et avec (à droite) les coupures d'isolation du lepton dans les canaux électron (en haut) et muon (en bas).

thèse. La première, basée entièrement sur les données a été utilisée dans la première analyse *single top* réalisée à  $D\bar{O}$  [99]. La seconde utilise des événements simulés  $W$ +jets pour obtenir les formes des distributions qui sont normalisées sur les données réelles. C'est cette seconde méthode qui est retenue pour l'analyse décrite dans cette thèse. Dans les deux cas la méthode de la matrice est utilisée pour la normalisation.

#### a. Estimation du fond $W$ +jets dans les données et TRF inclusive

Dans la première analyse [99] un seul schéma d'étiquetage est utilisé : chaque événement doit contenir au moins un jet étiqueté *Tight*. Les données sélectionnées sont principalement constituées d'événements  $W$ +jets. Les événements non étiquetés sont utilisés pour estimer les formes des distributions pour le fond  $W$ +jets. Ces distributions sont ensuite normalisées au nombre de vrais leptons isolés estimés par la méthode de la matrice.

Pour reproduire l'étiquetage des jets  $b$ , une *TRF* inclusive en saveur est dérivée à partir

des données multijet (2 à 4 jets). Cette  $TRF$  est utilisée pour déterminer un poids d'événement suivant la formule 4.8. Ceci suppose que le contenu en saveur ( $b$ ,  $c$ , légers) est similaire dans les données multijet et dans les événements  $W$ +jets sélectionnés. Cette hypothèse avait été validée au Run I [100], en appliquant une  $TRF$  similaire (pour l'identification des jets  $b$  par présence d'un muon dans le jet) à différents lots de données simulées (multijet,  $W$ +jets,  $Z$ +jets,  $\gamma$ +jets) produits avec le générateur HERWIG. Cette étude ne montre pas de déviation supérieure à 10 % entre les nombres d'événements observés et estimés avec la  $TRF$ . Ces déviations sont illustrées par la figure 4.24. Finalement, une erreur systématique conservatrice de 20 % est affectée à cette procédure dans l'analyse *single top* du Run II.

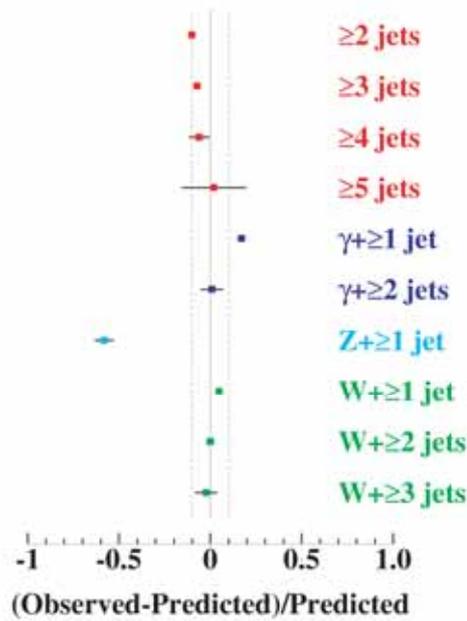


FIG. 4.24: Écart relatif entre l'application d'une  $TRF$  inclusive et de l'identification des jets  $b$  par recherche d'un muon dans le jet au Run I pour différents lots d'événements simulés.

La  $TRF$  inclusive pour *JLIP Tight* est donnée sur la figure 4.25. La figure 4.26 présente les histogrammes de contrôles de cette  $TRF$ .

La limitation principale de cette méthode est l'impossibilité d'étiqueter doublement les événements car la  $TRF$  est inclusive en saveur. En effet, la probabilité d'étiqueter un second jet est plus élevée que celle d'étiqueter le premier car les quarks  $b$  sont souvent produits par paires. Les schémas d'étiquetage asymétriques sont également inutilisables. C'est pour ces raisons, que la seconde méthode d'estimation du fond  $W$ +jets, partiellement basée sur la simulation a été développée.

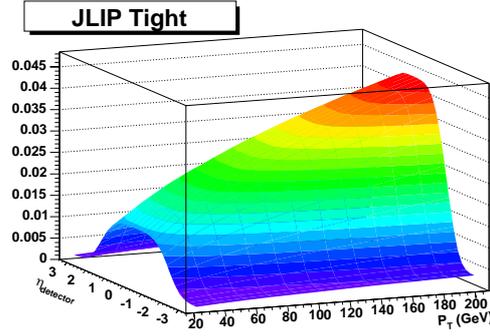


FIG. 4.25:  $TRF$  inclusive pour  $JLIP\ Tight$  dérivée sur des données multijet, en fonction de l'impulsion transverse et de la pseudorapidité du jet.

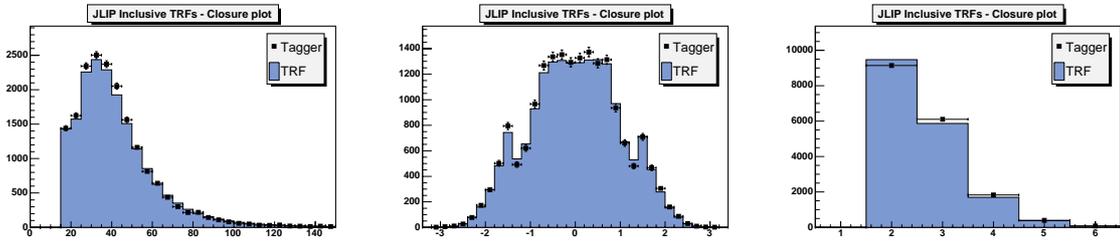


FIG. 4.26: Histogrammes de contrôle de la  $TRF$  inclusive  $JLIP\ Tight$  pour l'impulsion transverse (à gauche), la pseudorapidité (à droite) du jet et la multiplicité de jets des événements à un jet étiqueté. Les points représentent les données étiquetées, les histogrammes pleins les événements pondérés par la  $TRF$  inclusive.

### b. Estimation du fond $W$ +jets dans la simulation

La méthode d'estimation des bruits de fond  $W$ +jets utilisée dans notre analyse utilise les événements simulés  $Wjj$  et  $Wbb$ . Le nombre total d'événements  $W$ +jets avant application d'un schéma d'identification des jets  $b$  est estimé par la méthode de la matrice. Le nombre d'événements avec un vrai lepton isolé obtenus par la méthode de la matrice est corrigé des nombres d'événements  $t\bar{t}$ ,  $WW$  et  $WZ$ . Ces nombres d'événements sont évalués à partir des acceptances (sans étiquetage des jets  $b$ ) et des sections efficaces théoriques (c.f. tableau 4.1). Le rapport des sections efficaces théoriques de production  $Wbb$  et  $Wjj$  sert à la normalisation relative des lots simulés  $Wjj$  et  $Wbb$ . Les poids d'étiquetage sont ensuite appliqués pour reproduire l'identification des jets  $b$ . L'effet du déclenchement sur la normalisation des distributions est pris en compte implicitement puisque cette normalisation se fait sur les données réelles. Il faut néanmoins inclure l'effet du déclenchement sur la forme des distributions. Les poids de déclenchement sont

donc appliqués à chaque événement et la normalisation est corrigée de la valeur moyenne de ces poids. Finalement les nombres d'événements  $Wjj$  et  $Wbb$  attendus  $\mathcal{Y}_{Wjj,Wbb}^{\text{MC}}$  sont donnés par :

$$\mathcal{Y}_{Wjj}^{\text{MC}} = S_{Wjj} \cdot (N_{\text{vrai}}^s - N_{t\bar{t}} - N_{WW} - N_{WZ}) \cdot \frac{\sum_{N_{Wjj}^{\text{select}}} P_{\text{déclenchement}} P_{\text{étiquetage}}}{\sum_{N_{Wjj}^{\text{select}}} P_{\text{déclenchement}}} \quad (4.33)$$

$$\mathcal{Y}_{Wbb}^{\text{MC}} = S_{Wbb} \cdot (N_{\text{vrai}}^s - N_{t\bar{t}} - N_{WW} - N_{WZ}) \cdot \frac{\sum_{N_{Wbb}^{\text{select}}} P_{\text{déclenchement}} P_{\text{étiquetage}}}{\sum_{N_{Wbb}^{\text{select}}} P_{\text{déclenchement}}} \quad (4.34)$$

avec :

$$S_{Wjj} = \frac{\tilde{\sigma}_{Wjj}}{\tilde{\sigma}_{Wjj} + \tilde{\sigma}_{Wbb}} \quad \text{et} \quad S_{Wbb} = \frac{\tilde{\sigma}_{Wbb}}{\tilde{\sigma}_{Wjj} + \tilde{\sigma}_{Wbb}} \quad (4.35)$$

$$\text{où } \tilde{\sigma} = \frac{\sigma}{N_{\text{initial}}} \sum_{N_{\text{select}}} P_{\text{déclenchement}} \quad (4.36)$$

- $N_{\text{initial}}$  est le nombre total d'événements  $Wjj$  ou  $Wbb$  simulés,
- $N_{\text{select}}$  est le nombre d'événements  $Wjj$  ou  $Wbb$  simulés sélectionnés, sans application de facteurs correctifs,
- $S_{Wjj}$  et  $S_{Wbb}$  sont les fractions relatives d'événements  $Wjj$  et  $Wbb$  sélectionnés. Ces fractions incluent les sections efficaces théoriques ( $\sigma_{Wjj}$ ,  $\sigma_{Wbb}$ ) et l'efficacité de sélection (déclenchement inclus) de chacun des processus. Les sections efficaces  $NLO$   $\sigma_{Wjj} = 287$  pb et  $\sigma_{Wbb} = 3.35$  pb [79] sont calculées avec le générateur MCFM [88]. Les facteurs correctifs constants entre données et simulation qui devraient apparaître au numérateur et au dénominateur se simplifient.
- $N_{\text{vrai}}^s$  est le nombre d'événements contenant un vrai lepton dans les données sélectionnées avant identification des jets  $b$  obtenu par la méthode de la matrice.
- $N_{t\bar{t},WW,WZ} = \mathcal{A}_{t\bar{t},WW,WZ}^{\text{notag}} \times \mathcal{L} \times \sigma_{t\bar{t},WW,WZ}$  est le nombre d'événements  $t\bar{t}$ ,  $WW$  et  $WZ$  avant application de l'étiquetage des jets  $b$ .  $\mathcal{A}^{\text{notag}}$  est l'acceptance sans étiquetage.
- $P_{\text{déclenchement}}$  est le poids de déclenchement de l'événement, décrit au § 4.6.2.
- $P_{\text{étiquetage}}$  est le poids d'identification des jets  $b$  de l'événement correspondant à un des schémas décrits au § 4.6.3.

L'utilisation combinée de la simulation et des données dans l'estimation du fond  $W + jet$  a un double avantage. Les événements simulés permettent l'utilisation des poids d'étiquetage ainsi que des schémas d'étiquetage complexes (double étiquetage, étiquetage

asymétrique). La normalisation sur les données réelles réduit la sensibilité aux incertitudes théoriques des sections efficaces  $W$ +jets, qui n'interviennent que sous forme de rapport de sections efficaces.

### 4.7.3.3 Le bruit de fond multijet

Le bruit de fond multijet où l'identification d'un lepton isolé est erronée est entièrement estimé sur les données réelles. De tels événements improprement reconstruits correspondent à des queues de distribution et ne peuvent être proprement simulés. Une sélection spécifique est utilisée pour définir des lots de données enrichis en événements multijet. Ces lots sont statistiquement indépendants des lots d'analyse. Les différences par rapport aux coupures de sélection standard sont :

- dans le canal électron, l'association de l'électron à une trace n'est pas requise et la coupure sur la fonction de vraisemblance est inversée :  $\mathcal{L}_{EM} < 0.05$ ,
- dans le canal muon, les coupures sur les deux halos (traces et calorimètre) sont inversées :  $Rattrk > 0.06$  et  $Rat11 > 0.08$

Ces coupures sont illustrées dans les données sélectionnées sur la figure 4.27. Les schémas d'étiquetage des jets  $b$  sont appliqués comme dans les données réelles sélectionnées. Ces lots de données multijet permettent d'obtenir les formes des distributions. La normalisation avant identification des jets  $b$  est donnée par la méthode de la matrice. Le nombre d'événements multijet attendus après application d'un schéma d'étiquetage est donné par :

$$\mathcal{Y}_{Multijet} = \frac{N_{faux}^s}{N_{Multijet}} N_{étiqueté}^{Multijet} \quad (4.37)$$

avec :

- $N^{Multijet}$  est le nombre d'événements dans le lot multijet sans étiquetage des jets  $b$ .
- $N_{étiqueté}^{Multijet}$  est le nombre d'événements étiquetés dans le lot multijet.
- $\varepsilon_{faux} N_{faux}$  est le nombre d'événements avec un "faux" lepton isolé dans les données sélectionnées obtenu par la méthode de la matrice avant identification des jets  $b$ .

	canal électron		canal muon	
	avant	après	avant	après
STag	$22.6 \pm 3.2$	$22.6 \pm 4.4$	$19.1 \pm 4.9$	$12.6 \pm 1.8$
DTag	$0.61 \pm 0.18$	$0.63 \pm 0.62$	$1.46 \pm 0.25$	$0.11 \pm 0.19$
STagVeto	$21.8 \pm 3.7$	$21.4 \pm 4.2$	$17.9 \pm 4.5$	$12.4 \pm 1.65$
DTagAsym	$1.41 \pm 0.33$	$1.79 \pm 0.90$	$2.67 \pm 0.43$	$0.32 \pm 0.29$

TAB. 4.10: Nombres d'événements multijet estimés dans chaque canal d'analyse pour chaque schéma d'étiquetage en appliquant la méthode de la matrice avant et après identification des jets  $b$ .

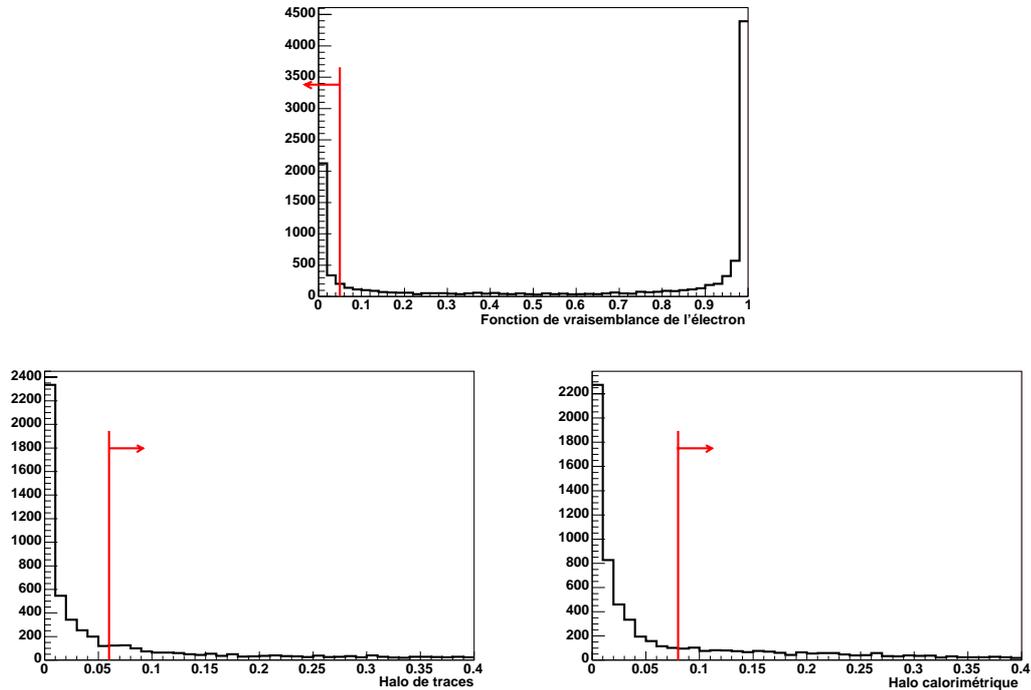


FIG. 4.27: Coupures inversées pour la sélection des lots multijet, dans les données sélectionnées. En haut : Fonction de vraisemblance de l'électron, pour le canal électron. En bas : Halo de traces (à gauche) et halo calorimétrique (à droite), pour le canal muon. Les flèches rouges indiquent les coupures de sélection inversées.

L'utilisation de la méthode de la matrice après application du schéma d'étiquetage permet également de mesurer le fond multijet résiduel. Néanmoins cette approche souffre d'une grande erreur statistique. De plus, la statistique disponible ne permet pas de mesurer les efficacités d'identification d'un vrai ou d'un faux lepton isolé après sélection des jets  $b$ . Les valeurs des tableaux 4.7 et 4.8 sont utilisées. Ces efficacités peuvent être erronées, en particulier dans le canal muon où les "faux" muons isolés peuvent provenir de jets  $b$ . Cette seconde méthode d'estimation du fond multijet sert de vérification à la première.

Le tableau 4.10 compare les taux d'événements attendus avec les deux méthodes. L'accord est excellent dans le canal électron. Les différences observées dans le canal muon s'expliquent par l'utilisation d'efficacités inappropriées.

### 4.7.4 Nombres d'événements attendus après sélection

Les tableaux 4.11 (canal électron) et 4.12 (canal muon) résument les nombres d'événements sélectionnés attendus pour le signal et les bruits de fond avant et après identification des jets  $b$ , pour chaque schéma d'étiquetage.

Après application du schéma d'étiquetage STag (DTag), les fonds  $W$ +jets représentent environ 58% (25%) du bruit de fond total soit 47% (8%) de  $Wjj$  et 11% (17%) de  $Wbb$ . L'autre contribution principale au bruit de fond est la production de paires  $t\bar{t}$  qui représente 29% (69%) du total soit 23% (54%) de lepton+jets et 6% (15%) de dilepton.

## 4.8 Analyse multivariable

### 4.8.1 Variables discriminantes

Après sélection des événements, les distributions de nombreuses variables cinématiques discriminantes sont étudiées. Les variables qui sont présentées par la suite sont sélectionnées parmi les variables utilisées pour la construction des réseaux de neurones de l'analyse *single top* publiée par DØ [83] :

1. L'impulsion transverse du jet de plus grande impulsion transverse,  $p_T^{j1}$ .
2. L'impulsion transverse du second jet de plus grande impulsion transverse,  $p_T^{j2}$ .
3. L'impulsion transverse du troisième jet de plus grande impulsion transverse, s'il existe,  $p_T^{j3}$ .
4. La somme de l'impulsion transverse du lepton et de l'énergie transverse manquante,  $METLep = \cancel{E}_T + p_T^\ell$ .
5. La masse invariante du système de tous les jets (somme des quadri-impulsions de tous les jets sélectionnés de l'événement),  $M_{jets}$ .
6. La masse transverse du système composé du lepton et de l'énergie transverse manquante,  $M_T^W = \sqrt{(p_T^\ell + \cancel{E}_T)^2 - (p_x^\ell + \cancel{E}_{Tx})^2 - (p_y^\ell + \cancel{E}_{Ty})^2}$ . Cette variable correspond à la masse transverse du boson  $W$  se désintégrant leptoniquement.
7. La masse invariante du système composé du jet  $b$  de plus grande impulsion transverse, du lepton chargé et du neutrino,  $M_{Wb}$ . Le neutrino est reconstruit à partir de l'énergie transverse manquante et de contraintes de masse. Cette procédure est décrite dans l'annexe B. Pour des événements *single top* et  $t\bar{t}$ , cette variable correspond à la masse reconstruite du quark top.
8. La séparation angulaire minimale  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$  entre deux jets dans l'événement,  $\Delta R_{min}^{jj}$ .
9. Le cosinus de l'angle entre le second jet de plus haut  $p_T$  et le lepton dans le référentiel du quark top reconstruit à partir du jet de plus haut  $p_T$  et du boson

Canal électron					
Schéma d'étiquetage	Aucun	S <b>T</b> ag	D <b>T</b> ag	S <b>T</b> agVeto	D <b>T</b> agAsym
Signal					
$tb$	8.60	3.82	1.10	3.31	1.62
$tqb$	19.33	7.27	0.62	6.91	0.99
Bruits de fond					
$t\bar{t}l$ +jets	107.85	47.11	13.60	40.39	20.33
$t\bar{t}ll$	28.78	13.00	3.78	11.28	5.50
$Wbb$	69.56	27.00	5.20	24.15	8.06
$Wjj$	5458.46	114.04	2.38	111.55	4.87
$WW$	52.22	2.30	0.01	2.28	0.03
$WZ$	16.73	2.26	0.48	1.99	0.74
Multijet	653.20	22.60	0.61	21.79	1.41
Données sélectionnées					
Données	6388	248	24	229	43

TAB. 4.11: Nombres d'événements de signal et de bruits de fond attendus dans le canal électron pour chaque schéma d'étiquetage des jets  $b$ .

Canal muon					
Schéma d'étiquetage	Aucun	S <b>T</b> ag	D <b>T</b> ag	S <b>T</b> agVeto	D <b>T</b> agAsym
Signal					
$tb$	7.77	3.49	1.05	3.00	1.55
$tqb$	17.03	6.55	0.57	6.22	0.90
Bruits de fond					
$t\bar{t}l$ +jets	101.0	44.37	13.81	37.58	20.60
$t\bar{t}ll$	25.98	11.81	3.61	10.17	5.24
$Wbb$	37.90	16.55	3.50	14.67	5.38
$Wjj$	3292.59	76.48	1.64	74.70	3.42
$WW$	53.06	2.42	0.01	2.40	0.03
$WZ$	17.79	2.08	0.46	1.82	0.72
Multijet	219.0	19.13	1.46	17.92	2.67
Données sélectionnées					
Données	3913	157	14	138	33

TAB. 4.12: Nombres d'événements de signal et de bruits de fond attendus dans le canal muon pour chaque schéma d'étiquetage des jets  $b$ .

$W$ ,  $\cos(\widehat{j_2, \ell})_{top}$ . Le second jet de plus grand  $p_T$  est généralement issu du quark produit avec le top (quark  $b$  en voie  $s$ , quark léger en voie  $t$ ).

10. La sphéricité de l'événement  $\mathcal{S}$ , définie par  $\mathcal{S} = \frac{3}{2}(\lambda_2 + \lambda_3)$ , où  $\lambda_2$  et  $\lambda_3$  sont les deux plus petites valeurs propres du tenseur d'impulsion normalisé :

$$\mathcal{M}_{ij} = \frac{\sum_{jets} p_i \cdot p_j}{\sum_{jets} |\mathbf{p}|^2}.$$

11. La centralité de l'événement,  $\mathcal{C} = \frac{\sum_{jets} p_T}{\sum_{jets} |\mathbf{p}|}$ .
12. La pseudorapidité du jet non étiqueté de plus grande impulsion transverse, multipliée par la charge du lepton,  $Q \times \eta$ . Un jet est non étiqueté s'il ne passe pas le critère *JLIP Loose*. Si tous les jets passent *JLIP Loose*, les jets non étiquetés sont ceux qui ne passent pas le critère *JLIP Tight*. Enfin si tous les jets passent *JLIP Tight*, le second jet de plus haut  $p_T$  est utilisé, le jet de plus haut  $p_T$  étant supposé provenir de la désintégration du top. Le jet sélectionné est ainsi supposé être un jet léger. Cette variable, discriminante pour la production de *single top* dans la voie  $t$ , a été discutée au § 4.2.1.

Les distributions obtenues pour chacune de ces variables en combinant les canaux électron et muon ainsi que les événements simplement et doublement étiquetés sont données par les figures 4.28, 4.29 et 4.30. Ces figures comparent les différents bruits de fond estimés aux données réelles sélectionnées. Seule l'erreur statistique sur les données est indiquée. Pour assurer une meilleure lisibilité le signal *single top* voie  $s$  et  $t$  est multiplié par 10. Dans toutes les figures qui sont présentées, le signal *single top* voie  $s$  est représenté en bleu, la voie  $s$  en cyan, les fonds  $t\bar{t}$  lepton+jet en rouge,  $t\bar{t}$  dilepton en rose,  $Wbb$  en vert foncé,  $Wjj$  en vert clair et multijet en marron. Des distributions supplémentaires (canaux séparés, autres variables, sous-lots pour le contrôle des bruits de fond) sont données dans l'annexe C.

Même après les sélections le bruit de fond domine largement. Aucune de ces variables ne permet de séparer nettement le signal (voie  $s$  ou  $t$ ) des principaux bruits de fond. Les fonds dominants  $W$ +jets et  $t\bar{t}$  possèdent des topologies très différentes et le signal *single top* a souvent des distributions intermédiaires entre celles-ci. Ceci s'observe particulièrement pour des variables telles que  $p_T^j$  ou  $M_{jets}$ . C'est pourquoi une méthode d'analyse combinant plusieurs variables discriminantes est nécessaire. Pour chaque canal d'analyse, quatre nouvelles variables discriminantes ou filtres sont ainsi construites afin d'offrir une discrimination accrue entre *single top* voie  $s$  et  $W$ +jets, *single top* voie  $s$  et  $t\bar{t}$ , *single top* voie  $t$  et  $W$ +jets et *single top* voie  $t$  et  $t\bar{t}$ .

L'analyse précédente de l'expérience  $D\bar{O}$  utilisait des réseaux de neurones ou des arbres de décision pour combiner différentes variables discriminantes [101]. Pour cette analyse, une troisième approche basée sur des fonctions de vraisemblance a été choisie. Les sensibilités de ces différentes techniques d'analyse seront comparées à la fin de cette étude.

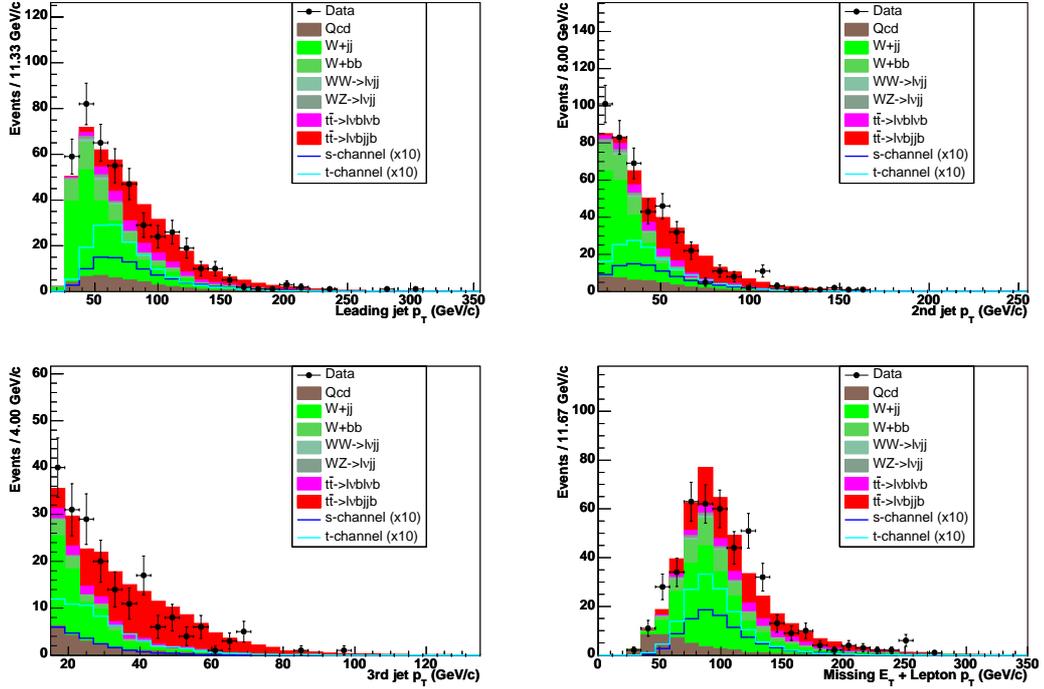


FIG. 4.28: Distributions des variables  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $p_T^{j3}$  et  $MET_{Lep}$  dans les canaux combinés électron et muon, simple et double étiquetage.

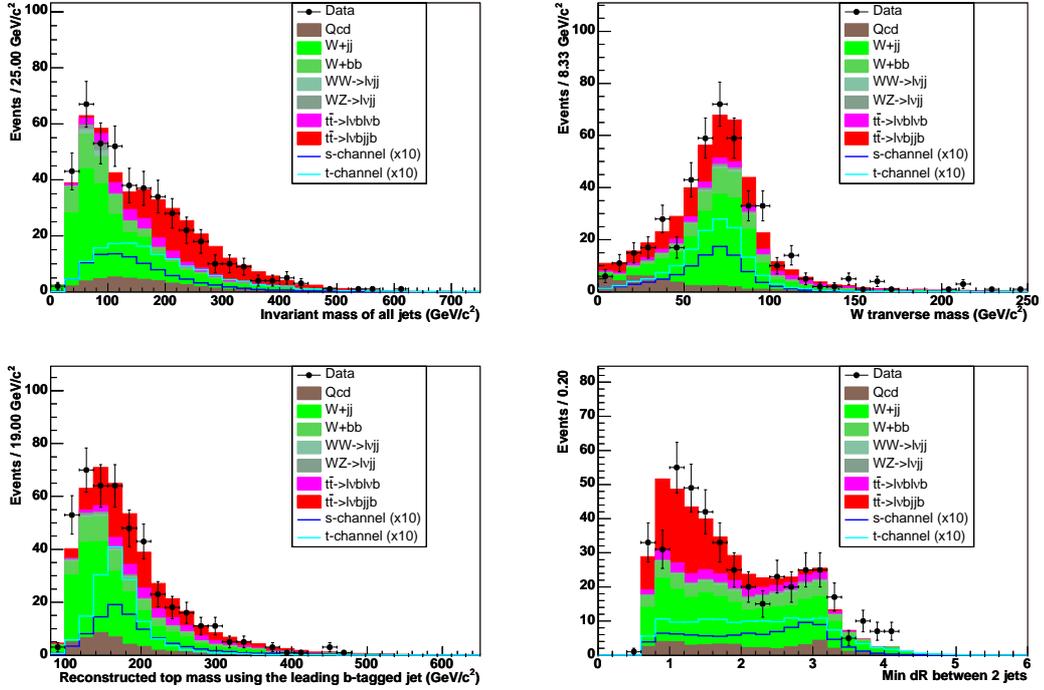


FIG. 4.29: Distributions des variables  $M_{jets}^W$ ,  $M_{Wb}$  et  $\Delta R_{min}^{jj}$  dans les canaux combinés électron et muon, simple et double étiquetage.

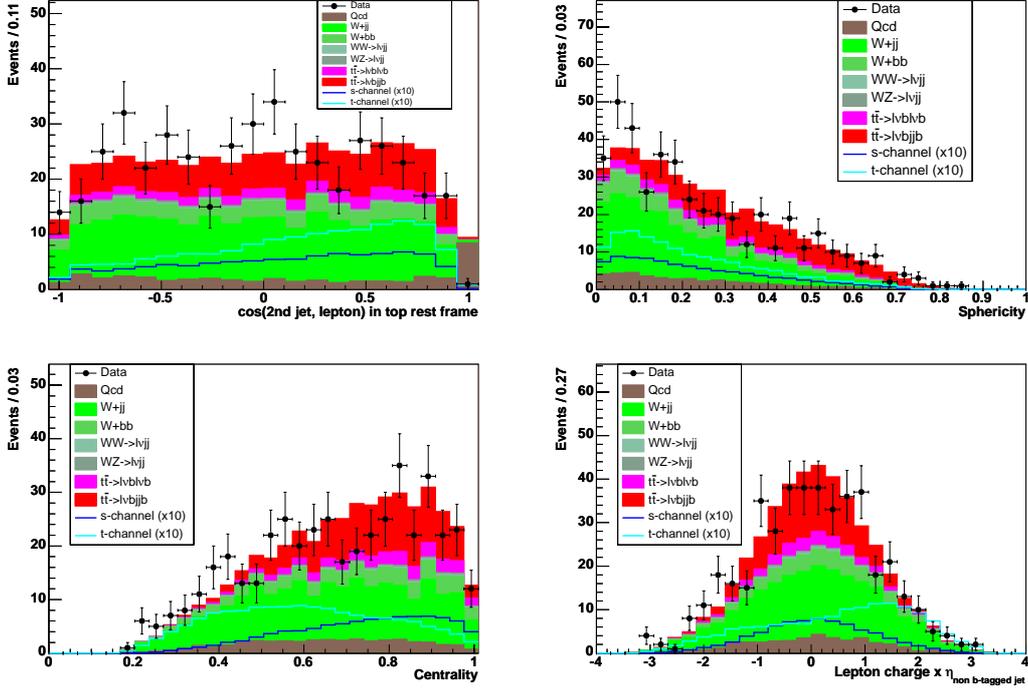


FIG. 4.30: Distributions des variables  $\cos(\hat{j}_2, \ell)_{top}$ ,  $\mathcal{S}$ ,  $\mathcal{C}$  et  $Q \times \eta$  dans les canaux combinés électron et muon, simple et double étiquetage.

## 4.8.2 Fonctions de vraisemblance

La fonction de vraisemblance est une méthode statistique robuste qui permet de combiner efficacement plusieurs variables de faible pouvoir discriminant. Chaque événement est supposé appartenir à une des deux classes : signal ou fond. La fonction de vraisemblance teste la compatibilité de chaque événement avec les hypothèses signal et signal+fond. Elle est définie en fonction d'un jeu de  $n$  variables  $\vec{x}$  par :

$$\mathcal{L}(\vec{x}) = \frac{\mathcal{P}_{\text{signal}}(\vec{x})}{\mathcal{P}_{\text{signal}}(\vec{x}) + \mathcal{P}_{\text{fond}}(\vec{x})} \quad (4.38)$$

Les fonctions  $\mathcal{P}_{\text{signal}}(\vec{x})$  et  $\mathcal{P}_{\text{fond}}(\vec{x})$  sont les densités de probabilité pour chaque classe. Les événements les plus compatibles avec l'hypothèse signal auront des valeurs de  $\mathcal{L}(\vec{x})$  voisines de l'unité alors que les événements compatibles avec le bruit de fond auront des valeurs proches de 0. Pour des variables entièrement décorrélées, les densités de probabilité à  $n$  dimensions  $\mathcal{P}_{\text{signal}}(\vec{x})$  et  $\mathcal{P}_{\text{fond}}(\vec{x})$  peuvent se factoriser en produit de densités de probabilité ne dépendant que d'une seule variable  $x_i = 1, \dots, n$  :

$$\mathcal{P}(\vec{x}) = \prod_{i=1}^n P(x_i) \quad (4.39)$$

Si  $p$  sources indépendantes sont utilisées pour définir le signal ou le fond, l'équation 4.39 se généralise en :

$$\mathcal{P}(\vec{x}) = \sum_{j=1}^p \left( f_j \prod_{i=1}^n P_j(x_i) \right) \quad (4.40)$$

où les termes  $f_j$  servent à normaliser chaque source relativement à son abondance  $\mathcal{Y}_j$  :

$$f_j = \frac{\mathcal{Y}_j}{\sum_{k=1}^p \mathcal{Y}_k} \quad (4.41)$$

Si les variables  $x_i$  sont corrélées, cette factorisation n'est plus valable. Elle peut cependant être utilisée pour approximer la véritable densité de probabilité. Ceci n'invalide pas la méthode mais réduit le pouvoir discriminant de la fonction de vraisemblance ainsi construite.

Les distributions déterminées dans la simulation et les données (voir § 4.7) sont utilisées afin d'estimer les densités de probabilité du signal et du fond en utilisant l'équation 4.40. Afin de réduire une éventuelle sensibilité à des fluctuations statistiques une interpolation *spline* d'ordre 3 de chaque histogramme est utilisée. Les fractions  $f_j$  sont dérivées des taux d'événements attendus dans chaque canal donnés dans les tableaux 4.11 et 4.12. Pour chaque canal d'analyse et chaque schéma d'étiquetage, quatre fonctions de vraisemblance sont construites :

- Un filtre **voie s-W+jets**
- Un filtre **voie s-t $\bar{t}$**
- Un filtre **voie t-W+jets**
- Un filtre **voie t-t $\bar{t}$**

Pour les schémas d'étiquetage STag et STagVeto, W+jets inclut les fonds  $Wjj$ ,  $Wbb$  et multijet. Pour les schémas DTag et DTagAsym seul le lot  $Wbb$  est utilisé. Le lot de données multijet souffre d'une trop faible statistique d'événements doublement étiquetés pour modéliser la densité de probabilité de manière satisfaisante. Le lot  $Wjj$  contient une faible fraction d'événements  $Wcc$  qui possèdent des probabilités de double étiquetage très supérieures (de trois ordres de grandeur environ) aux événements  $Wjj$ , ce qui crée artificiellement des pics dans les distributions pondérées. Le fond  $t\bar{t}$  regroupe les modes de désintégration lepton+jet et dilepton.

Au total, 32 fonctions de vraisemblance (4 filtres  $\times$  2 canaux d'analyse  $\times$  4 schémas d'étiquetage) sont utilisées dans l'analyse. Divers ensembles de variables discriminantes sélectionnées parmi les variables utilisées dans les réseaux de neurones de l'analyse *single top* précédente [101] ou dans la mesure de la section efficace  $t\bar{t}$  utilisant également une fonction de vraisemblance [102] ont été testés afin de sélectionner les variables à utiliser dans la construction de nos fonctions de vraisemblance. Les jeux de variables retenus sont ceux qui optimisent le pouvoir séparateur de la fonction de vraisemblance. Les mêmes jeux de variables sont utilisés dans les canaux électron et muon. Les fonctions de vraisemblance ainsi construites ont le même pouvoir discriminant dans les deux canaux. Les jeux de variables pour chaque classe de filtres sont donnés dans le tableau 4.13. Les

	voie $s$ - $W$ +jets	voie $s$ - $t\bar{t}$	voie $t$ - $W$ +jets	voie $t$ - $t\bar{t}$
1. $p_T^{j_1}$	✓	✓	✓	✓
2. $p_T^{j_2}$	✓	✓	✓	✓
3. $p_T^{j_3}$	—	✓	—	✓
4. $METLep$	✓	✓	✓	✓
5. $M_{jets}$	✓	✓	✓	✓
6. $M_T^W$	✓	✓	✓	✓
7. $M_{Wb}$	✓	—	✓	—
8. $\Delta R_{min}^{jj}$	—	✓	—	✓
9. $\cos(j_2, \ell)_{top}$	✓	—	✓	—
10. <i>Sphéricité</i>	—	✓	—	✓
11. <i>Centralité</i>	—	✓	—	✓
12. $Q \times \eta$	—	—	✓	✓

TAB. 4.13: Jeux de variables retenus (✓) pour les quatre classes de fonctions de vraisemblance.

variables utilisées sont celles décrites au § 4.8.1. La masse transverse du boson  $W$  ( $M_T^W$ ) et  $METLep$  servent principalement avec les schémas d'étiquetage STag et STagVeto, où le fond multijet est inclus dans la densité de probabilité. La variable  $Q \times \eta$  ne sert que pour la voie  $t$ . Les densités de probabilité sont illustrées par les figures de l'annexe D.

Pour quantifier le pouvoir séparateur des fonctions de vraisemblance, des courbes de performance sont déterminées pour chaque filtre. Chaque courbe représente l'efficacité de sélection d'un signal (voie  $s$  ou  $t$ ) en fonction de celle d'un bruit de fond ( $t\bar{t}$  lepton+jet,  $t\bar{t}$  dilepton,  $Wbb$ ,  $Wjj$ ,  $WW$ ,  $WZ$ , multijet) pour différentes coupures sur la valeur de la fonction de vraisemblance. L'optimisation du pouvoir séparateur est réalisée en minimisant l'aire située au dessus de la courbe de performance correspondant au signal et au fond utilisés dans la construction de la fonction de vraisemblance. Ces courbes de performance pour le schéma d'étiquetage STag sont données sur la figure 4.31.

Les distributions des fonctions de vraisemblance pour les différents canaux sont représentées sur les figures 4.33 à 4.36. Ces figures montrent un bon accord entre les données observées et les estimations des différents bruits de fond. Ces figures incluent également les représentations à deux dimensions, pour un même signal du filtre signal- $W$ +jets en fonction du filtre signal- $t\bar{t}$ . Ce sont ces distributions qui sont utilisées pour le calcul des limites sur les sections efficaces de production présentées au § 4.10. Afin d'augmenter le nombre d'événements utilisés pour déterminer les densités de probabilité et ainsi limiter la sensibilité à des fluctuations statistiques, l'intégralité des lots de données simulées est utilisée. Ainsi les densités de probabilité sont déterminées sur les mêmes événements que ceux utilisés pour calculer les distributions de ces fonctions de vraisemblance. Il est donc nécessaire de vérifier que cette approche n'introduit pas de biais dans l'analyse. Pour cela, chaque lot de données est divisé en deux en séparant les événements pairs des événements impairs. Deux séries de distributions de probabilité sont ensuite dérivées à

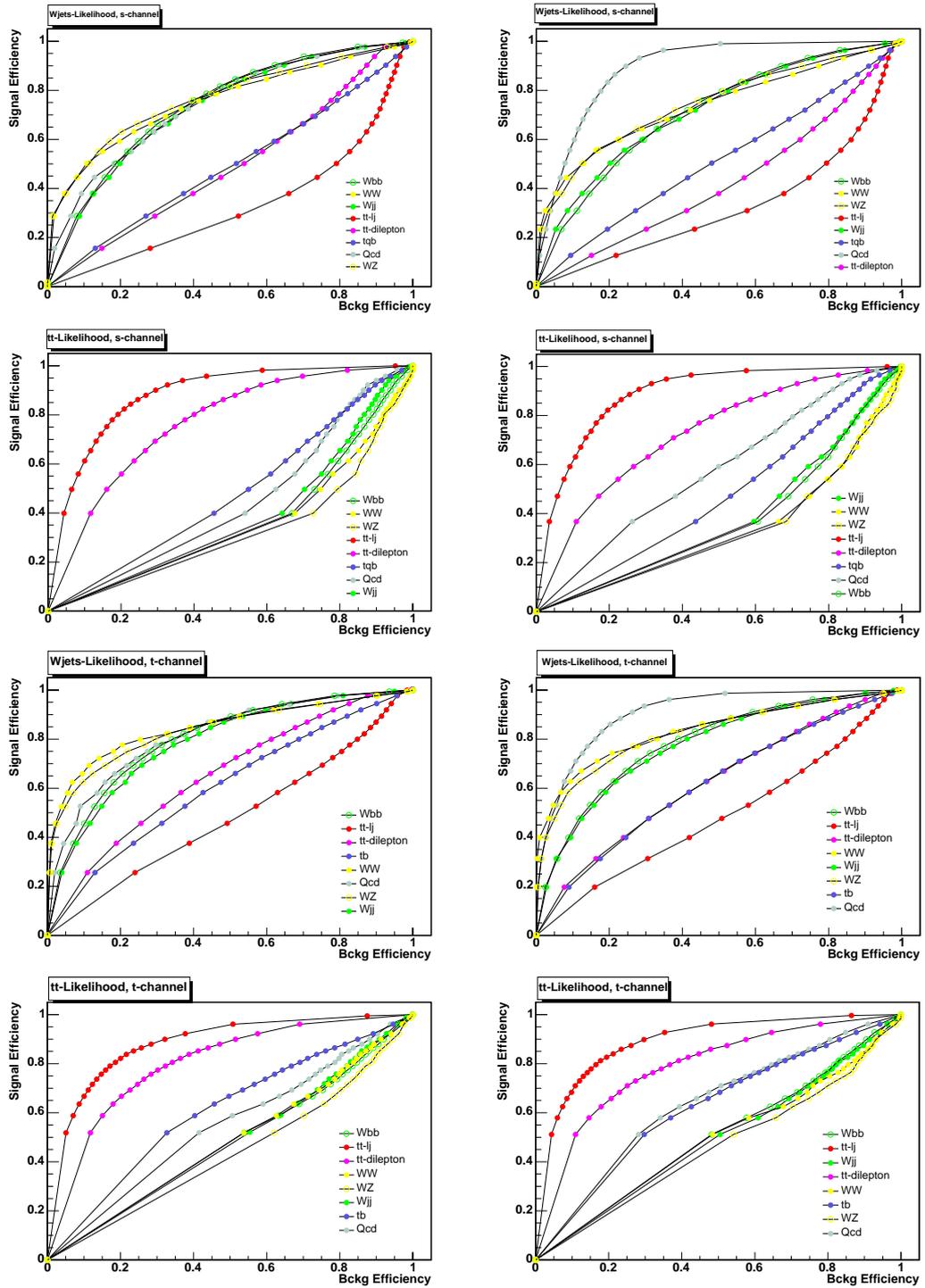


FIG. 4.31: Courbes de performance pour le schéma d'étiquetage STag dans les canaux électron (à gauche) et muon (à droite).

De haut en bas les filtres voie  $s$ - $W$ +jets, voie  $s$ - $t\bar{t}$ , voie  $t$ - $W$ +jets et voie  $t$ - $t\bar{t}$ .

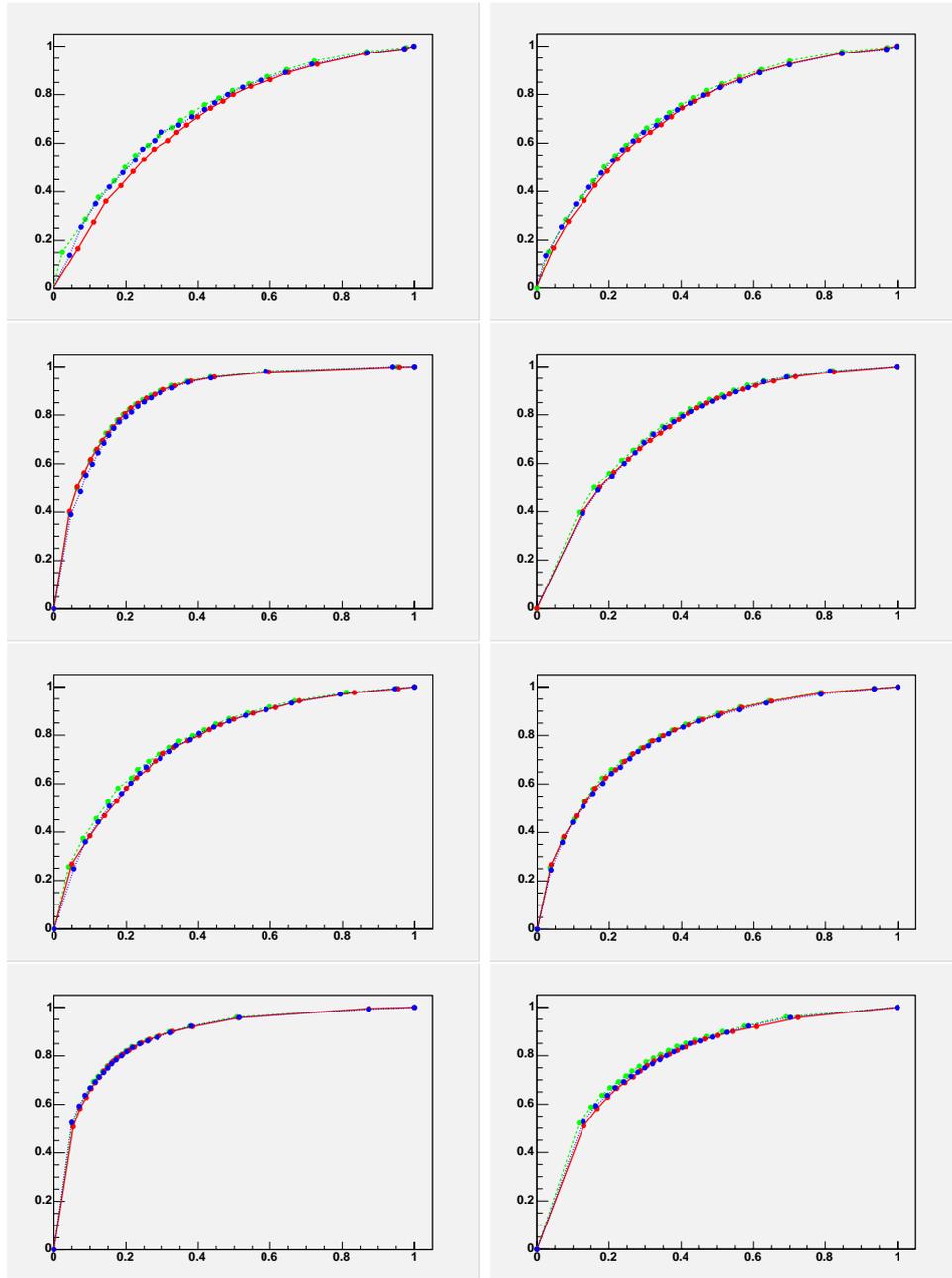


FIG. 4.32: Courbes de performance pour les quatre filtres dans le canal électron, STag.  
 Première ligne : filtre voie  $s$ - $W$ +jet, à gauche fond  $Wjj$ , à droite fond  $Wbb$ .  
 Deuxième ligne : filtre voie  $s$ - $t\bar{t}$ , à gauche fond  $t\bar{t}$   $\ell$ +jets, à droite fond  $t\bar{t}$  dilepton.  
 Troisième ligne : filtre voie  $t$ - $W$ +jet, à gauche fond  $Wjj$ , à droite fond  $Wbb$ .  
 Quatrième ligne : filtre voie  $s$ - $t\bar{t}$ , à gauche fond  $t\bar{t}$   $\ell$ +jets, à droite fond  $t\bar{t}$  dilepton.  
 La courbe verte est obtenue avec la statistique complète pour les probabilités et l'analyse. La courbe rouge (bleu) utilise les événements pairs (impairs) pour les probabilités et les événements impairs (pairs) pour l'analyse.

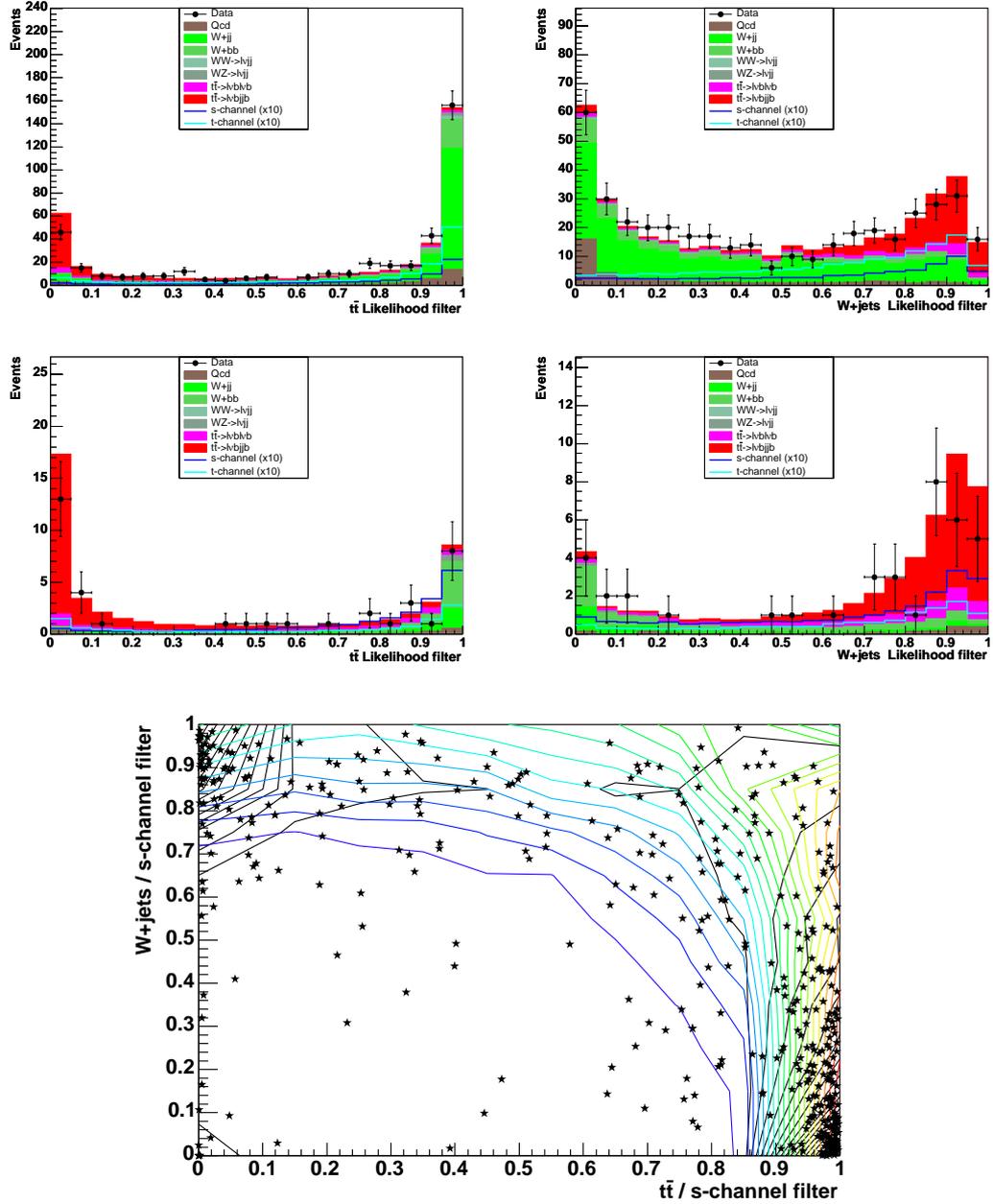


FIG. 4.33: En haut : comparaison entre données observées, signal et bruits de fond estimés pour les fonctions de vraisemblance voie  $s$ - $t\bar{t}$  (à gauche) et voie  $s$ - $W$ +jets (à droite) pour les schémas d'étiquetage STag (en haut) et DTag (en bas).

En bas : distribution à 2 dimensions du filtre  $s$ - $W$ +jets en fonction du filtre  $s$ - $t\bar{t}$ , combinant les événements STag et DTag. Les étoiles indiquent les données réelles, les contours le signal (couleur) et les fonds (noir).

Les canaux électron et muon sont combinés.

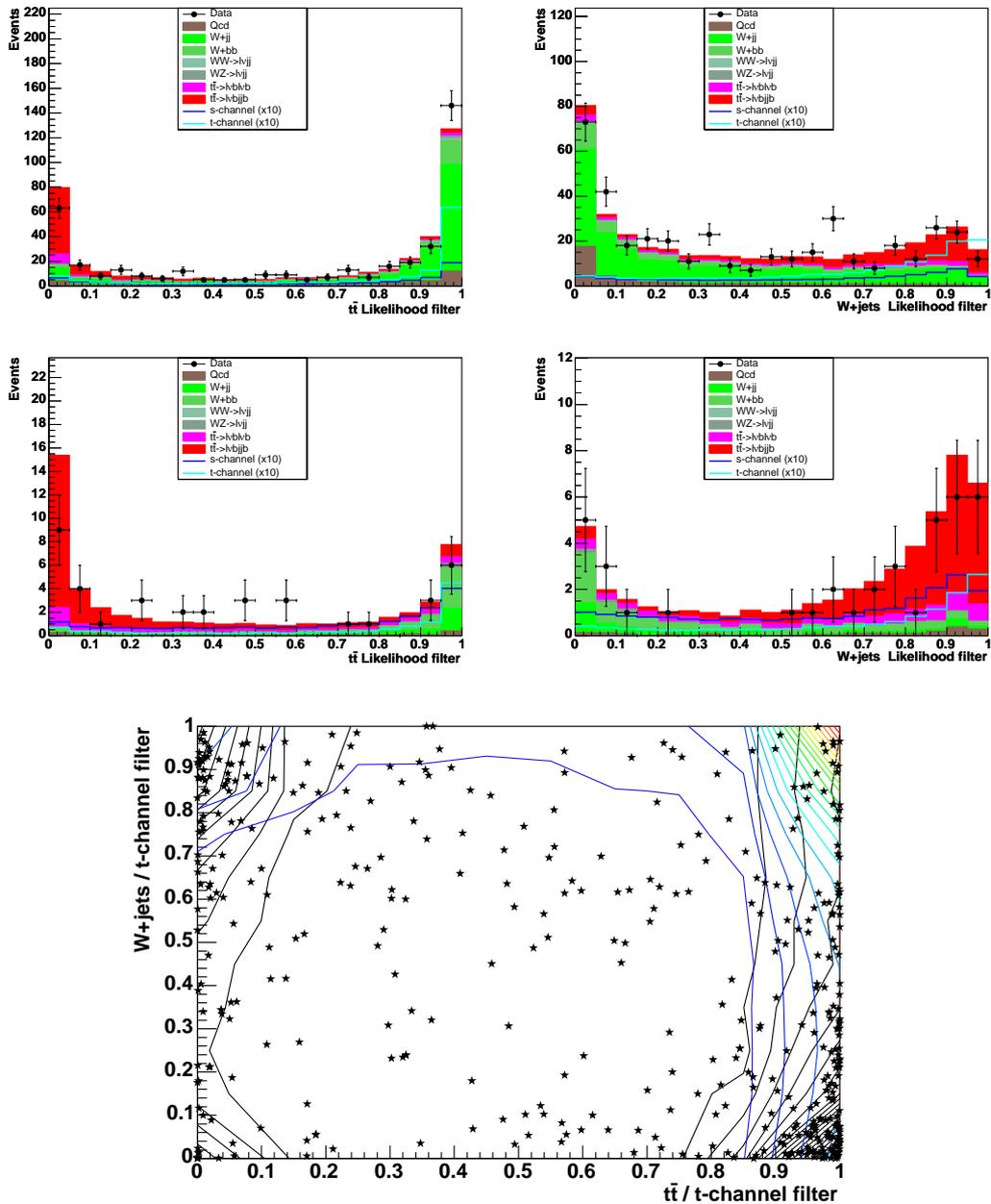


FIG. 4.34: En haut : comparaison entre données observées, signal et bruits de fond estimés pour les fonctions de vraisemblance voie  $t\bar{t}\bar{t}$  (à gauche) et voie  $t\bar{t}W+jets$  (à droite) pour les schémas d'étiquetage STag (en haut) et DTag (en bas).

En bas : distribution à 2 dimensions du filtre  $t\bar{t}W+jets$  en fonction du filtre  $t\bar{t}\bar{t}$ , combinant les événements STag et DTag. Les étoiles indiquent les données réelles, les contours le signal (couleur) et les fonds (noir).

Les canaux électron et muon sont combinés.

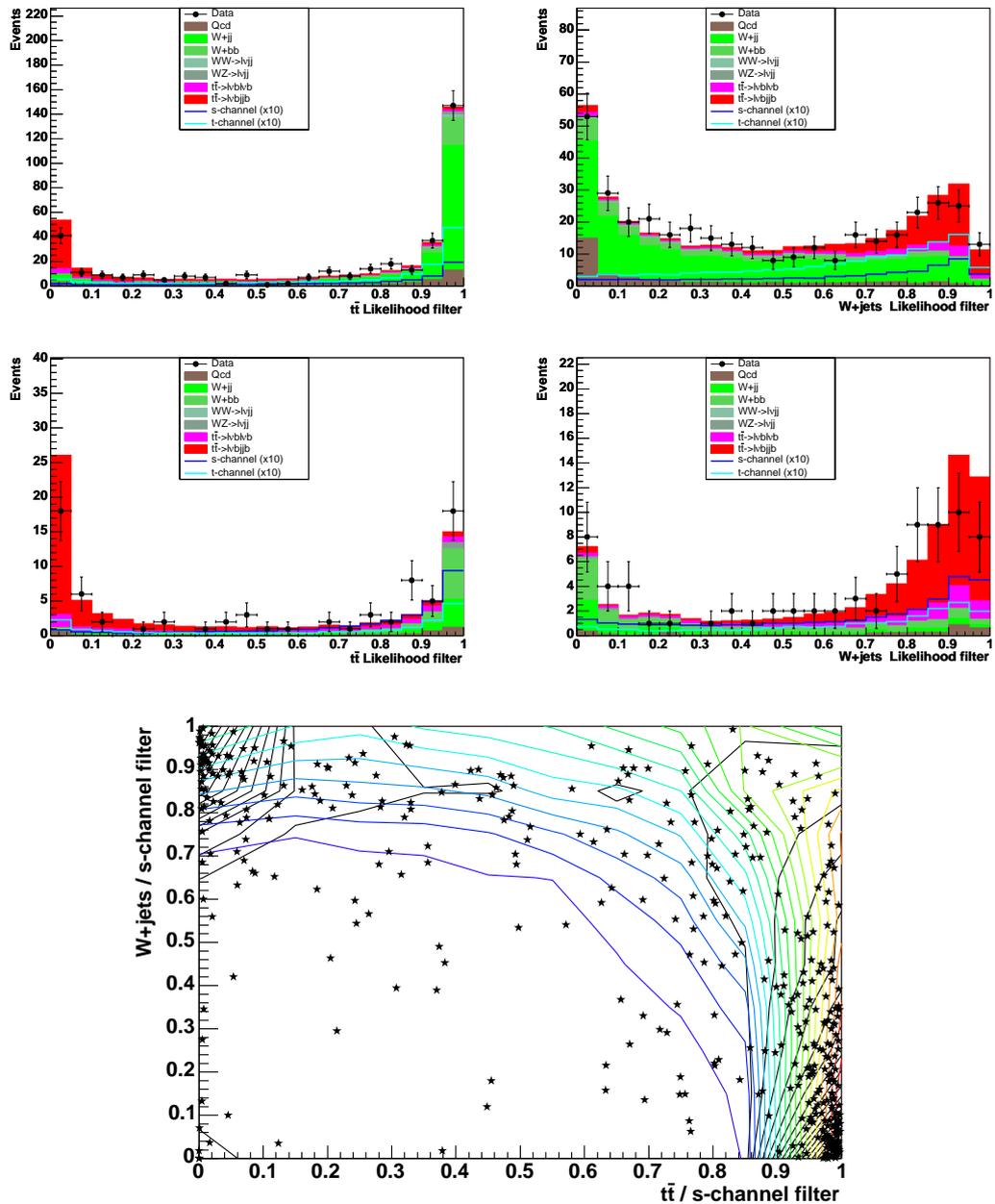


FIG. 4.35: En haut : comparaison entre données observées, signal et bruits de fond estimés pour les fonctions de vraisemblance voie  $s\text{-}t\bar{t}$  (à gauche) et voie  $s\text{-}W+\text{jets}$  (à droite) pour les schémas d'étiquetage STagVeto (en haut) et DTagAsym (en bas).

En bas : distribution à 2 dimensions du filtre  $s\text{-}W+\text{jets}$  en fonction du filtre  $s\text{-}t\bar{t}$ , combinant les événements STagVeto et DTagAsym. Les étoiles indiquent les données réelles, les contours le signal (couleur) et les fonds (noir).

Les canaux électron et muon sont combinés.

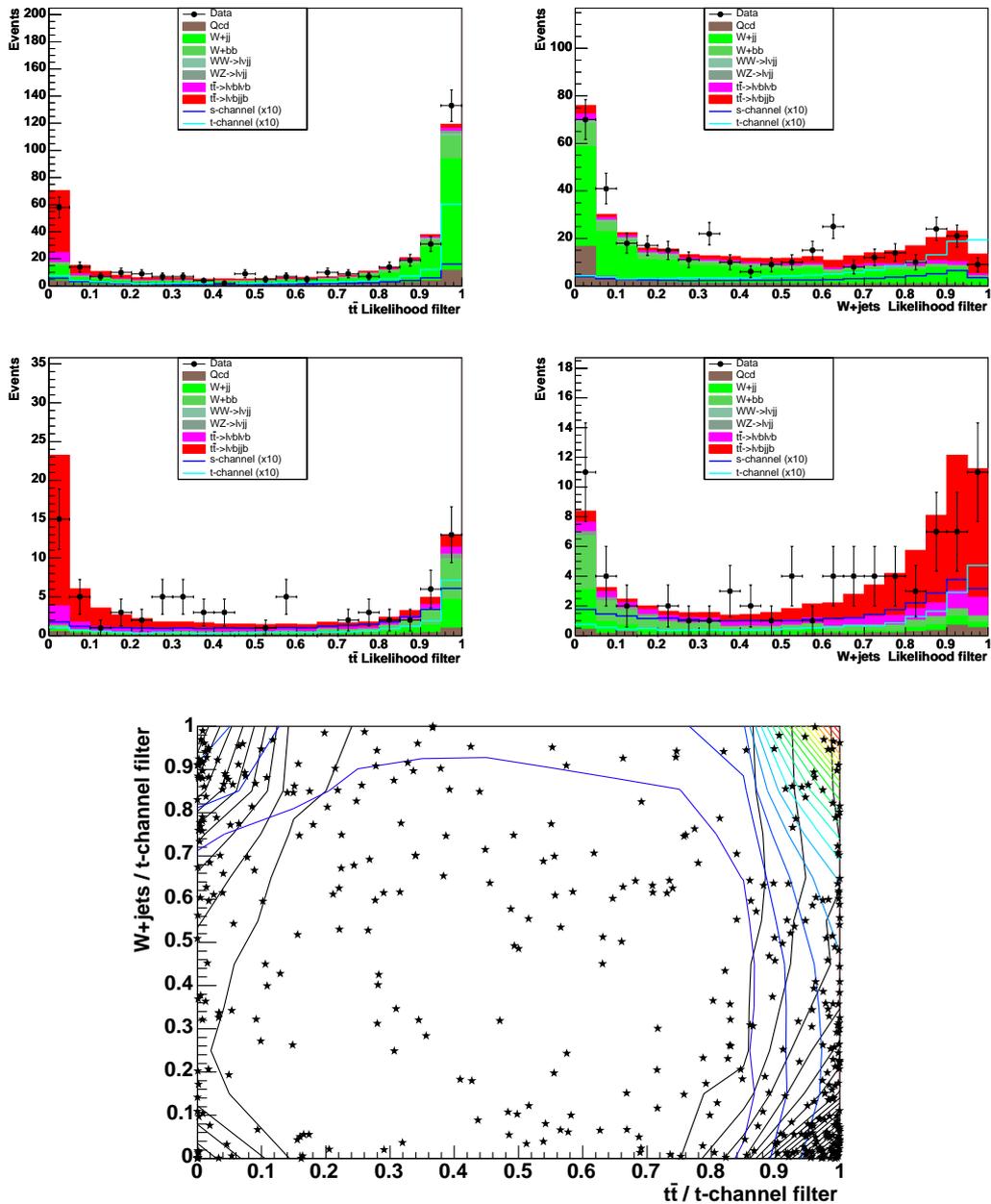


FIG. 4.36: En haut : comparaison entre données observées, signal et bruits de fond estimés pour les fonctions de vraisemblance voie  $t\bar{t}\bar{t}$  (à gauche) et voie  $t\bar{t}\bar{t}$ -W+jets (à droite) pour les schémas d'étiquetage STagVeto (en haut) et DTagAsym (en bas).

En bas : distribution à 2 dimensions du filtre  $t\bar{t}\bar{t}$ -W+jets en fonction du filtre  $t\bar{t}\bar{t}$ , combinant les événements STagVeto et DTagAsym. Les étoiles indiquent les données réelles, les contours le signal (couleur) et les fonds (noir).

Les canaux électron et muon sont combinés.

partir de chacun de ces lots. Les courbes de performance sont alors calculées sur le lot pair, en utilisant les densités de probabilité du lot impair et réciproquement. Les deux courbes de performance non biaisées ainsi produites sont comparées à celles obtenues avec la statistique complète. Ces courbes sont présentées sur la figure 4.32 pour le canal électron dans le schéma d'étiquetage STag. Dans tous les canaux d'analyse, aucun biais significatif n'est observé dans la limite des erreurs statistiques, ce qui valide l'approche avec la statistique complète.

## 4.9 Erreurs systématiques

De nombreuses sources d'erreurs systématiques, associées aux différentes étapes de sélection et de correction sont considérées dans cette analyse. Leur détermination peut être complexe et souvent l'approche la plus conservative est adoptée. On distingue deux classes d'erreurs systématiques. La première regroupe les erreurs qui affectent la normalisation globale des distributions et qui proviennent des incertitudes sur les facteurs multiplicatifs constants appliqués dans chaque canal d'analyse. Ces erreurs ne concernent que le signal et les fonds entièrement étudiés au moyen de la simulation (*single top* voies  $s$  et  $t$ ,  $t\bar{t}$ ,  $WW$ ,  $WZ$ ), à l'exception de l'erreur sur la méthode de la matrice qui n'affecte que les fonds  $W$ +jets et multijet :

- l'incertitude sur la mesure de la luminosité qui affecte l'estimation du signal et du fond. Elle vaut 6.5% à DØ au Run II [60].
- l'incertitude théorique sur les sections efficaces (*single top* voies  $s$  et  $t$ ,  $t\bar{t}$  et dibosons) varie de 2% à 18% (c.f. tableaux 1.4 et 4.1) selon le processus physique.
- l'incertitude sur les rapports d'embranchement correspondant aux lots de données simulées. Cette erreur est fixée à 2% (c.f. tableaux 1.4 et 4.1)
- l'erreur relative liée aux différences d'efficacité de reconstruction du vertex et des leptons entre données et simulation. Ces différences sont corrigées par des facteurs multiplicatifs. Il demeure une erreur résiduelle liée à l'estimation de ces facteurs correctifs. Cette erreur varie de 2 à 5% (c.f. § 4.6.1).
- l'incertitude sur la fragmentation des jets dans la simulation. Cette incertitude est dérivée des différences de modélisation des radiations initiales et finales et de la fragmentation entre les générateurs PYTHIA et HERWIG (5-7%). Les valeurs utilisées ont été déterminées dans l'analyse *single top* du Run I [82].
- les incertitudes sur la normalisation des fonds  $W$ +jets et multijet. Ces erreurs proviennent de la méthode de la matrice et de la normalisation relative des fonds  $Wjj$  et  $Wbb$ . La combinaison de ces erreurs est détaillée dans l'annexe A.

L'autre catégorie correspond aux incertitudes qui dépendent de la topologie de l'événement et affectent à la fois la normalisation et la forme des distributions cinématiques, ce qui se répercute sur les fonctions de vraisemblance. Seules les distributions de la fonction de vraisemblance sont affectées par ces erreurs systématiques, pas les densités de

probabilité utilisées dans la définition de cette fonction de vraisemblance. Les erreurs systématiques entrant dans cette seconde catégorie sont :

- l’incertitude sur l’efficacité de reconstruction des jets. Cette efficacité dépend principalement de l’impulsion transverse du jet. Les différences entre données réelles et simulation sont prises en compte en retirant de manière aléatoire, avant sélection, des jets dans les événements simulés. L’énergie transverse manquante est recalculée en conséquence. La fraction de jets rejetés correspond au rapport d’efficacité d’identification entre données réelles et simulation. L’erreur systématique associée à cette procédure est estimée en variant cette fraction de  $\pm 1$  écart standard. Les lots simulés ainsi modifiés sont processés à travers toute la chaîne d’analyse.
- l’incertitude sur la correction en énergie des jets (*JES*). Elle est mesurée en variant la correction en énergie de  $\pm 1$  écart standard. Cette variation inclut l’erreur sur *JES* dans les données réelles et dans la simulation.
- l’incertitude systématique sur la résolution en énergie des jets. Cette erreur est estimée en variant de  $\pm 1$  écart standard les facteurs de dégradation de la résolution des jets dans la simulation.
- l’incertitude liée à la modélisation des efficacités de déclenchement. Les erreurs supérieures et inférieures associées à chaque terme des critères de déclenchement sont déterminées indépendamment en variant l’efficacité du terme de  $\pm 1$  écart standard. Les erreurs provenant de chaque terme sont ensuite sommées quadratiquement pour obtenir les erreurs supérieure et inférieure finales sur chaque poids de déclenchement.
- l’incertitude attachée à la modélisation des efficacités d’identification des jets *b* et de la *taggabilité*. Cette erreur est calculée en utilisant les *TRF* à  $\pm 1$  écart standard (voir § 3.5.4 et § 4.6.3.2). Les contributions de la *taggabilité* et des *TRF* sont traitées indépendamment et sommées quadratiquement.
- l’erreur systématique résultant de la statistique limitée des lots de données simulées.

L’estimation des trois sources d’erreurs systématiques sur les jets (identification, correction en énergie et résolution) conduit à modifier le nombre de jets ou leur impulsion dans les événements simulés, avant sélection. Ces variations sont répercutées sur l’énergie transverse manquante. Toute la chaîne d’analyse (sélection et pondération) est alors appliquée à ces données simulées modifiées. L’effet des sélections et des pondérations sur ces incertitudes est totalement pris en compte.

Les huit tableaux 4.14 à 4.21 présentent les contributions de chaque source à l’erreur finale pour chaque canal d’analyse et chaque schéma d’étiquetage. Les différentes sources d’erreur pour un même fond sont traitées de manière décorrélée, et combinées quadratiquement. L’erreur totale sur les acceptances n’inclut pas les erreurs sur la luminosité et les sections efficaces (c.f. équation 4.18). Les erreurs affectant la forme des distributions varient suivant la cellule de l’histogramme de la fonction de vraisemblance considérée. A titre indicatif, ces tableaux incluent des valeurs moyennes pour ces erreurs. Quand l’erreur est asymétrique, la plus grande valeur est présentée. Le calcul des limites prend

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	5.0	5.0	—	—	16.2
Vertex primaire	1.8	1.8	1.8	1.8	—	—	1.8	1.8	—
Electron ID	3.6	3.6	3.6	3.6	—	—	3.6	3.6	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	4.8	5.9	1.1	3.7	—	—	13.3	17.2	—
Echelle en énergie	4.0	4.2	0.6	2.7	—	—	9.4	9.0	—
Résolution	1.1	1.6	0.4	0.9	—	—	9.1	9.7	—
Déclenchement	3.9	4.3	1.8	2.5	—	—	5.8	6.6	—
$TRF$	6.4	7.2	7.2	6.9	5.5	8.4	7.9	5.4	—
Statistique	0.8	1.0	0.5	0.6	1.0	2.4	3.7	3.1	4.2
Total									
Erreur sur l'acceptance	13.6	14.6	13.0	12.8	—	—	23.4	25.3	—
Erreur sur le nombre									
Total d'événements	15.8	18.3	22.2	22.1	7.5	10.1	23.5	26.8	16.7

TAB. 4.14: Erreurs systématiques relatives, canal électron, schéma STag.

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	4.6	4.6	—	—	25.6
Vertex primaire	2.5	2.5	2.5	2.5	—	—	2.5	2.5	—
Muon ID	5.3	5.3	5.3	5.3	—	—	5.3	5.3	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	3.8	5.7	0.4	3.2	—	—	12.8	8.5	—
Echelle en énergie	4.5	5.3	0.06	2.4	—	—	11.8	10.4	—
Résolution	0.8	1.4	0.4	0.8	—	—	3.8	3.4	—
Déclenchement	6.9	7.4	2.2	3.0	—	—	9.9	11.3	—
$TRF$	7.1	7.9	7.9	7.7	6.5	9.2	8.9	6.3	—
Statistique	0.9	1.2	0.5	0.7	1.1	2.6	3.7	3.3	2.0
Total									
Erreur sur l'acceptance	15.5	16.9	14.1	13.8	—	—	24.8	21.8	—
Erreur sur le nombre									
Total d'événements	17.4	20.2	22.8	22.7	8.0	10.6	24.9	23.6	25.6

TAB. 4.15: Erreurs systématiques relatives, canal muon, schéma STag.

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	17.2	17.2	—	—	15.8
Vertex primaire	1.8	1.8	1.8	1.8	—	—	1.8	1.8	—
Electron ID	3.6	3.6	3.6	3.6	—	—	3.6	3.6	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	5.0	6.7	1.9	3.7	—	—	16.3	16.9	—
Echelle en énergie	5.4	7.3	2.4	4.2	—	—	11.5	10.8	—
Résolution	2.0	3.7	0.6	1.3	—	—	10.2	10.2	—
Déclenchement	3.0	3.3	1.7	2.1	—	—	4.5	5.2	—
$TRF$	17.0	15.4	17.1	17.5	14.0	14.2	19.2	13.2	—
Statistique	1.3	2.3	0.7	1.1	1.8	12.5	18.6	5.4	25.8
Total									
Erreur sur l'acceptance	21.1	21.3	20.4	20.8	—	—	36.4	28.7	—
Erreur sur le nombre									
Total d'événements	22.6	24.0	27.2	27.5	22.3	25.6	36.4	30.0	30.3

TAB. 4.16: Erreurs systématiques relatives, canal électron, schéma DTag.

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	17.0	17.0	—	—	15.3
Vertex primaire	2.5	2.5	2.5	2.5	—	—	2.5	2.5	—
Muon ID	5.3	5.3	5.3	5.3	—	—	5.3	5.3	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	3.5	4.0	1.4	3.6	—	—	8.6	7.0	—
Echelle en énergie	5.5	7.6	2.5	3.8	—	—	14.2	10.6	—
Résolution	0.9	0.1	0.7	1.0	—	—	3.5	2.3	—
Déclenchement	5.6	6.1	1.9	2.3	—	—	9.8	9.2	—
$TRF$	18.4	16.8	18.5	19.0	16.0	15.1	20.5	15.1	—
Statistique	1.5	2.6	0.8	1.2	2.0	12.4	9.9	5.9	7.4
Total									
Erreur sur l'acceptance	22.8	22.5	22.6	22.4	—	—	31.7	24.9	—
Erreur sur le nombre									
Total d'événements	24.2	25.0	28.9	28.7	23.4	25.9	31.8	26.4	17.0

TAB. 4.17: Erreurs systématiques relatives, canal muon, schéma DTag.

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	4.7	4.7	—	—	16.2
Vertex primaire	1.8	1.8	1.8	1.8	—	—	1.8	1.8	—
Electron ID	3.6	3.6	3.6	3.6	—	—	3.6	3.6	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	4.7	5.8	1.4	3.6	—	—	13.3	17.2	—
Echelle en énergie	3.9	4.1	0.8	2.6	—	—	9.4	8.9	—
Résolution	1.0	1.5	0.4	0.9	—	—	9.2	9.6	—
Déclenchement	3.9	4.4	1.9	2.6	—	—	5.8	6.7	—
$TRF$	5.2	6.9	6.2	5.8	4.8	8.3	7.9	4.8	—
Statistique	0.9	1.1	0.6	0.7	1.1	2.5	3.8	3.4	4.3
Total									
Erreur sur l'acceptance	13.0	14.4	12.5	12.2	—	—	23.4	25.2	—
Erreur sur le nombre									
Total d'événements	15.3	18.1	21.9	21.7	6.8	9.9	23.5	26.7	16.8

TAB. 4.18: Erreurs systématiques relatives, canal électron, schéma STagVeto.

	$tb$	$tqb$	$t\bar{t}lj$	$t\bar{t}ll$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	4.2	4.2	—	—	25.1
Vertex primaire	2.5	2.5	2.5	2.5	—	—	2.5	2.5	—
Muon ID	5.3	5.3	5.3	5.3	—	—	5.3	5.3	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	3.8	5.7	0.8	3.0	—	—	12.8	8.8	—
Echelle en énergie	4.5	5.3	0.2	2.3	—	—	11.7	10.5	—
Résolution	1.0	1.4	0.4	0.8	—	—	3.8	3.6	—
Déclenchement	7.0	7.5	2.2	3.1	—	—	9.9	11.4	—
$TRF$	5.7	7.6	6.6	6.3	5.6	9.1	8.8	5.4	—
Statistique	1.0	1.2	0.6	0.8	1.2	2.7	3.8	3.7	2.1
Total									
Erreur sur l'acceptance	15.0	16.8	13.4	13.1	—	—	24.7	21.9	—
Erreur sur le nombre									
Total d'événements	17.0	20.1	22.4	22.2	7.1	10.4	24.8	23.6	25.2

TAB. 4.19: Erreurs systématiques relatives, canal muon, schéma STagVeto.

	$tb$	$tqb$	$t\bar{l}j$	$t\bar{l}l$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	15.6	15.6	—	—	15.9
Vertex primaire	1.8	1.8	1.8	1.8	—	—	1.8	1.8	—
Electron ID	3.6	3.6	3.6	3.6	—	—	3.6	3.6	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	5.2	7.0	1.9	3.8	—	—	13.5	17.2	—
Echelle en énergie	5.1	7.1	2.1	3.9	—	—	11.6	10.3	—
Résolution	2.0	3.7	0.5	1.3	—	—	8.1	10.3	—
Déclenchement	3.1	3.4	1.7	2.2	—	—	4.8	5.4	—
$TRF$	16.0	14.2	15.9	16.5	12.9	13.4	15.9	12.2	—
Statistique	1.0	1.5	0.5	0.8	1.2	6.7	7.3	3.7	16.9
Total									
Erreur sur l'acceptance	20.3	20.4	19.4	19.9	—	—	28.3	28.1	—
Erreur sur le nombre									
Total d'événements	21.8	23.2	26.5	26.8	20.3	21.6	28.4	29.4	23.2

TAB. 4.20: Erreurs systématiques relatives, canal électron, schéma DTagAsym.

	$tb$	$tqb$	$t\bar{l}j$	$t\bar{l}l$	$Wbb$	$Wjj$	$WW$	$WZ$	Multijet
Normalisation									
Luminosité	6.5	6.5	6.5	6.5	—	—	6.5	6.5	—
Section efficace	16.0	15.0	18.0	18.0	—	—	2.2	8.9	—
Rapport d'embranchement	2.0	2.0	2.0	2.0	—	—	2.0	2.0	—
Méthode de la matrice	—	—	—	—	15.3	15.3	—	—	16.3
Vertex primaire	2.5	2.5	2.5	2.5	—	—	2.5	2.5	—
Muon ID	5.3	5.3	5.3	5.3	—	—	5.3	5.3	—
Fragmentation	5.0	5.0	7.0	5.0	—	—	5.0	5.0	—
Normalisation et forme									
Jets :									
Reconstruction	3.7	4.5	1.5	3.8	—	—	8.4	6.9	—
Echelle en énergie	5.2	7.3	2.1	3.5	—	—	14.2	10.4	—
Résolution	0.8	0.3	0.7	1.0	—	—	4.4	2.2	—
Déclenchement	5.8	6.3	1.9	2.4	—	—	9.7	9.6	—
$TRF$	17.6	15.7	17.3	18.2	15.0	14.5	17.2	14.2	—
Statistique	1.1	1.8	0.5	0.9	1.4	6.6	4.2	4.1	5.4
Total									
Erreur sur l'acceptance	22.2	21.6	21.0	21.7	—	—	28.4	24.0	—
Erreur sur le nombre									
Total d'événements	23.6	24.3	27.6	28.2	21.5	22.1	28,5	25.6	17.2

TAB. 4.21: Erreurs systématiques relatives, canal muon, schéma DTagAsym.

en compte les distributions et les asymétries de ces sources d'erreurs. Le tableau 4.22 présente les nombres d'événements attendus pour chaque schéma d'étiquetage avec leur erreurs systématiques.

Pour combiner différents fonds, les erreurs provenant d'une même source sont supposées entièrement corrélées et sommées linéairement. Ceci correspond à sommer linéairement les erreurs présentées sur une même ligne des tableaux 4.14 à 4.17 (à l'exception des sections efficaces qui ne sont supposées corrélées qu'entre les deux fonds  $t\bar{t}$ ). Les bruits de fond  $W$ +jets et multijet sont combinés séparément. L'anticorrélation des erreurs introduite par la méthode de la matrice est prise en compte. Le calcul de l'erreur totale sur la combinaison des fonds  $W$ +jets et multijet est détaillé dans l'annexe A. L'erreur totale sur le fond combiné est la somme quadratique de toutes les sources d'erreurs.

## 4.10 Calcul des limites

### 4.10.1 Limites bayésiennes sur la production *single top* dans les voies $s$ et $t$

Les distributions à deux dimensions des filtres signal- $W$ +jets/signal- $t\bar{t}$ , découpées en  $10 \times 10$  cellules sont utilisées pour calculer des limites supérieures à 95% de niveau de confiance sur les sections efficaces de production *single top* dans les voies  $s$  et  $t$ . Les limites sont mesurées en utilisant une méthode bayésienne [103]. Toutes les cellules des histogrammes à deux dimensions sont traitées comme des canaux statistiquement indépendants. Le détail du calcul des limites, incluant le traitement des erreurs systématiques et leur corrélation entre chaque cellule, est donné dans l'annexe E. La distribution (voie  $s$ - $W$ +jets, voie  $s$ - $t\bar{t}$ ) est utilisée pour extraire la limite sur la production *single top* dans la voie  $s$ . La voie  $t$  est alors considérée comme un bruit de fond, normalisé sur sa section efficace théorique dans le Modèle Standard (voir tableau 1.4). Ce procédé est inversé pour le calcul de la limite dans la voie  $t$ . La limite obtenue pour chaque mode de production est donc dépendante du Modèle Standard.

Les canaux électron et muon sont combinés, de même que les schémas d'étiquetage : soit STag et DTag, soit STagVeto et DTagAsym. Les limites sont également calculées dans chaque canal séparément et pour des combinaisons partielles.

Deux types de limites sont calculées : la limite attendue où le nombre d'événements observés est fixé à la même valeur que le nombre d'événements de bruits de fond attendus et la limite mesurée qui utilise le nombre d'événements observés dans les données réelles. La limite attendue permet d'estimer la sensibilité de l'analyse. La limite mesurée est le résultat final de l'analyse.

Afin d'illustrer l'intérêt des distributions des fonctions de vraisemblance, le calcul des limites a également été effectué par une méthode de comptage des événements sélectionnés. Ces limites attendues et mesurées dans les différents canaux simples et combinés,

	S <sub>Tag</sub>	D <sub>Tag</sub>	S <sub>Tag</sub> Veto	D <sub>Tag</sub> Asym
Canal électron				
<i>tb</i>	3.82 ± 0.60	1.10 ± 0.25	3.31 ± 0.51	1.62 ± 0.35
<i>tqb</i>	7.27 ± 1.33	0.62 ± 0.15	6.91 ± 1.25	0.99 ± 0.23
<i>t<math>\bar{t}</math> → l + jets</i>	47.11 ± 10.46	13.60 ± 3.70	40.39 ± 8.85	20.33 ± 5.39
<i>t<math>\bar{t}</math> → ll</i>	13.00 ± 2.87	3.78 ± 1.04	11.28 ± 2.45	5.50 ± 1.47
<i>Wbb</i>	27.00 ± 2.03	5.20 ± 1.16	24.15 ± 1.64	8.06 ± 1.64
<i>Wjj</i>	114.04 ± 11.52	2.38 ± 0.61	111.55 ± 11.04	4.87 ± 1.05
<i>WW</i>	2.30 ± 0.54	0.01 ± 0.004	2.28 ± 0.54	0.03 ± 0.008
<i>WZ</i>	2.26 ± 0.61	0.48 ± 0.14	1.99 ± 0.51	0.74 ± 0.22
Multijet	22.60 ± 3.19	0.61 ± 0.18	21.79 ± 3.66	1.41 ± 0.33
Canal muon				
<i>tb</i>	3.49 ± 0.61	1.05 ± 0.25	3.00 ± 0.51	1.55 ± 0.37
<i>tqb</i>	6.55 ± 1.32	0.57 ± 0.14	6.22 ± 1.25	0.90 ± 0.22
<i>t<math>\bar{t}</math> → l + jets</i>	44.37 ± 10.12	13.81 ± 3.99	37.58 ± 8.42	20.60 ± 5.69
<i>t<math>\bar{t}</math> → ll</i>	11.81 ± 2.68	3.61 ± 1.04	10.17 ± 2.26	5.24 ± 1.48
<i>Wbb</i>	16.55 ± 1.32	3.50 ± 0.82	14.67 ± 1.04	5.38 ± 1.16
<i>Wjj</i>	76.48 ± 8.11	1.64 ± 0.42	74.70 ± 7.76	3.42 ± 0.76
<i>WW</i>	2.42 ± 0.60	0.01 ± 0.003	2.40 ± 0.60	0.03 ± 0.009
<i>WZ</i>	2.08 ± 0.49	0.46 ± 0.12	1.82 ± 0.43	0.72 ± 0.72
Multijet	19.13 ± 4.90	1.46 ± 0.25	17.92 ± 4.52	2.67 ± 0.43
Événements de signal (voies <i>s</i> et <i>t</i> ) attendus, avec erreur systématique :				
Electron	11.10 ± 1.93	1.73 ± 0.39	10.22 ± 1.75	2.60 ± 0.58
Muon	10.04 ± 1.92	1.62 ± 0.39	9.22 ± 1.75	2.45 ± 0.58
Événements de bruit de fond (voies <i>s</i> et <i>t</i> ) attendus, avec erreur systématique :				
Electron	228.31 ± 23.26	26.06 ± 5.76	213.43 ± 20.69	40.94 ± 8.45
Muon	171.99 ± 20.26	24.45 ± 5.73	158.47 ± 17.57	37.98 ± 8.39
Événements observés dans les données, avec l'erreur statistique :				
Electron	248.0 ± 15.7	24.0 ± 4.9	229.0 ± 15.1	43.0 ± 6.6
Muon	157.0 ± 12.5	14.0 ± 3.7	138.0 ± 11.7	33.0 ± 5.7

TAB. 4.22: Nombres d'événements observés dans les données et attendus pour le signal et les différents fonds après sélection dans les canaux électron et muon, pour les quatre schémas d'étiquetage, avec leurs erreurs systématiques.

Limites supérieures attendues/mesurées à 95% CL sur les sections efficaces de production <i>single top</i> (pb)						
	voie <i>s</i>			voie <i>t</i>		
	STag	DTag	Combiné	STag	DTag	Combiné
Etiquetage standard						
Sans incertitudes systématiques						
Canal électron	7.1/12.3	9.0/9.7	5.4/8.8	8.4/15.4	31.7/32.4	8.1/14.7
Canal muon	6.8/5.3	9.1/5.3	5.3/3.2	8.1/6.8	32.5/21.5	7.9/5.8
Combinés	4.8/6.0	6.1/4.4	3.7/3.6	5.7/7.9	24.6/17.3	5.5/7.0
Avec incertitudes systématiques						
Canal électron	16.7/23.2	18.4/19.3	15.4/19.3	21.9/28.4	35.9/36.0	18.6/27.4
Canal muon	18.3/17.2	19.0/13.4	16.1/12.0	24.3/24.6	36.2/34.6	20.3/24.8
Combinés	16.1/19.4	15.8/15.7	14.2/18.4	22.8/25.5	35.7/34.7	16.7/29.0
Etiquetage asymétrique						
Sans incertitudes systématiques						
Canal électron	8.1/12.5	7.4/10.6	5.3/9.2	8.6/14.5	26.8/33.7	8.1/14.5
Canal muon	7.7/5.0	7.5/7.2	5.2/3.7	8.2/5.8	28.0/26.3	7.8/5.4
Combinés	5.4/5.4	5.1/6.4	3.6/4.3	5.8/6.6	19.2/23.1	5.5/6.6
Avec incertitudes systématiques						
Canal électron	17.7/22.5	16.1/20.4	15.2/19.4	20.8/26.4	35.1/36.5	18.2/24.1
Canal muon	19.7/14.7	19.2/17.8	16.8/14.2	24.6/20.6	35.9/35.5	20.2/15.7
Combinés	17.4/18.9	14.9/18.2	13.6/17.3	19.1/25.0	34.7/35.9	15.7/21.0

TAB. 4.23: Limites supérieures attendues/mesurées à 95% CL sur les sections efficaces de production *single top* en picobarn, calculées à partir de l'acceptance du signal et du nombre d'événements de bruit de fond attendus après sélection pour les schémas d'étiquetage standard (en haut) et asymétriques (en bas).

avec et sans prise en compte des incertitudes systématiques sont regroupées dans le tableau 4.23.

Les limites calculées à partir des distributions des fonctions de vraisemblance sont données dans le tableau 4.24 pour les schémas d'étiquetage standard (STag et DTag) et asymétriques (STagVeto et DTagAsym). Les densités de probabilité bayésienne a posteriori, décrites dans l'annexe E, permettent de calculer ces limites. Elles sont données sur les figures 4.37 à 4.40.

Les limites attendues obtenues avec l'étiquetage standard ou asymétrique sont similaires après combinaison des canaux. Néanmoins l'étiquetage asymétrique donne de meilleurs résultats pour les événements doublement étiquetés et des résultats comparables entre STag et STagVeto. L'étiquetage asymétrique permet d'augmenter la statistique d'événements doublement étiquetés et offre une meilleure séparation entre les voies *s* et *t*. Les erreurs systématiques sur les poids d'étiquetage sont plus faibles, en

Limites supérieures attendues/mesurées à 95% CL sur les sections efficaces de production <i>single top</i> (pb)						
	voie <i>s</i>			voie <i>t</i>		
	S <sub>Tag</sub>	D <sub>Tag</sub>	Combiné	S <sub>Tag</sub>	D <sub>Tag</sub>	Combiné
Etiquetage standard						
Sans incertitudes systématiques						
Canal électron	5.6/6.3	6.7/7.5	4.1/5.0	5.7/6.9	9.3/9.3	5.5/6.6
Canal muon	5.5/5.3	7.0/3.8	4.2/2.7	5.7/4.3	9.3/8.6	5.4/3.6
Combinés	3.8/4.1	4.7/3.4	2.8/2.5	3.8/3.7	9.2/8.2	3.7/3.2
Avec incertitudes systématiques						
Canal électron	6.3/6.5	7.4/8.8	4.8/5.5	6.5/7.2	29.8/29.1	6.32/7.3
Canal muon	6.1/8.8	7.7/4.2	4.7/4.5	6.4/6.7	30.9/18.1	6.23/5.9
Combinés	4.5/5.9	5.1/4.1	<b>3.4/3.9</b>	4.4/5.0	22.5/15.4	<b>4.3/5.0</b>
Etiquetage asymétrique						
Sans incertitudes systématiques						
Canal électron	6.3/5.9	5.7/8.0	4.1/5.2	5.7/7.0	9.2/9.4	5.5/7.0
Canal muon	6.1/5.1	5.9/6.1	4.1/3.7	5.7/3.8	9.2/8.8	5.4/3.3
Combinés	4.3/3.6	3.9/5.4	2.8/3.2	3.9/3.4	9.0/8.8	3.7/3.1
Avec incertitudes systématiques						
Canal électron	7.1/6.2	6.2/9.4	4.7/5.8	6.6/7.4	24.7/29.4	6.3/7.6
Canal muon	6.9/8.3	6.4/7.4	4.7/6.4	6.5/5.6	25.9/20.4	6.2/5.0
Combinés	5.0/5.2	4.4/6.8	<b>3.3/5.0</b>	4.4/4.6	17.6/17.9	<b>4.3/4.4</b>

TAB. 4.24: Limites supérieures attendues/mesurées à 95% CL sur les sections efficaces de production *single top* en picobarn, calculées à partir des distributions des fonctions de vraisemblance du signal et des différents bruits de fond après sélection pour les schémas d'étiquetage standard (en haut) et asymétriques (en bas). Les résultats finals obtenus en combinant tous les canaux sont indiqués en gras.

particulier pour le double étiquetage, néanmoins cet effet est dilué par la forme des distributions. Avec l'augmentation de la statistique, les événements doublement étiquetés contribueront de manière plus significative à la limite totale. L'étiquetage asymétrique pourra alors améliorer les limites combinées. Pour ces raisons, le résultat avec étiquetage asymétrique a été retenu comme résultat final de notre analyse.

L'optimisation du choix des points de fonctionnement de l'algorithme *JLIP* constituerait l'étape suivante de cette analyse. A priori, l'optimisation de l'analyse passe par la minimisation de la limite attendue. Ceci implique de reproduire cette analyse pour différents couples de points de fonctionnement, en prenant en compte les modifications des erreurs systématiques. Cette procédure a été utilisée au début de l'analyse, alors que de nombreux éléments (facteurs correctifs, poids de déclenchement) n'étaient pas encore finalisés. Les résultats obtenus sont peu fiables. Seules les limites par comptage

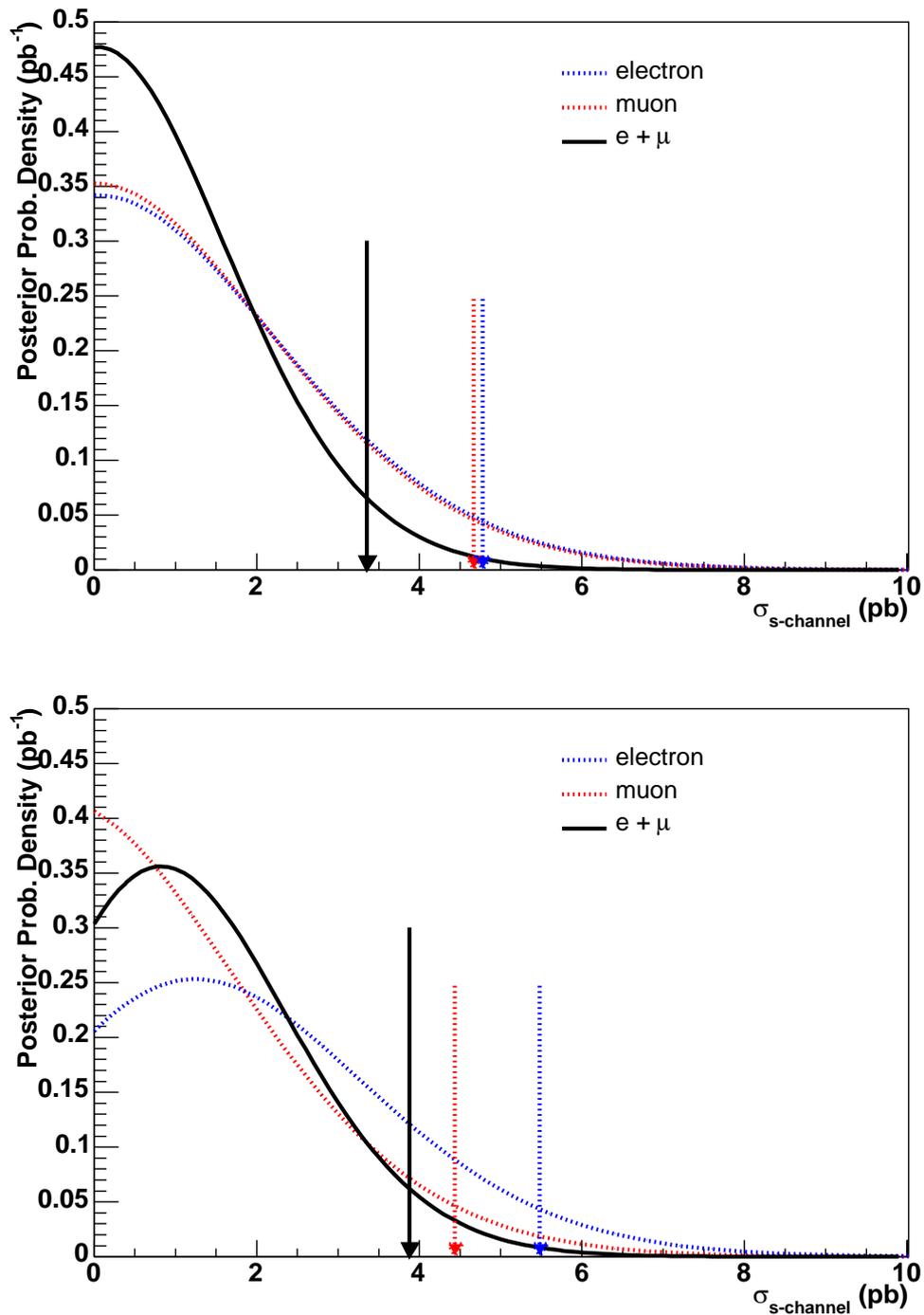


FIG. 4.37: Densités de probabilité bayésienne a posteriori de la section efficace *single top* dans la voie *s* obtenues en combinant les schémas d'étiquetage standard STag et DTag. Le canal électron est indiqué en bleu, le canal muon en rouge et la combinaison des deux canaux en noir. Les flèches indiquent les limites à 95% de niveau de confiance.

En haut : limites attendues.

En bas : limites mesurées.

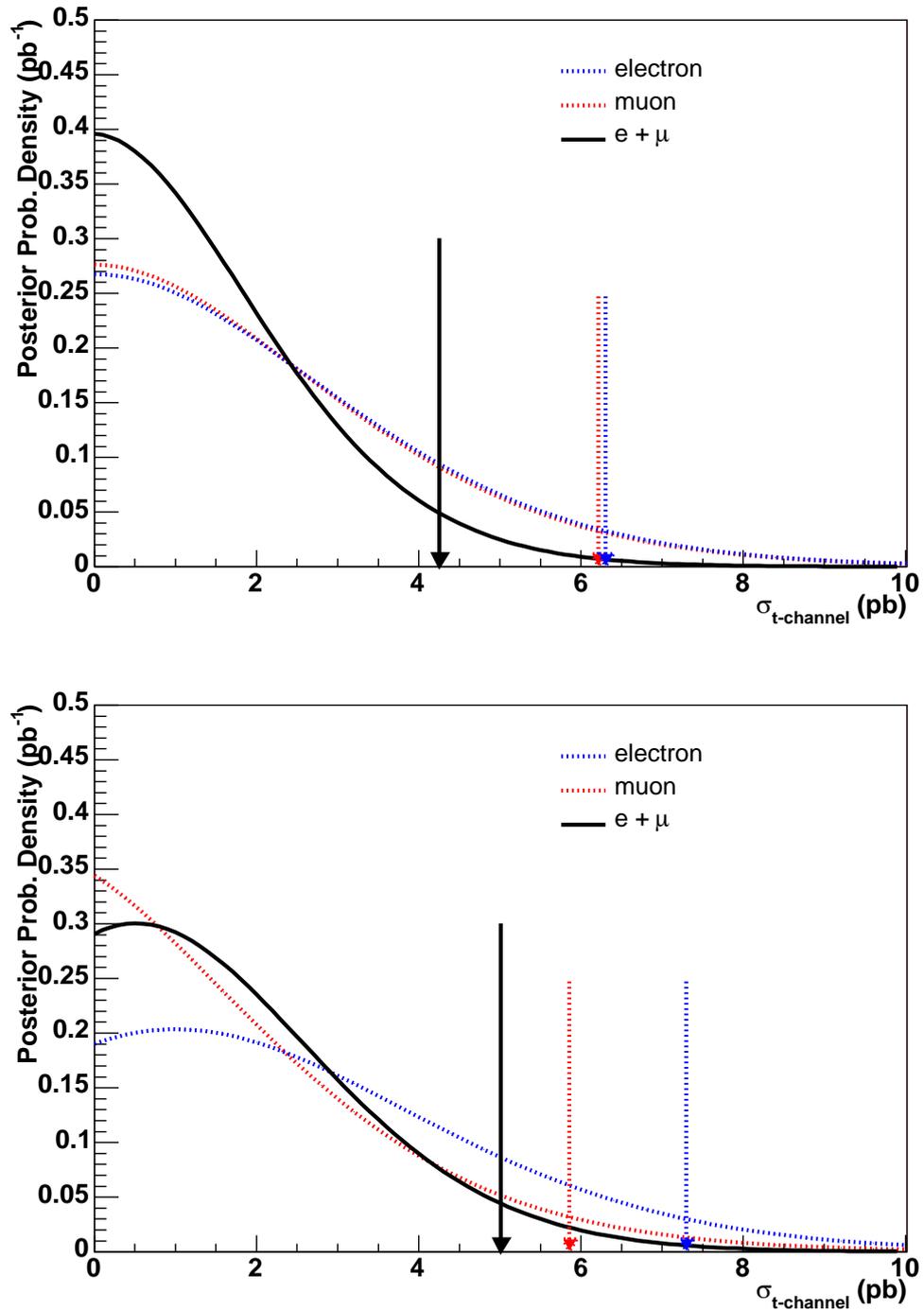


FIG. 4.38: Densités de probabilité bayésienne a posteriori de la section efficace *single top* dans la voie *t* obtenues en combinant les schémas d'étiquetage standard STag et DTag. Le canal électron est indiqué en bleu, le canal muon en rouge et la combinaison des deux canaux en noir. Les flèches indiquent les limites à 95% de niveau de confiance.

En haut : limites attendues.

En bas : limites mesurées.

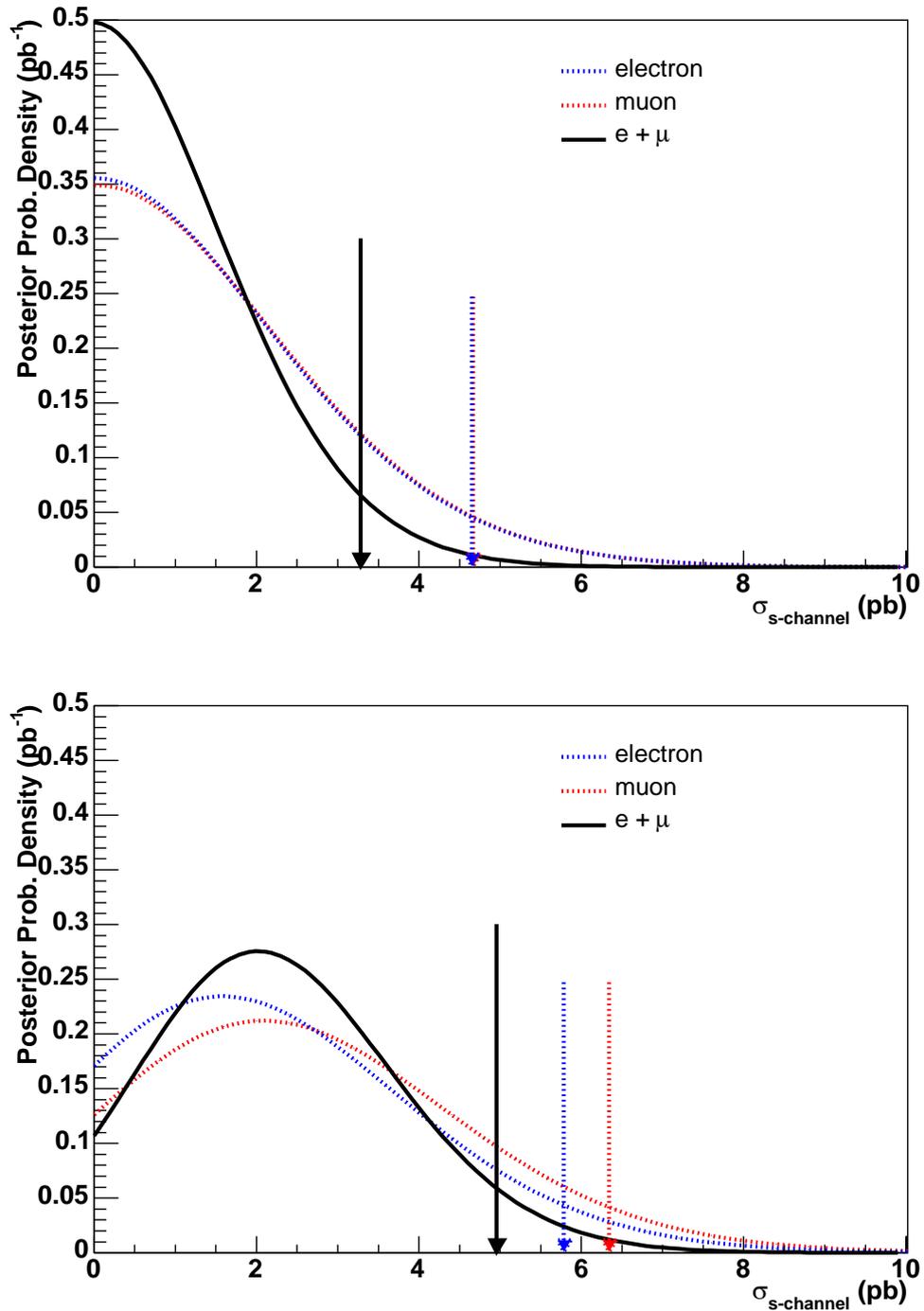


FIG. 4.39: Densités de probabilité bayésienne a posteriori de la section efficace *single top* dans la voie *s* obtenues en combinant les schémas d'étiquetage asymétriques STagVeto et DTagAsym. Le canal électron est indiqué en bleu, le canal muon en rouge et la combinaison des deux canaux en noir. Les flèches indiquent les limites à 95% de niveau de confiance.

En haut : limites attendues.

En bas : limites mesurées.

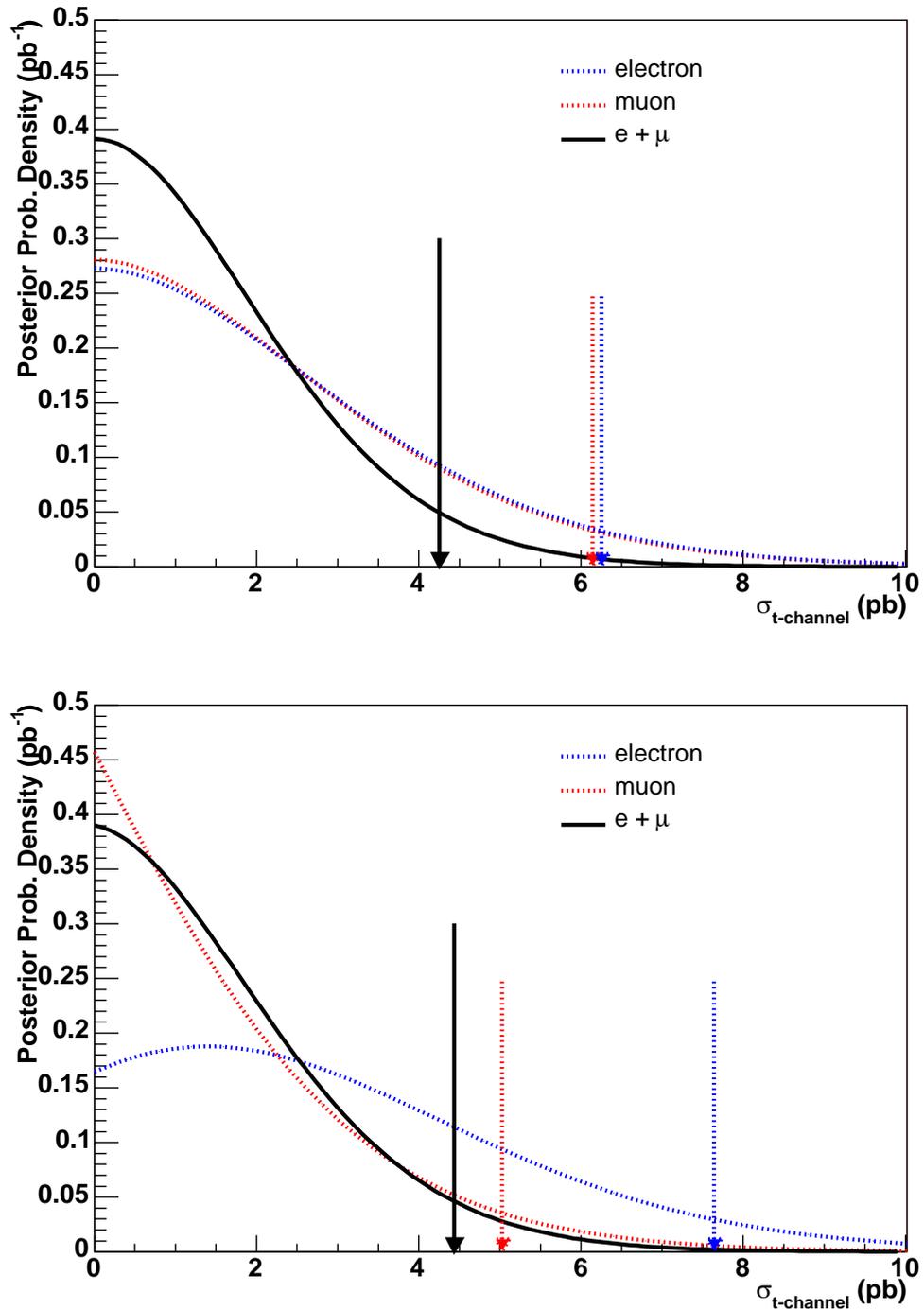


FIG. 4.40: Densités de probabilité bayésienne a posteriori de la section efficace *single top* dans la voie *t* obtenues en combinant les schémas d'étiquetage asymétriques STagVeto et DTagAsym. Le canal électron est indiqué en bleu, le canal muon en rouge et la combinaison des deux canaux en noir. Les flèches indiquent les limites à 95% de niveau de confiance.

En haut : limites attendues.

En bas : limites mesurées.

sont calculées dans la voie  $t$  (étiquetage simple) et la voie  $s$  (étiquetage double) en n'incluant que des erreurs systématiques partielles, dont celles sur les poids d'étiquetage. Ces résultats préliminaires sont présentés dans le tableau 4.25. L'amélioration des limites apportée par les schémas asymétriques est clairement mise en évidence. Il existe peu de différences entre les divers modes asymétriques. La combinaison *Tight/Loose* a finalement été retenue comme le meilleur compromis.

	voie $t$ , simple étiquetage		voie $s$ , double étiquetage	
	électron	muon	électron	muon
Schémas standard				
<i>UltraTight</i>	14.4	12.5	17.6	18.9
<i>Tight</i>	14.1	12.8	15.0	16.7
<i>Medium</i>	14.3	12.8	13.9	15.2
Schémas asymétriques				
<i>UltraTight/Loose</i>	13.2	11.0	14.6	13.6
<i>Tight/Loose</i>	13.4	11.7	14.1	13.8
<i>Medium/Loose</i>	13.8	11.9	14.0	13.6

TAB. 4.25: Limites attendues pour différents schémas d'étiquetage. Ces limites ont été déterminées avant finalisation de l'analyse.

L'optimisation complète, sur les sélections finales, en utilisant les fonctions de vraisemblance et toutes les combinaisons de points de fonctionnement, serait longue à mettre en œuvre et n'a pas été développée au cours de cette thèse.

#### 4.10.2 Limites à deux dimensions et physique au-delà du MS

Il est également possible de déterminer simultanément les limites sur les sections efficaces dans les voies  $s$  et  $t$ . Pour ce faire les filtres  $W$ +jets et  $t\bar{t}$  sont combinés. Comme les filtres sont des fonctions de vraisemblance, elles peuvent s'interpréter comme les probabilités d'être compatible avec le signal ou le bruit de fond. Pour chaque filtre, le signal a des valeurs voisines de l'unité alors que le fond tend vers 0. La topologie des événements *single top* étant intermédiaire entre  $W$ +jets et  $t\bar{t}$ , le filtre  $t\bar{t}$  tend à sélectionner également les événements  $W$ +jets, et réciproquement. La manière la plus simple, bien que non optimale mathématiquement, de combiner les deux filtres consiste simplement à multiplier leur valeur :

$$\mathcal{K}_{\text{voie } s} = \mathcal{L}_{\text{voie } s-W+\text{jets}} \times \mathcal{L}_{\text{voie } s-t\bar{t}} \quad (4.42)$$

$$\mathcal{K}_{\text{voie } t} = \mathcal{L}_{\text{voie } t-W+\text{jets}} \times \mathcal{L}_{\text{voie } t-t\bar{t}} \quad (4.43)$$

Les distributions à deux dimensions ( $\mathcal{K}_{\text{voie } s}$ ,  $\mathcal{K}_{\text{voie } t}$ ) de ces deux variables sont alors utilisées pour déterminer une limite de manière analogue aux calculs précédents. Cette

fois, les densités de probabilité a priori et a posteriori dépendent des deux sections efficaces. Les contours correspondant aux limites mesurées et attendues à 68%, 90% et 95% de niveau de confiance sont donnés sur la figure 4.41. Ces limites à deux dimensions

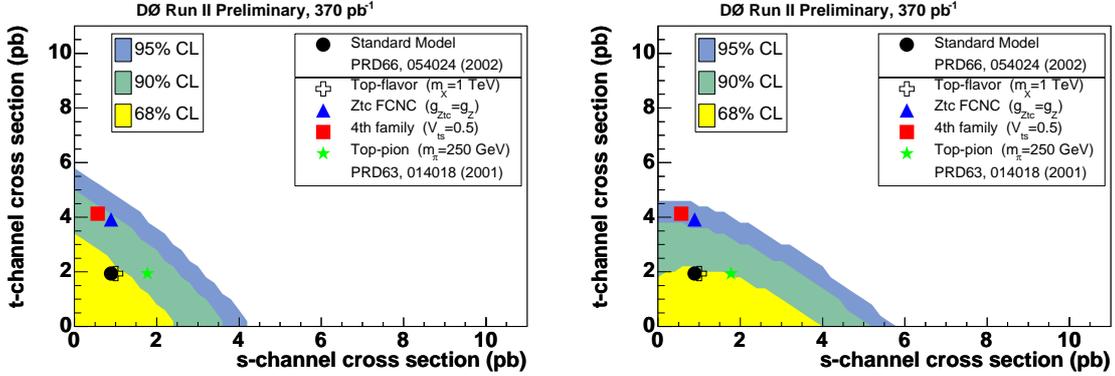


FIG. 4.41: Contours des limites attendues (à gauche) et mesurées (à droite) à 68%, 90% et 95% de niveau de confiance dans le plan  $(\sigma_s, \sigma_t)$  obtenues en combinant les schémas d'étiquetage asymétriques STagVeto et DTagAsym. Les différents points indiquent les valeurs attendues dans le MS (disque noir) et pour différents modèles au-delà du MS (voir texte).

permettent également de contraindre certains modèles de physique au-delà du Modèle Standard. Quelques exemples ont été étudiés dans la référence [104] et les points correspondants dans l'espace  $(\sigma_{\text{voie } s}, \sigma_{\text{voie } t})$  sont indiqués sur la figure 4.41. Ces modèles introduisent soit de nouvelles particules, soit de nouveaux couplages. Nous donnons ici une description succincte de ces différents modèles :

- *Boson de jauge supplémentaire.* Les théories de grande unification incluent le MS dans une symétrie de jauge plus large (groupes  $SU(5)$ ,  $SO(10)$  ou  $E_6$  par exemple). Cette symétrie serait brisée à basse énergie et se réduirait aux symétries du MS. Une telle symétrie de jauge introduit naturellement des nouveaux bosons de jauge massifs,  $W'$  et  $Z'$  qui se mélangent avec les bosons de jauge électrofaibles. L'exemple considéré dans [104] est le modèle *top-flavour* qui introduit une nouvelle symétrie  $SU(2)_h$  responsable de la masse élevée du quark top. Le nouveau boson  $W'$  se couple préférentiellement au quark top. Un tel boson de jauge peut augmenter la section efficace dans la voie  $s$  avec un faible impact dans la voie  $t$  ( $m'_W \gg m_W$ ). Sur la figure 4.41, ce modèle est indiqué par une croix blanche pour  $m'_W = 1 \text{ TeV}/c^2$  et un angle de mélange  $\sin^2 \phi = 0.05$ .
- *Boson scalaire supplémentaire.* Les bosons scalaires apparaissent souvent pour briser spontanément une symétrie. Dans certains modèles, dits *technicouleur*, les bosons scalaires tels le boson de Higgs sont décrits par des états liés de deux particules. Dans le modèle *top-color* présenté ici, un état lié  $t\bar{b}$  dit top-pion peut jouer ce rôle. Une telle particule peut contribuer à la production *single top*. Comme

- dans le cas d'un boson  $W'$ , la section efficace dans la voie  $s$  peut être augmentée, sans affecter la section efficace dans la voie  $t$ . Les sections efficaces pour un topion de masse  $m_\pi = 250 \text{ GeV}/c^2$  sont représentées par une étoile verte sur la figure 4.41.
- *Quarks supplémentaires (quatrième famille)*. Le nombre de familles de fermions n'est pas fixé dans le MS. Il est donc possible de rajouter une quatrième famille de quarks ( $t', b'$ ), qui se mélange aux trois autres familles par une généralisation de la matrice CKM. Les conditions d'unitarité de la matrice CKM à trois familles ne sont alors plus respectées. Sans ces conditions d'unitarité, les éléments de la matrice CKM mesurés aujourd'hui sont bien moins contraints. En particulier la contrainte sur  $|V_{ts}|$  devient  $0 < |V_{ts}| < 0.55$  [105] ( $0.034 < |V_{ts}| < 0.046$  [105] pour trois familles). Une valeur de  $|V_{ts}|$  élevée pourrait augmenter la section efficace de production *single top* dans la voie  $t$  car les densités de parton sont plus élevées pour le quark  $s$  que pour le quark  $b$ . Pour satisfaire l'unitarité de la matrice à quatre familles,  $|V_{tb}|$  serait alors réduit, ce qui diminuerait également la section efficace *single top* dans la voie  $s$ . Le carré rouge sur la figure 4.41 correspond au cas extrême,  $|V_{ts}| = 0.5$  et  $|V_{tb}| = 0.84$ . Ces contraintes sur  $|V_{ts}|$  sont celles données dans [104] qui date de l'année 2001. Les contraintes récentes sur  $|V_{ts}|$ , sans condition d'unitarité, sont  $0 < |V_{ts}| < 0.11$  [5], ce qui exclut a priori ce cas particulier.
  - *Couplages supplémentaires*. Un dernier modèle envisage l'existence de couplages supplémentaires qui n'interviendraient qu'au delà de l'échelle d'énergie de la brisure électrofaible, notamment des couplages par courant neutre modifiant la saveur (**FCNC** : **F**lavour **C**hanging **N**eutral **C**urrents) via des vertex  $Ztc$ ,  $\gamma tc$  ou  $g tc$ . L'effet de ces nouveaux couplages pourrait se faire particulièrement sentir dans le secteur du quark top, dans la production de *single top* par exemple. Dans la voie  $t$ , la présence d'un quark  $c$  au lieu d'un quark  $b$  dans la voie d'entrée augmenterait la section efficace. Sur la figure 4.41 les sections efficaces correspondant à un couplage  $Ztc$  de même amplitude que le couplage électrofaible du MS est indiqué par un triangle bleu.

Les modèles de quatrième famille et de couplage par courant neutre (**FCNC**) présentés ci-dessus sont exclus à 90% de niveau de confiance par notre analyse.

En calculant la valeur moyenne de la densité de probabilité a posteriori à deux dimensions selon  $\sigma_s$  ou  $\sigma_t$ , on peut également obtenir des densités à une dimension :

$$P(\sigma_{s(t)}) = \int_0^\infty \sigma_{t(s)} P(\sigma_s, \sigma_t) d\sigma_{t(s)}, \quad (4.44)$$

et calculer des limites séparées sur les voies  $s$ ,  $t$  et  $s+t$  combinées. Ces limites inclusives diffèrent des limites calculées précédemment : elles ne sont plus dépendantes du MS et n'ont pas été dérivées avec exactement les mêmes variables discriminantes (combinaison des filtres). Ces limites sont données dans le tableau 4.26. Elles sont illustrées sur la figure 4.42. On constate que les limites attendues sont proches de celles estimées précé-

demment, ce qui indique que cette méthode est consistante avec la précédente. De fait, l'impact du bruit de fond voie  $t$  dans l'analyse voie  $s$  et réciproquement est faible. La méthode inclusive est cependant plus sensible à des fluctuations statistiques dans les données observées, c'est pourquoi les résultats finals de l'analyse, rendus publics par la collaboration  $D\bar{O}$ , sont les limites dépendantes du  $MS$ .

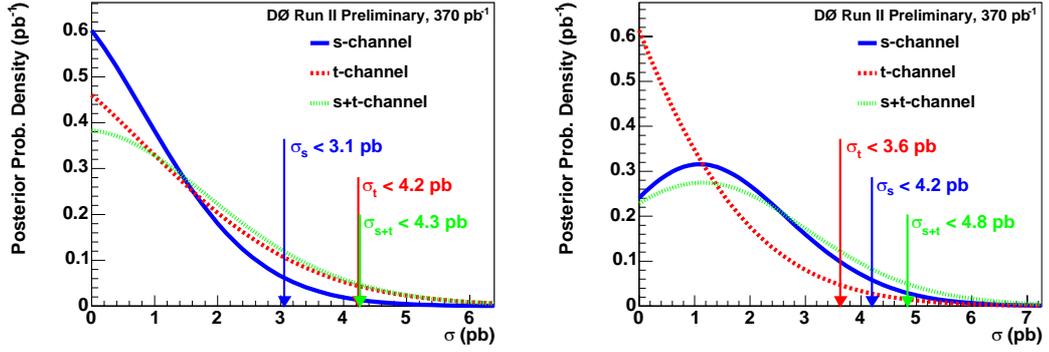


FIG. 4.42: Densités de probabilité bayésienne a posteriori pour les limites inclusives attendues (à gauche) et mesurées (à droite), pour la voie  $s$  (en bleu), la voie  $t$  (en rouge) et la combinaison  $s + t$  (en vert). Les flèches indiquent les limites à 95% de niveau de confiance.

	voie $s$	voie $t$	voies $s + t$
Limite attendue	3.1 pb	4.2 pb	4.3 pb
Limite mesurée	4.2 pb	3.6 pb	4.8 pb

TAB. 4.26: Limites inclusives obtenues à partir des densités de probabilité à deux dimensions.

### 4.10.3 Comparaison avec l'analyse par réseaux de neurones de $D\bar{O}$

Les résultats présentés dans cette thèse améliorent les limites publiées par  $D\bar{O}$  [83], obtenues à partir de réseaux de neurones. Les deux analyses possèdent la même structure, les principales différences résidant dans l'algorithme d'identification des jets  $b$  et la méthode d'analyse multivariable. La statistique a également été accrue et certaines erreurs systématiques, notamment sur la correction en énergie des jets et sur les  $TRF$  ont été réduites. Ces différences sont résumées dans le tableau 4.27.

Afin de comparer la sensibilité des deux analyses, les nombres d'événements attendus dans notre analyse sont renormalisés pour une luminosité intégrée de  $230 \text{ pb}^{-1}$ . Les erreurs statistiques sont réajustées en conséquence. Les limites attendues sont alors recalculées avec et sans incertitudes systématiques. Les résultats sont présentés dans le

tableau 4.28. Les limites sans incertitudes systématiques permettent de comparer la sensibilité des méthodes d'analyse. Dans la mesure où les algorithmes *SVT* et *JLIP* ont des performances voisines, ces limites permettent de comparer les deux techniques multivariées (fonctions de vraisemblance et réseaux de neurones). Les deux analyses ont des sensibilités voisines. Les limites avec erreurs systématiques permettent de quantifier l'impact de la réduction des incertitudes systématiques sur les limites. Le gain est de l'ordre de 0.3 pb. L'essentiel du gain sur les limites entre les deux analyses provient de l'accroissement de la statistique.

	Résultat publié [83]	Nouveau résultat préliminaire
Luminosité	230 pb <sup>-1</sup>	370 pb <sup>-1</sup>
Étiquetage des jets <i>b</i>	<i>SVT</i>	<i>JLIP</i>
Schéma d'étiquetage	Standard ( <i>Tight</i> )	Asymétrique ( <i>Tight/Loose</i> )
Analyse multivariée	Réseaux de neurones	Fonctions de vraisemblance

TAB. 4.27: Principales différences entre l'analyse publiée par DØ et le résultat préliminaire décrit dans cette thèse.

	Résultat publié [83]	Nouveau résultat préliminaire
	voie <i>s</i>	
Sans erreurs systématiques	3.6 pb	3.6 pb
Avec erreurs systématiques	4.4 pb	4.1 pb
	voie <i>t</i>	
Sans erreurs systématiques	4.6 pb	4.8 pb
Avec erreurs systématiques	5.8 pb	5.5 pb

TAB. 4.28: Comparaisons des limites bayésiennes attendues à 95% de niveau de confiance sans et avec erreurs systématiques pour une luminosité intégrée de 230 pb<sup>-1</sup> entre l'analyse publiée par DØ et le résultat préliminaire décrit dans cette thèse.

## 4.11 Conclusion

Pour cette étude, environ 370 pb<sup>-1</sup> de données du Run II de l'expérience DØ ont été analysées. L'analyse repose sur une sélection topologique (2 à 4 jets dont 1 ou 2 jets *b*, un lepton chargé, de l'énergie manquante) et sur la construction de fonctions de vraisemblance pour séparer le signal des principaux bruits de fond. Aucune évidence de signal n'est observée et des limites à 95% de niveau de confiance sur les sections efficaces sont déterminées en utilisant une méthode bayésienne. De nouvelles limites sur les sections efficaces de production du quark top par interaction électrofaible dans

les voies  $s$  et  $t$  ont été obtenues. Ce résultat améliore les limites précédentes publiées par DØ et CDF [83, 85]. Les densités de probabilité pour chaque section efficace sont données sur la figure 4.43 et les limites mesurées à 95% de niveau de confiance sont :

$$\begin{cases} \sigma_s < 5.0 \text{ pb}, \\ \sigma_t < 4.4 \text{ pb}. \end{cases} \quad (4.45)$$

Les valeurs les plus probables pour les sections efficaces de production *single top* sont :

$$\begin{cases} \sigma_s = 2.1 \pm 1.6 \text{ pb}, \\ \sigma_t = 0.0 \pm 1.8 \text{ pb}. \end{cases} \quad (4.46)$$

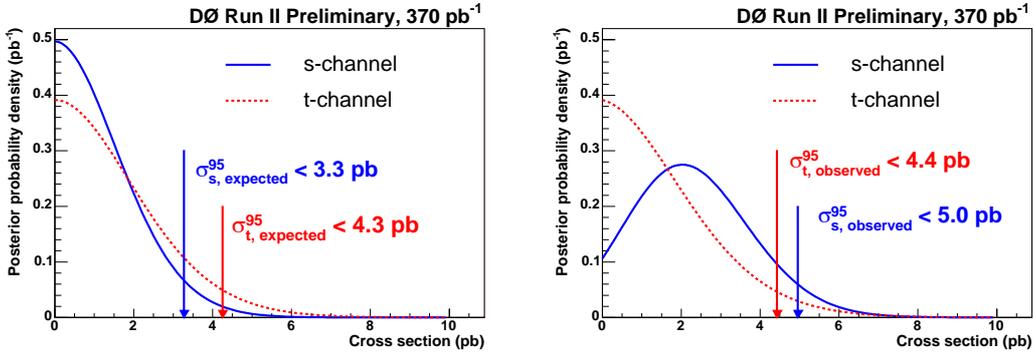


FIG. 4.43: Densités de probabilité bayésienne a posteriori de la section efficace attendues (à gauche) et mesurées (à droite) pour le résultat final. La voie  $s$  est représentée en bleu, la voie  $t$  en rouge. Les flèches indiquent les limites à 95% de niveau de confiance.

La section efficace mesurée dans la voie  $s$  possède une significativité de 1.3 écart standard, ce qui est largement insuffisant pour prétendre à une observation (3 écarts standard). Il n'y a aucune évidence de signal dans la voie  $t$ . Une statistique plus importante ainsi que la réduction des erreurs systématiques sont nécessaires pour mettre en évidence la production de *single top*. Avec l'analyse actuelle, une luminosité intégrée de plusieurs  $fb^{-1}$  sera nécessaire pour observer ce signal. Une estimation de l'évolution de la sensibilité des analyses actuelles (DØ et CDF) est donnée sur la figure 4.44 [106]. Le Tevatron devrait pouvoir observer (mesure à 3 écarts standard) les deux modes de production dans les voies  $s$  et  $t$ . La découverte (mesure à 5 écarts standard) dans la voie  $t$  devrait également être possible. De nombreuses études sont actuellement en cours afin d'augmenter l'acceptance du signal en modifiant les critères de sélection. L'amélioration des performances de l'identification des jets  $b$  (combinaison des différents algorithmes par un réseau de neurones) devrait également contribuer à l'augmentation de l'acceptance. De nombreuses erreurs systématiques devraient diminuer au cours du temps avec l'augmentation de la luminosité (normalisations et pondérations dérivées dans les données) et une meilleure description du détecteur dans la simulation. Enfin des méthodes discriminantes plus élaborées sont également en développement. Néanmoins, la faible différence

de sensibilité entre les analyses avec fonctions de vraisemblance (qui ne tiennent pas compte des corrélations entre variables) et avec réseaux de neurones (utilisant plus de variables et utilisant leurs corrélations) peut faire craindre que la sensibilité maximale ait déjà été atteinte.

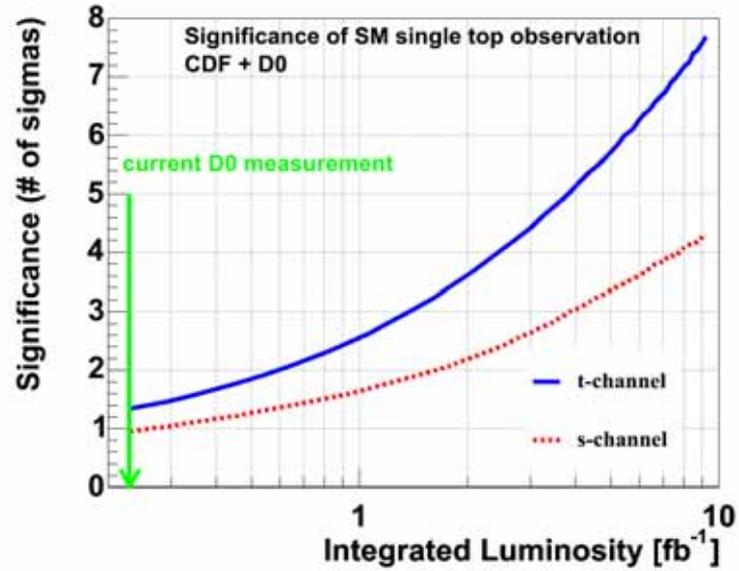


FIG. 4.44: Projection de la signification des analyses de recherche du *single top* en fonction de la luminosité intégrée pour les voies *s* (ligne pointillée rouge) et *t* (ligne continue bleue).

Now my charms are all o'erthrown,  
And what strength I have's mine own,  
Which is most faint : now, 'tis true, [...]

William Shakespeare, “*The Tempest*”, Epilogue

## Conclusion

Cette thèse présente une analyse portant sur la recherche de la production électrofaible du quark top (*single top*) dans l'expérience  $D\bar{O}$  au Run II du Tevatron. Ce mode de production du quark top n'a jamais été observé et les limites sur les sections efficaces des deux processus dominants (voies  $s$  et  $t$ ) au Tevatron ont été améliorées.

Les événements sont sélectionnés en recherchant un électron ou un muon isolé, associé à de l'énergie transverse manquante. La présence de deux à quatre jets, dont au moins un jet de quark  $b$ , est également requise. Pour améliorer la sensibilité de l'analyse à chacun des deux processus étudiés, les événements contenant exactement un jet et au moins deux jets  $b$  identifiés sont traités séparément.

L'algorithme *JLIP* utilise le temps de vol élevé des hadrons beaux pour identifier les jets issus de la fragmentation d'un quark  $b$ . La signification en paramètre d'impact des traces chargées reconstruites par le trajectographe de  $D\bar{O}$  est convertie en probabilité pour le jet auquel elles appartiennent de provenir du vertex primaire de l'interaction. La calibration et la validation, éléments essentiels de l'élaboration de cet algorithme, représentent une contribution majeure développée dans ce travail de thèse. Une étude détaillée de l'efficacité et du taux de mauvaise identification de l'algorithme *JLIP* a été menée. Une nouvelle méthode de détermination de l'efficacité d'étiquetage à partir des données réelles et utilisant la simulation au minimum, a été développée à cet effet. Cette méthode est devenue un outil standard de la collaboration  $D\bar{O}$ . L'utilisation de l'algorithme *JLIP*, une fois certifié, ne se limite pas à notre étude du *single top* mais s'étend à de nombreuses analyses de  $D\bar{O}$ .

L'analyse finale repose sur la construction de fonctions de vraisemblance combinant plusieurs variables cinématiques discriminantes afin de maximiser la séparation entre le signal et les bruits de fond. La distribution combinée de ces fonctions de vraisemblance est utilisée dans le calcul des limites par une méthode bayésienne. L'évaluation des différentes composantes du bruit de fond à partir de la simulation et des données réelles ainsi que le calcul des erreurs systématiques associées à chaque fond constitue un point crucial de l'analyse, en particulier l'estimation des pondérations permettant de modéliser la réponse du système de déclenchement et l'effet de l'identification des jets de quarks  $b$  dans la simulation.

Les limites supérieures mesurées à 95% de niveau de confiance sur les sections efficaces de production électrofaible du quark top sont de 5 pb dans la voie  $s$  et 4.3 pb dans la voie  $t$ . Ces limites sont, à ce jour, les meilleures contraintes existantes sur ces

sections efficaces. Ce résultat a été rendu public par la collaboration DØ en juillet 2005 et présenté à diverses conférences. Elles devraient être améliorées par DØ et CDF avec l'augmentation de la statistique disponible. L'observation à trois écarts standard du processus *single top* nécessitera entre 1.5 et 3 fb<sup>-1</sup> de luminosité intégrée, ce qui devrait être disponible d'ici la fin 2007. L'amélioration de la sensibilité des analyses est également à l'étude par exemple en utilisant un algorithme d'identification de jets de quark *b* plus performant, tel l'algorithme de réseau de neurones développé à DØ. Certaines incertitudes systématiques (liées à la correction en énergie des jets, aux divers facteurs correctifs, à l'efficacité d'identification des jets *b*,...) dépendent de la statistique et devraient donc être réduites avec l'augmentation de la luminosité.

Si le processus *single top* est observé à Tevatron, les sections efficaces mesurées conserveront une erreur totale importante qui se repercutera sur la détermination de  $|V_{tb}|$ . Ces mesures ne permettront guère de contraindre des modèles de nouvelle physique. Ces études pourront être menées au LHC, où les sections efficaces attendues sont plus élevées.

# A

## Incertitudes systématiques sur la combinaison des fonds $W$ +jets et multijet

### A.1 L'erreur systématique sur la méthode de la matrice

La méthode de la matrice décrite au § 4.7.3.1 permet de déterminer le nombre d'événements contenant un vrai ou un "faux" lepton isolé, c'est-à-dire les nombres d'événements contenant un boson  $W$  (signal et bruits de fond physiques :  $W$ +jets,  $t\bar{t}$ , dibosons) et multijet. Elle consiste en la résolution d'un système de deux équations à deux inconnues. Les termes utilisés pour la normalisation des fonds  $W$  + jets et multijet provenant de la méthode de la matrice sont :

$$N_{faux}^s = \varepsilon_{faux} N_{faux} = \varepsilon_{faux} \frac{N^s - \varepsilon_{faux} \cdot N}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (\text{A.1})$$

$$N_{vrai}^s = \varepsilon_{vrai} N_{vrai} = \varepsilon_{vrai} \frac{\varepsilon_{vrai} \cdot N - N^s}{\varepsilon_{vrai} - \varepsilon_{faux}} \quad (\text{A.2})$$

Les erreurs systématiques sur ces deux termes ont deux origines :

- l'erreur statistique sur  $N$  et  $N^s$
- l'erreur systématique sur les efficacités  $\varepsilon_{faux}$  et  $\varepsilon_{vrai}$ . Cette erreur est liée à la méthode et à la statistique du lot utilisé pour déterminer chaque efficacité.

Pour pouvoir calculer l'erreur statistique, on pose :

$$N^{\bar{s}} = N - N^s \quad (\text{A.3})$$

Les nombres d'événements  $N^{\bar{s}}$  et  $N^s$  sont statistiquement indépendants. Pour  $N^{\bar{s}}$  et  $N^s$  suffisamment grand ( $>25$ ), l'erreur statistique est donnée par  $\Delta N^{\bar{s}} = \sqrt{N^{\bar{s}}}$  et  $\Delta N^s = \sqrt{N^s}$ . L'erreur statistique sur  $N_{faux}^s$  et  $N_{vrai}^s$  s'écrit alors :

$$\Delta_{stat} N_{faux}^s = \sqrt{\left(\frac{\partial N_{faux}^s}{\partial N^s}\right)^2 N^s + \left(\frac{\partial N_{faux}^s}{\partial N^{\bar{s}}}\right)^2 N^{\bar{s}}} \quad (\text{A.4})$$

$$\Delta_{stat} N_{vrai}^s = \sqrt{\left(\frac{\partial N_{vrai}^s}{\partial N^s}\right)^2 N^s + \left(\frac{\partial N_{vrai}^s}{\partial N^{\bar{s}}}\right)^2 N^{\bar{s}}} \quad (\text{A.5})$$

L'erreur systématique s'écrit en fonction des incertitudes sur les efficacités  $\Delta \varepsilon_{faux}$  et  $\Delta \varepsilon_{vrai}$  :

$$\Delta_{syst} N_{faux}^s = \sqrt{\left(\frac{\partial N_{faux}^s}{\partial \varepsilon_{faux}}\right)^2 (\Delta \varepsilon_{faux})^2 + \left(\frac{\partial N_{faux}^s}{\partial \varepsilon_{vrai}}\right)^2 (\Delta \varepsilon_{vrai})^2} \quad (\text{A.6})$$

$$\Delta_{syst} N_{vrai}^s = \sqrt{\left(\frac{\partial N_{vrai}^s}{\partial \varepsilon_{faux}}\right)^2 (\Delta \varepsilon_{faux})^2 + \left(\frac{\partial N_{vrai}^s}{\partial \varepsilon_{vrai}}\right)^2 (\Delta \varepsilon_{vrai})^2} \quad (\text{A.7})$$

Les erreurs totales sont alors données par :

$$\Delta N_{faux}^s = \sqrt{(\Delta_{stat} N_{faux}^s)^2 + (\Delta_{syst} N_{faux}^s)^2} \quad (\text{A.8})$$

$$\Delta N_{vrai}^s = \sqrt{(\Delta_{stat} N_{vrai}^s)^2 + (\Delta_{syst} N_{vrai}^s)^2} \quad (\text{A.9})$$

Les expressions algébriques de toutes les dérivées partielles sont données dans la référence [107]

Dans le cadre de l'analyse, pour prendre en compte les dépendances des efficacités, la méthode de la matrice est appliquée événement par événement. L'erreur est approximée globalement en utilisant des valeurs moyennes des efficacités et de leurs erreurs :

$$\langle \varepsilon_{faux} \rangle = \frac{N_{faux}^s}{N_{faux}} \quad (\text{A.10})$$

$$\langle \varepsilon_{vrai} \rangle = \frac{N_{vrai}^s}{N_{vrai}} \quad (\text{A.11})$$

Les valeurs numériques des erreurs relatives sur les efficacités sont :

$$\left\langle \frac{\Delta \varepsilon_{faux}}{\varepsilon_{faux}} \right\rangle_{\text{électron}} = 12.7\% \quad (\text{A.12})$$

$$\left\langle \frac{\Delta \varepsilon_{faux}}{\varepsilon_{faux}} \right\rangle_{\text{muon}} = 4.8\% \quad (\text{A.13})$$

$$\left\langle \frac{\Delta \varepsilon_{vrai}}{\varepsilon_{vrai}} \right\rangle_{\text{électron}} = 0.8\% \quad (\text{A.14})$$

$$\left\langle \frac{\Delta \varepsilon_{vrai}}{\varepsilon_{vrai}} \right\rangle_{\text{muon}} = 0.9\% \quad (\text{A.15})$$

## A.2 Incertitude sur la normalisation de la combinaison des fonds $Wjj$ et $Wbb$

Les erreurs sur la normalisation relative de chacun des fonds  $Wjj$  et  $Wbb$  sont corrélées. Le nombre total d'événements  $W + jets$  peut s'écrire à partir des équations 4.33 et 4.34 :

$$\mathcal{Y}_{W+jets} = \mathcal{Y}_{Wjj}^{\text{MC}} + \mathcal{Y}_{Wbb}^{\text{MC}} \quad (\text{A.16})$$

$$= N_W \cdot (S_{Wjj} \langle P_{Wjj} \rangle + S_{Wbb} \langle P_{Wbb} \rangle) \quad (\text{A.17})$$

$$= N_W \cdot ((1 - S_{Wbb}) \langle P_{Wjj} \rangle + S_{Wbb} \langle P_{Wbb} \rangle) \quad (\text{A.18})$$

$$\text{où } \langle P_{Wjj,Wbb} \rangle = \frac{\sum_{N_{Wjj,Wbb}^{\text{select}}} P_{\text{déclenchement}} P_{\text{étiquetage}}}{\sum_{N_{Wjj,Wbb}^{\text{select}}} P_{\text{déclenchement}}} \quad (\text{A.19})$$

$$\text{et } N_W = N_{vrai}^s - N_{t\bar{t}} - N_{WW} - N_{WZ} \quad (\text{A.20})$$

L'erreur sur le fond fond  $W+jets$  combiné, provenant de la normalisation relative est alors :

$$\Delta_{S_{Wbb}} \mathcal{Y}_{W+jets} = \frac{\partial \mathcal{Y}_{W+jets}}{\partial S_{Wbb}} \Delta S_{Wbb} \quad (\text{A.21})$$

$$= N_W | \langle P_{Wbb} \rangle - \langle P_{Wjj} \rangle | \Delta S_{Wbb} \quad (\text{A.22})$$

$$= \left| \frac{\mathcal{Y}_{Wbb}^{\text{MC}}}{S_{Wbb}} - \frac{\mathcal{Y}_{Wjj}^{\text{MC}}}{1 - S_{Wbb}} \right| \Delta S_{Wbb} \quad (\text{A.23})$$

L'erreur relative sur  $S_{Wbb}$  représente notre compréhension de la composition en saveur et de la simulation du fond  $W+jets$  [101]. Sa valeur est choisie de manière conservative à 25%. Les termes soustraits à  $N_{vrai}^s$  dans le calcul de  $N_W$  sont petits et l'on peut les négliger dans le calcul de l'erreur. L'erreur relative de normalisation du fond  $W+jets$  est ainsi :

$$\frac{\Delta_{Norm} \mathcal{Y}_{W+jets}}{\mathcal{Y}_{W+jets}} = \sqrt{\left( \frac{\Delta_{S_{Wbb}} \mathcal{Y}_{W+jets}}{\mathcal{Y}_{W+jets}} \right)^2 + \left( \frac{\Delta N_{vrai}^s}{N_{vrai}^s} \right)^2} \quad (\text{A.24})$$

Cette erreur apparaît dans la ligne "Méthode de la matrice" des tableaux 4.14 à 4.17.

### A.3 Incertitudes systématiques sur la combinaison des fonds $W$ +jets et multijet

Les fonds  $W$ +jets et multijet sont tous deux normalisés sur les données réelles par la méthode de la matrice. Il est important de prendre en compte correctement la corrélation des erreurs car elle conduit à une erreur totale plus faible. La combinaison de ces erreurs est décrite en détail dans la référence [108]. Le nombre total d'événements  $W$ +jets et multijet peut s'écrire :

$$\mathcal{Y}_{W+jets + multijet} = \mathcal{Y}_{W+jets} + \mathcal{Y}_{Multijet} \quad (\text{A.25})$$

L'erreur totale sur la normalisation du fond combiné  $W$ +jets et multijet est alors donnée par :

$$(\Delta_{Norm} \mathcal{Y}_{W+jets + multijet})^2 = \left( \frac{\partial \mathcal{Y}_{W+jets + multijet}}{\partial N^s} \right)^2 N^s \quad (\text{A.26})$$

$$+ \left( \frac{\partial \mathcal{Y}_{W+jets + multijet}}{\partial N^{\bar{s}}} \right)^2 N^{\bar{s}} \quad (\text{A.27})$$

$$+ \left( \frac{\partial \mathcal{Y}_{W+jets + multijet}}{\partial \varepsilon_{faux}} \right)^2 (\Delta \varepsilon_{faux})^2 \quad (\text{A.28})$$

$$+ \left( \frac{\partial \mathcal{Y}_{W+jets + multijet}}{\partial \varepsilon_{vrai}} \right)^2 (\Delta \varepsilon_{vrai})^2 \quad (\text{A.29})$$

$$+ (\Delta_{S_{Wbb}} \mathcal{Y}_{W+jets})^2 \quad (\text{A.30})$$

Les autres composantes de l'erreur totale (pondérations, erreurs statistiques) sont sommées en quadrature avec cette erreur de normalisation. L'expression de chaque dérivée est donnée dans la référence [108].

# B

## Reconstruction du neutrino

Les neutrinos n'interagissent pas dans le détecteur et ne sont pas détectés. Néanmoins la construction de certaines variables cinématiques nécessite de connaître la quadri-impulsion du neutrino, par exemple la masse invariante entre le jet  $b$  de plus haut  $p_T$ , le lepton chargé et le neutrino, qui correspond à la masse du top pour un événement *single top*. Dans un événement contenant un seul boson  $W$  se désintégrant en leptons comme c'est le cas dans le signal recherché, l'énergie transverse manquante peut s'interpréter comme l'impulsion transverse du neutrino. L'énergie totale  $E^\nu$  et l'impulsion selon l'axe du faisceau  $p_z^\nu$  demeurent inconnues. Ces grandeurs peuvent être déterminées en imposant une masse nulle pour le neutrino et la masse du boson  $W$  pour la masse invariante entre le lepton chargé et le neutrino :

$$m_\nu^2 = E^{\nu 2} - \cancel{E}_{T_x}^2 - \cancel{E}_{T_y}^2 - p_z^{\nu 2} \quad (\text{B.1})$$

$$m_W^2 = (E^\ell + E^\nu)^2 - (p_x^\ell + \cancel{E}_{T_x})^2 - (p_y^\ell + \cancel{E}_{T_y})^2 - (p_z^\ell + p_z^\nu)^2 \quad (\text{B.2})$$

$$m_\nu = 0 \text{ GeV}/c^2 \quad (\text{B.3})$$

$$m_W = 80.425 \text{ GeV}/c^2 \quad (\text{B.4})$$

L'équation B.1 permet d'exprimer  $E^{\nu 2}$  en fonction de  $p_z^{\nu 2}$  :

$$E^{\nu 2} = m_\nu^2 + \cancel{E}_{T_x}^2 + \cancel{E}_{T_y}^2 + p_z^{\nu 2} \quad (\text{B.5})$$

En substituant cette valeur dans l'équation B.1, on obtient un équation de degré 2 en  $p_z^\nu$ . En posant :

$$K_\nu^2 = m_\nu^2 + \cancel{E}_T^2 \quad (\text{B.6})$$

$$K_\ell^2 = m_\ell^2 + p_T^\ell{}^2 \quad (\text{B.7})$$

$$K_W^2 = m_W^2 + p_T^W{}^2 \quad (\text{B.8})$$

$$A = 4K_\ell^2 \quad (\text{B.9})$$

$$B = -4p_z^\ell(K_W^2 - K_\ell^2 - K_\nu^2) \quad (\text{B.10})$$

$$C = 4E^{\ell 2} K_\nu^2 - (K_W^2 - K_\ell^2 - K_\nu^2)^2 \quad (\text{B.11})$$

il vient :

$$0 = Ap_z^{\nu 2} + Bp_z^\nu + C \text{ et } p_z^\nu = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \quad (\text{B.12})$$

Si le discriminant est positif,  $B^2 - 4AC > 0$  il existe deux solutions. La plus petite des deux, en valeur absolue, est choisie comme valeur de  $p_z^\nu$ . On vérifie dans des événements simulés  $Wjj$  que ce critère sélectionne la solution correspondant au  $p_z$  du neutrino généré dans environ 70 % des cas. La figure B.1 illustre cette sélection dans la simulation. Si le discriminant est négatif,  $B^2 - 4AC < 0$  la solution est complexe. Dans ce cas la masse transverse lepton-neutrino est supérieure à la masse du  $W$ . Afin d'avoir une valeur utilisable dans l'analyse, la partie réelle de la solution complexe est retenue [109].

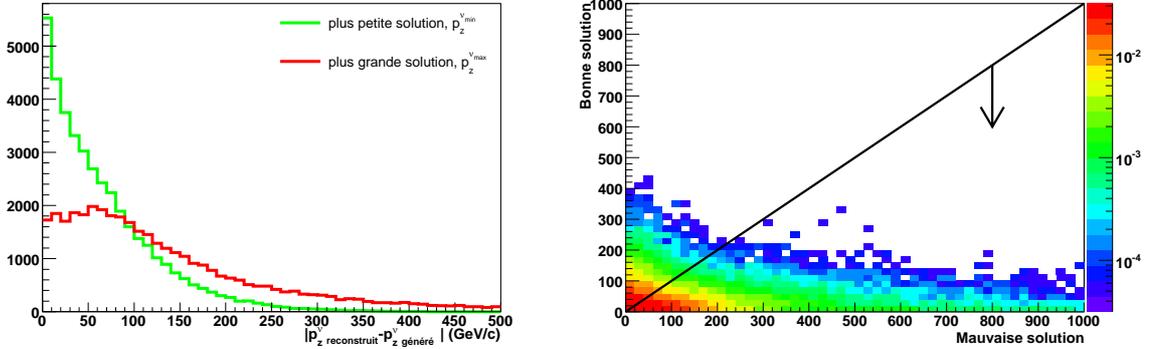


FIG. B.1: Solutions pour  $p_z^\nu$  dans la simulation  $Wjj$ .

A gauche : Différence entre le  $p_z^\nu$  reconstruit et le  $p_z^\nu$  généré pour la plus petite solution (en vert) et la plus grande solution (en rouge).

A droite : Graphe présentant en ordonnée, la valeur absolue de la solution la plus proche du  $p_z$  du neutrino généré ( $|p_z^{bon}|$ ) et en abscisse la valeur absolue de la seconde solution ( $|p_z^{mauvais}|$ ). La droite correspond à  $|p_z^{bon}| = |p_z^{mauvais}|$ .

Dans environ 70% des cas la solution la plus proche du  $p_z^\nu$  généré est la plus petite en valeur absolue, ce qui correspond au domaine en dessous de la droite.

Les distributions de  $p_z'$  obtenues dans les données et la simulation sont présentées sur la figure B.2. Les formes sont en bon accord.

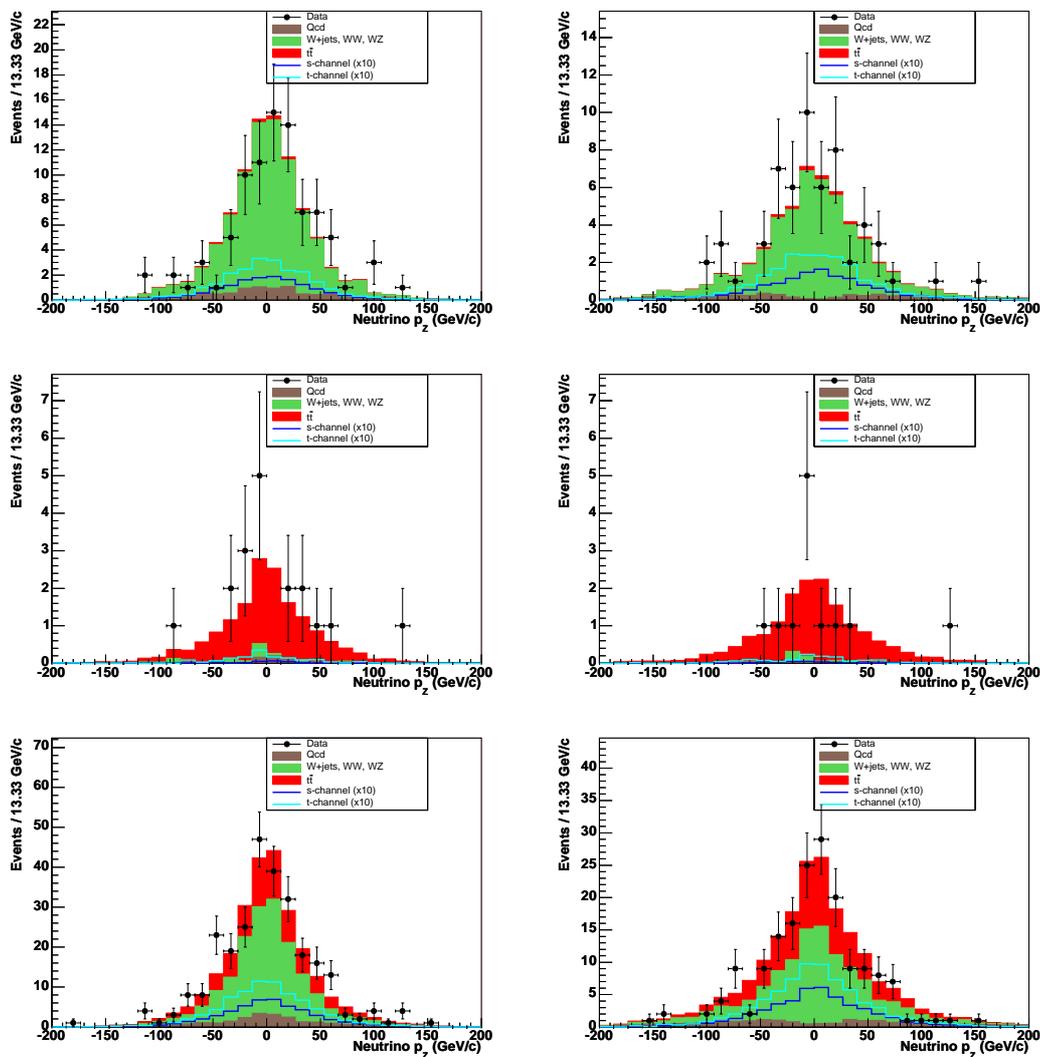


FIG. B.2: Comparaison entre données sélectionnées (disques noirs), le signal (lignes bleue et cyan) et les bruits de fond estimés ( $t\bar{t}$  en rouge,  $W$ +jets en vert, multijet en marron) pour le  $p_z$  du neutrino reconstruit. Les événements simplement et doublement étiquetés sont combinés. Les distributions sont données pour le canal électron (à gauche) et le canal muon (à droite), dans le lot de contrôle  $W$ +jets (en haut), le lot de contrôle  $t\bar{t}$  (au centre) et l'ensemble des données sélectionnées (en bas). Les lots de contrôle sont définis dans l'annexe C.



# C

## Histogrammes de contrôle

Cette annexe présente plusieurs séries d'histogrammes de contrôle comparant les données observées aux bruits de fond et au signal attendus. Chaque série d'histogrammes se compose de quatre planches de 18 histogrammes. Chaque planche illustre 6 variables, soit un total de 24 variables. Quand un histogramme n'est pas défini, un espace blanc est inséré à la place. Les 24 variables illustrées incluent les 12 variables utilisées dans la construction des fonctions de vraisemblance définies au § 4.8.1. Ces variables sont :

- 1<sup>ère</sup> planche :
  - $p_T^{j1}$  : l'impulsion transverse du jet de plus grande impulsion transverse,
  - $p_T^b$  : l'impulsion transverse du jet *b* étiqueté *Tight* de plus grand  $p_T$ ,
  - $p_T^u$  : l'impulsion transverse du jet non étiqueté de plus grand  $p_T$ ,
  - $p_T^{j2}$ ,  $p_T^{j3}$  et  $p_T^{j4}$  : les impulsions transverses des second, troisième et quatrième jets, par ordre de  $p_T$  décroissant.
- 2<sup>ème</sup> planche :
  - $\eta^{j=1..4}$  : les pseudorapidités des quatre jets, classés par ordre de  $p_T$  décroissant,
  - $\Delta R_{min}^{jj}$  : la séparation angulaire minimale entre 2 jets,
  - $N_{jet}$  : la multiplicité de jets sélectionnés dans l'événement.
- 3<sup>ème</sup> planche :
  - $p_T^\ell$  : l'impulsion transverse du lepton chargé,
  - $\eta^\ell$  : la pseudorapidité du lepton chargé,
  - $\cancel{E}_T$  : l'énergie transverse manquante,
  - $METLep = p_T^\ell + \cancel{E}_T$ ,
  - $M_T^W$ , la masse transverse du boson *W*,
  - $H_T = p_T^{j1} + p_T^{j2} + p_T^\ell + \cancel{E}_T$

- 4<sup>ème</sup> planche :
  - $M_{Wb}$ , la masse reconstruite du quark top,
  - $\mathcal{C}$ , la centralité,
  - $\mathcal{S}$ , la sphéricité,
  - $M_{jets}$ , la masse invariante du système composé de tous les jets sélectionnés,
  - $Q \times \eta$ ,
  - $\cos(j_2, \ell)_{top}$ , le cosinus de l'angle entre le second jet de plus haut  $p_T$  et le lepton dans le référentiel du quark top reconstruit avec le jet de plus haut  $p_T$  et le boson  $W$ .

Afin de valider la modélisation des fonds dominants  $W$ +jets et  $t\bar{t}$ , deux sous-lots sont définis dans les données sélectionnées en fonction de  $N_{jet}$  (multiplicité de jet) et  $H_T$ . Le **lot de contrôle  $W$ +jets** est défini par  $N_{jet} = 2$  et  $H_T < 200$ . Ce lot est dominé par les fonds  $W$ +jets et multijet. Le **lot de contrôle  $t\bar{t}$**  est défini par  $N_{jet} = 4$  et  $H_T > 270$  et contient principalement des événements  $t\bar{t}$  lepton+jets.

La première série d'histogrammes de contrôle (figures C.1 à C.4) donne les distributions des variables avant application d'un schéma d'étiquetage des quarks  $b$ , pour les canaux électron (colonne de gauche), muon (colonne centrale) et combinés (colonne de droite). Les variables qui dépendent de l'identification des jets  $b$  sont remplacées par une ligne vide.

La seconde série d'histogrammes de contrôle (figures C.5 à C.8) donne les distributions dans le canal électron pour les événements simplement et doublement étiquetés combinés dans le lot de contrôle  $W$ +jets (colonne de gauche), le lot de contrôle  $t\bar{t}$  (colonne centrale) et l'ensemble des données sélectionnées (colonne de droite).

La troisième série d'histogrammes de contrôle (figures C.9 à C.12) donne les distributions dans le canal muon pour les événements simplement et doublement étiquetés combinés dans le lot de contrôle  $W$ +jets (colonne de gauche), le lot de contrôle  $t\bar{t}$  (colonne centrale) et l'ensemble des données sélectionnées (colonne de droite).

La quatrième série d'histogrammes de contrôle (figures C.13 à C.16) donne les distributions dans les canaux électron et muon combinés pour les schémas d'étiquetage standard : STag (colonne de gauche), DTag (colonne centrale) et STag+Dtag (colonne de droite).

La cinquième série d'histogrammes de contrôle (figures C.17 à C.20) donne les distributions dans les canaux électron et muon combinés pour les schémas d'étiquetage asymétriques : STagVeto (colonne de gauche), DTagAsym (colonne centrale) et STag-Veto+DTagAsym (colonne de droite).

Tous ces histogrammes utilisent la même légende :

- Bleu : signal *single top* voie  $s$  (multiplié par 10),
- Cyan : signal *single top* voie  $t$  (multiplié par 10),
- Rouge : fonds  $t\bar{t}$  lepton+jet et dilepton,
- Vert : fonds  $W$ +jets,  $WW$  et  $WZ$ ,
- Marron : fond multijet.

- Points noirs : données observées avec les barres d'erreurs statistiques. Les erreurs systématiques sur les fonds ne sont pas indiquées. Les luminosités intégrées correspondantes sont  $\mathcal{L} = 366 \text{ pb}^{-1}$  dans le canal électron et  $\mathcal{L} = 363 \text{ pb}^{-1}$  dans le canal muon.

Des histogrammes de contrôle supplémentaires sont fournis en annexe de la référence [110].

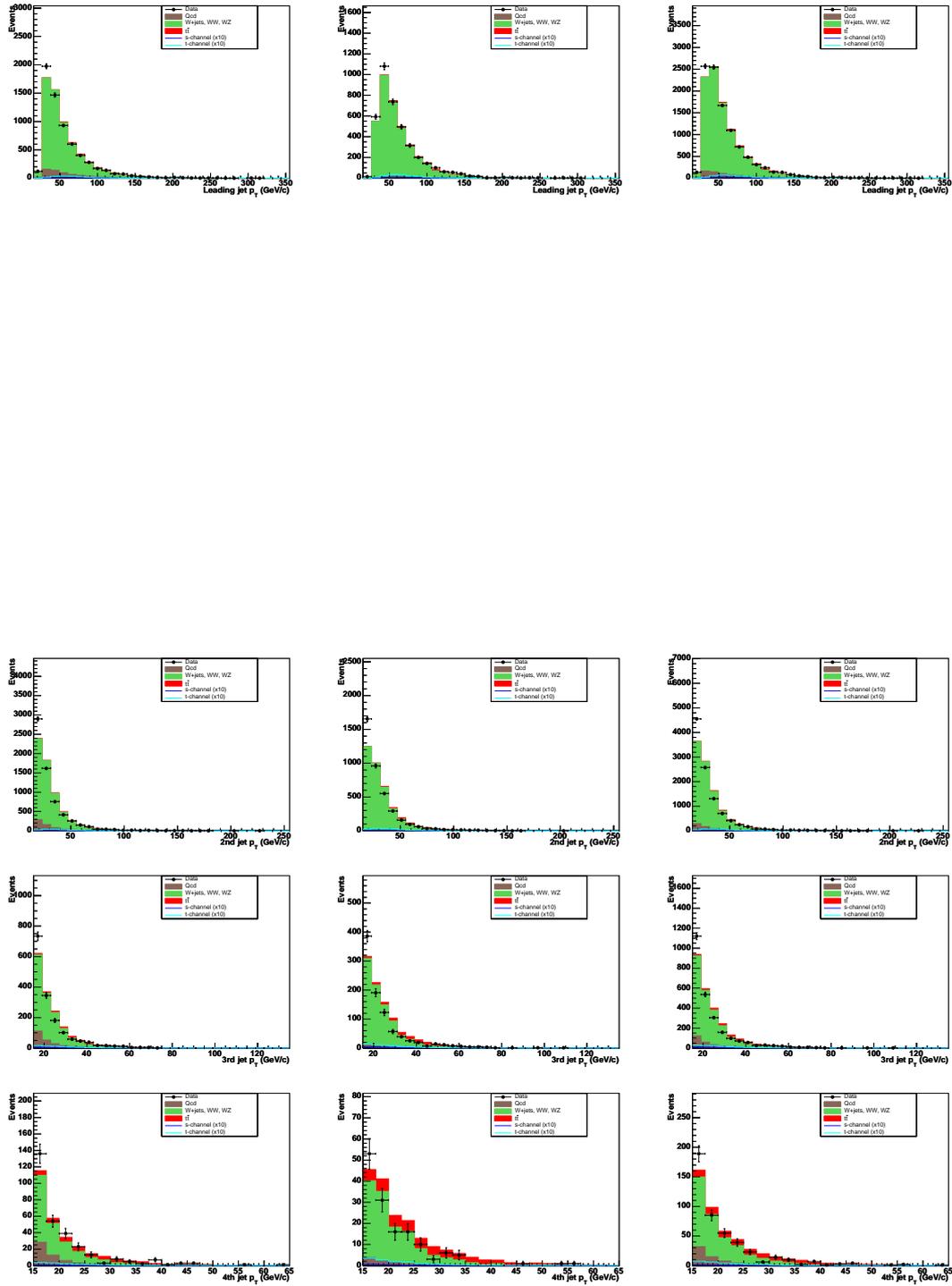


FIG. C.1: Histogrammes de contrôle, avant étiquetage.  
 De haut en bas :  $p_T^{j1}$ ,  $p_T^b$ ,  $p_T^u$ ,  $p_T^{j2}$ ,  $p_T^{j3}$  et  $p_T^{j4}$ .  
 De gauche à droite : canal électron, canal muon, canaux combinés.

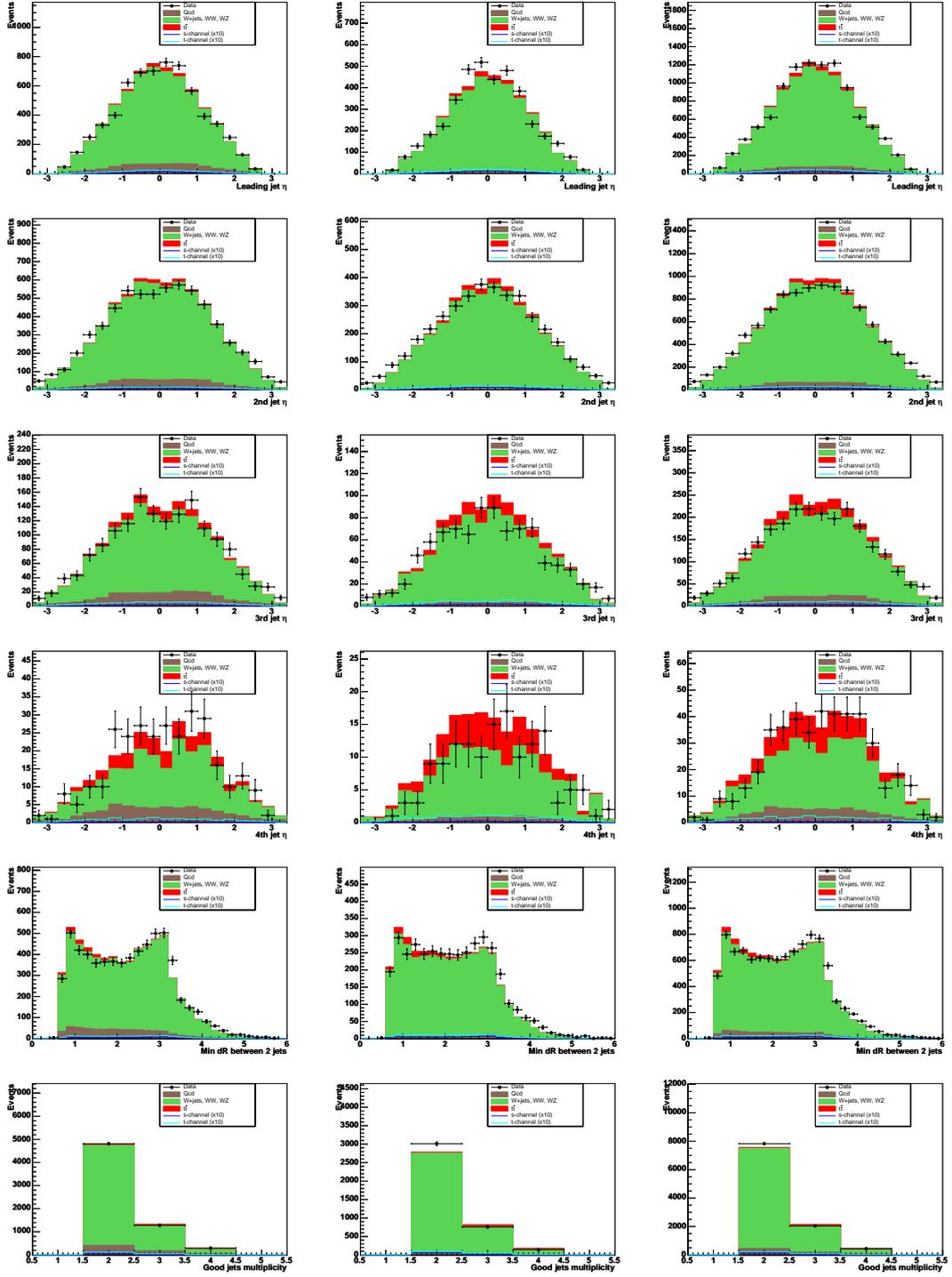


FIG. C.2: Histogrammes de contrôle, avant étiquetage.  
 De haut en bas :  $\eta^{j1}, \eta^{j2}, \eta^{j3}, \eta^{j4}, \Delta R_{min}^{jj}, N_{jet}$ .  
 De gauche à droite : canal électron, canal muon, canaux combinés.

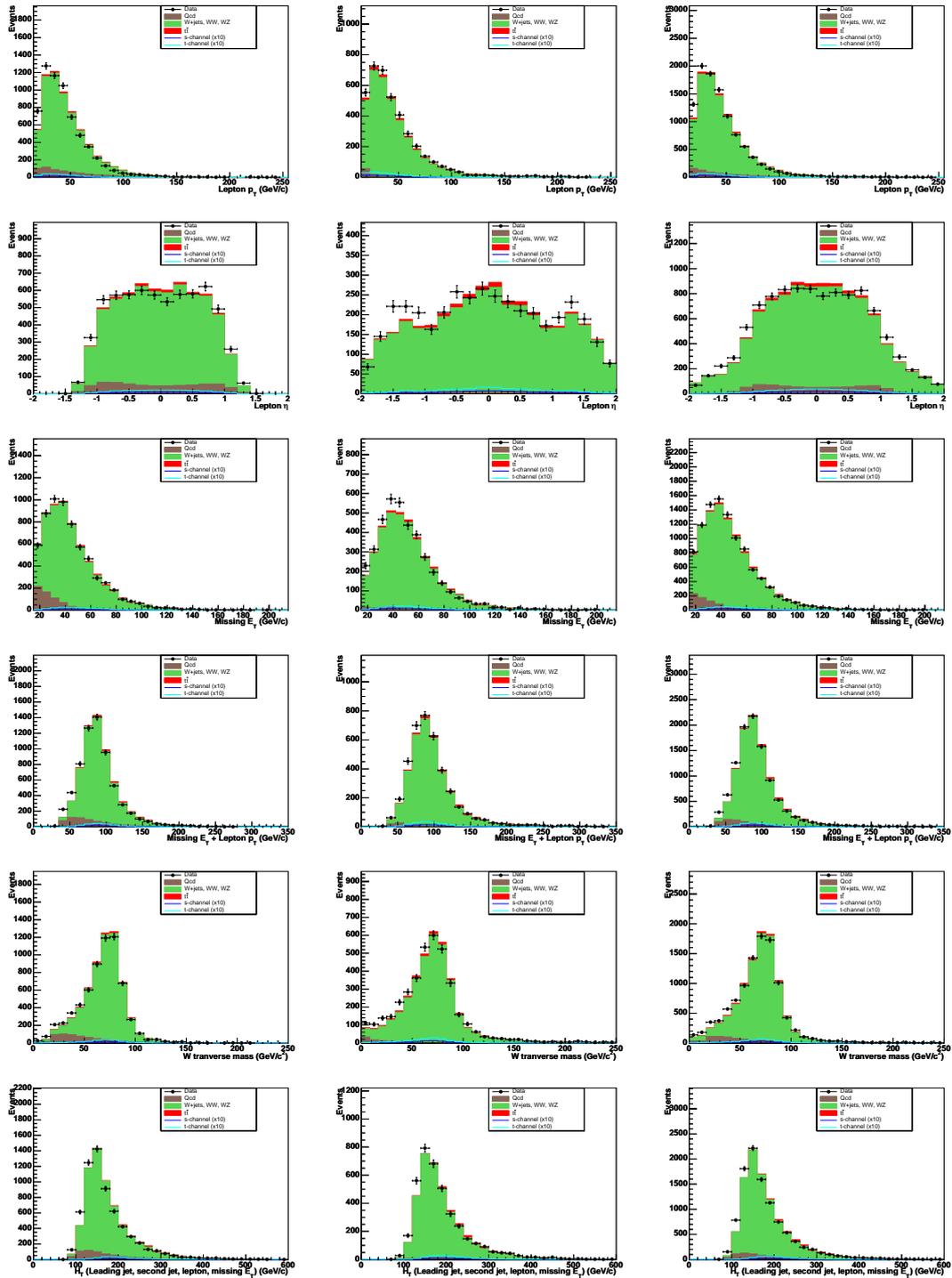


FIG. C.3: Histogrammes de contrôle, avant étiquetage.  
 De haut en bas :  $p_T^\ell$ ,  $\eta^\ell$ ,  $E_T^{\text{Missing}}$ ,  $M_T^W$ ,  $H_T$ .  
 De gauche à droite : canal électron, canal muon, canaux combinés.

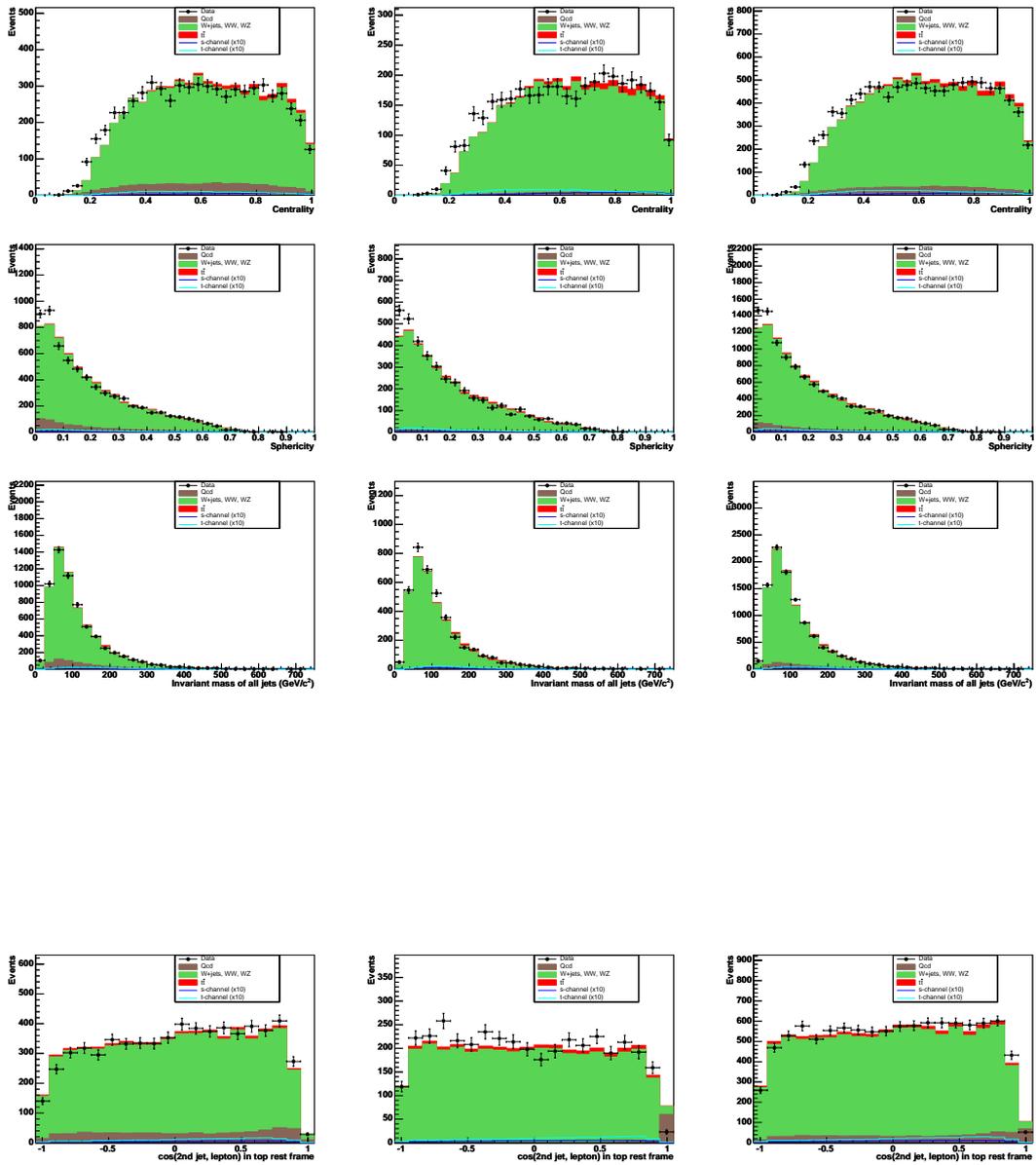


FIG. C.4: Histogrammes de contrôle, avant étiquetage.  
 De haut en bas :  $M_{Wb}$ ,  $\mathcal{C}$ ,  $\mathcal{S}$ ,  $M_{jets}$ ,  $Q \times \eta$ ,  $\cos(j_2, \ell)_{top}$ .  
 De gauche à droite : canal électron, canal muon, canaux combinés.

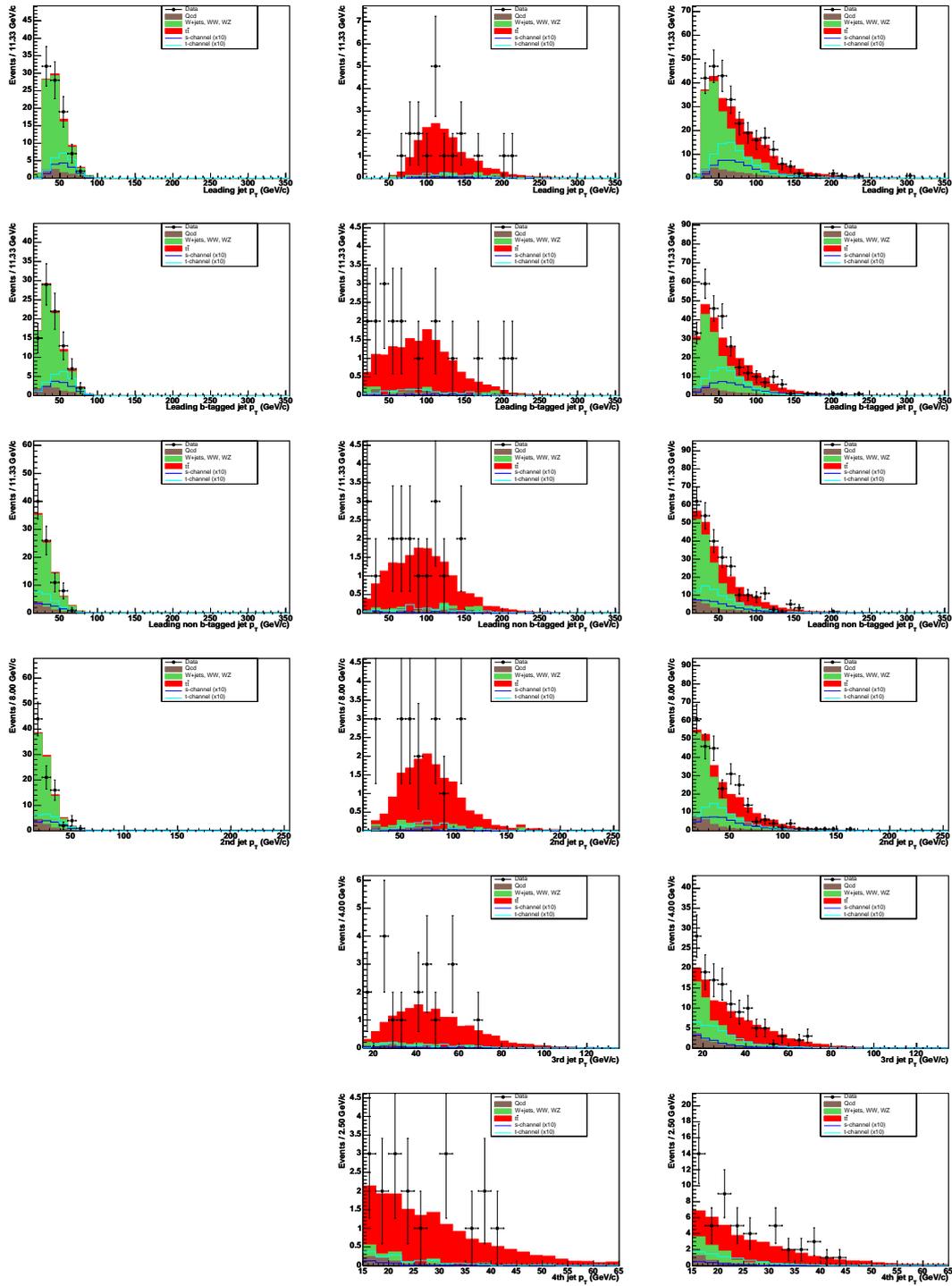


FIG. C.5: Histogrammes de contrôle, canal électron, événements simplement et doublement étiquetés combinés.

De haut en bas :  $p_T^{j1}$ ,  $p_T^b$ ,  $p_T^u$ ,  $p_T^{j2}$ ,  $p_T^{j3}$  et  $p_T^{j4}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

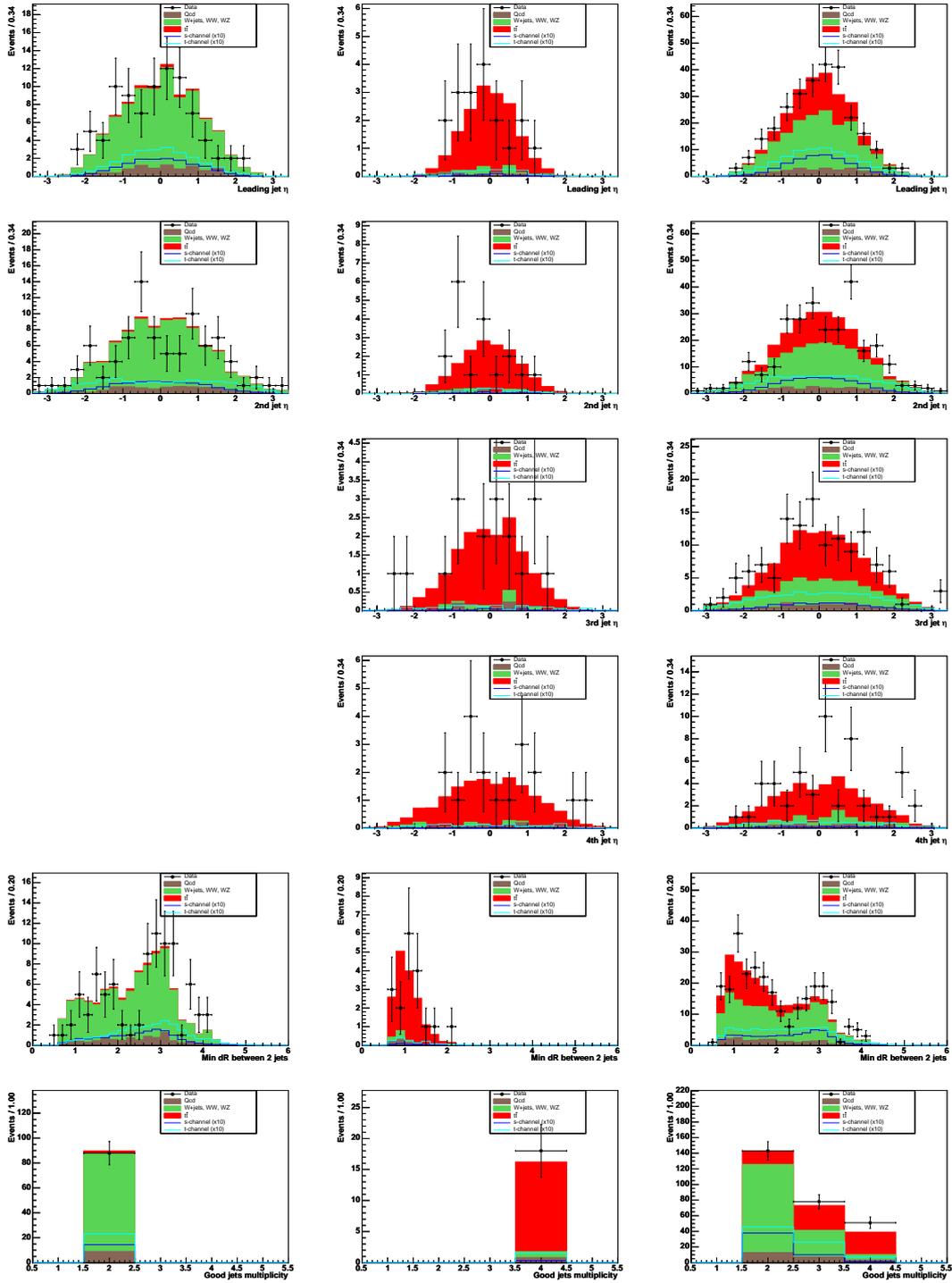


FIG. C.6: Histogrammes de contrôle, canal électron, événements simplement et doublement étiquetés combinés.

De haut en bas :  $\eta^{j1}, \eta^{j2}, \eta^{j3}, \eta^{j4}, \Delta R_{min}^{jj}, N_{jet}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

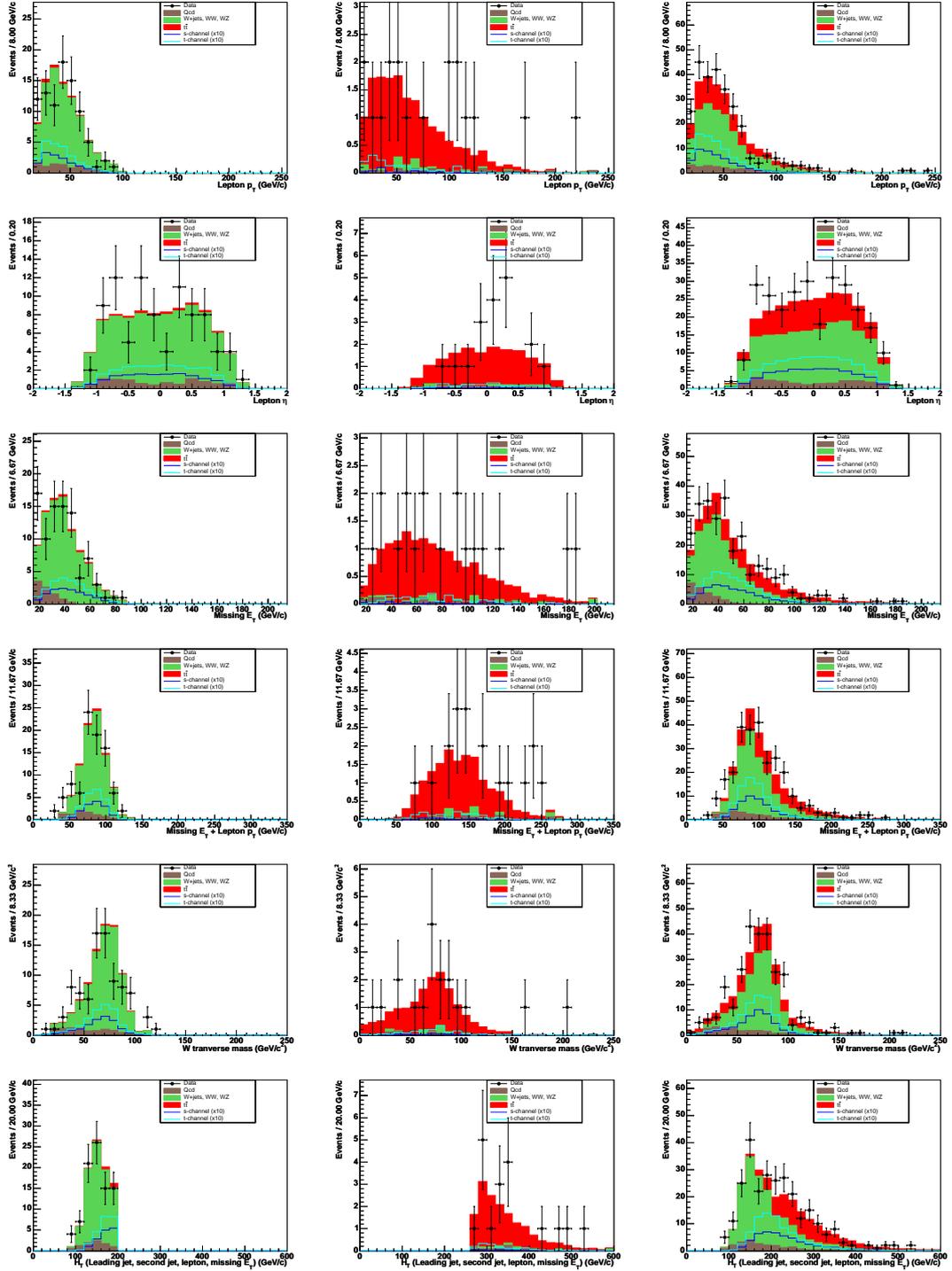


FIG. C.7: Histogrammes de contrôle, canal électron, événements simplement et doublement étiquetés combinés.

De haut en bas :  $p_T^\ell$ ,  $\eta^\ell$ ,  $\cancel{E}_T$ ,  $M_T^W$ ,  $H_T$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

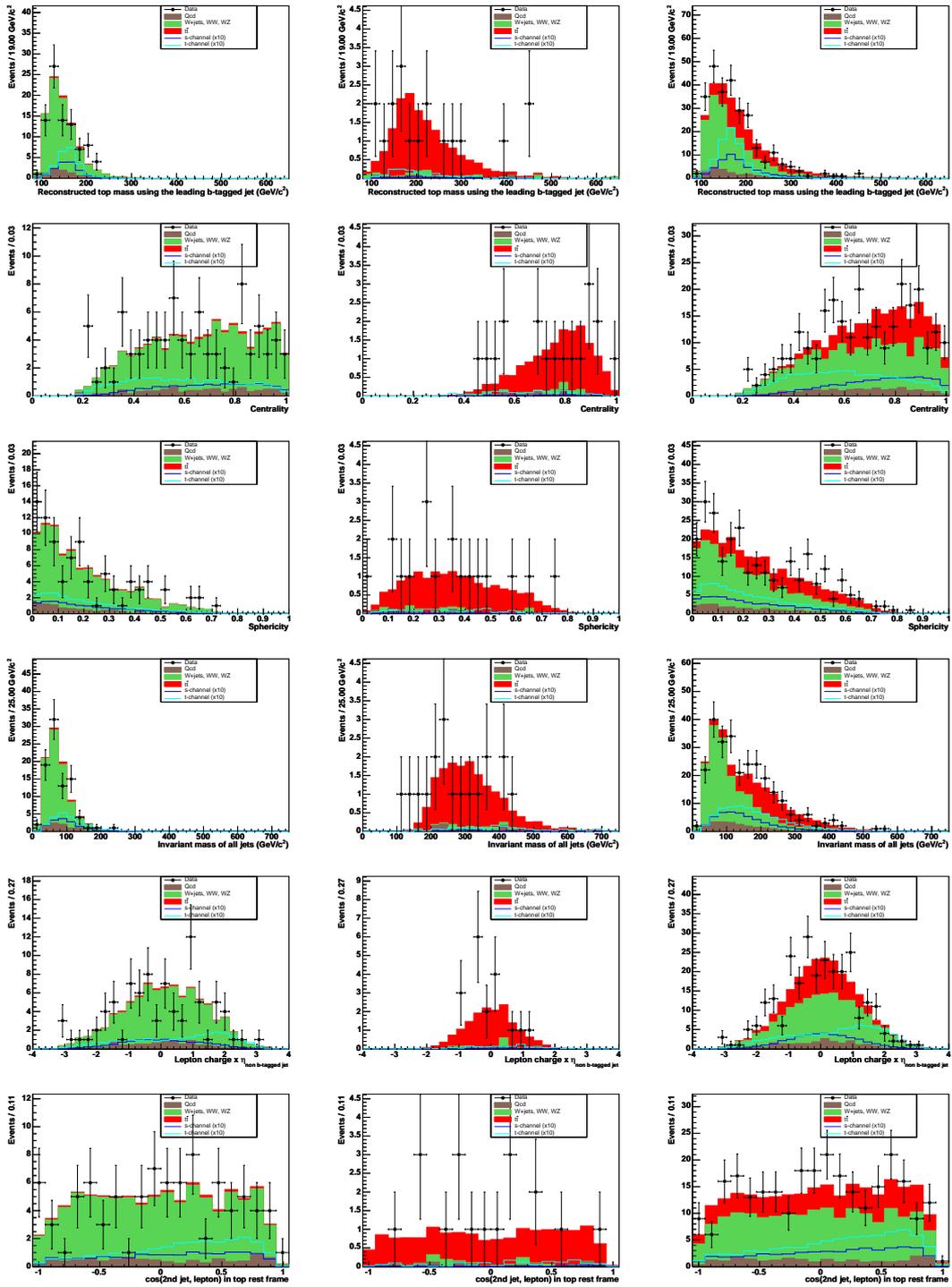


FIG. C.8: Histogrammes de contrôle, canal électron, événements simplement et doublement étiquetés combinés.

De haut en bas :  $M_{Wb}$ ,  $C$ ,  $S$ ,  $M_{jets}$ ,  $Q \times \eta$ ,  $\cos(j_2, \ell)_{top}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

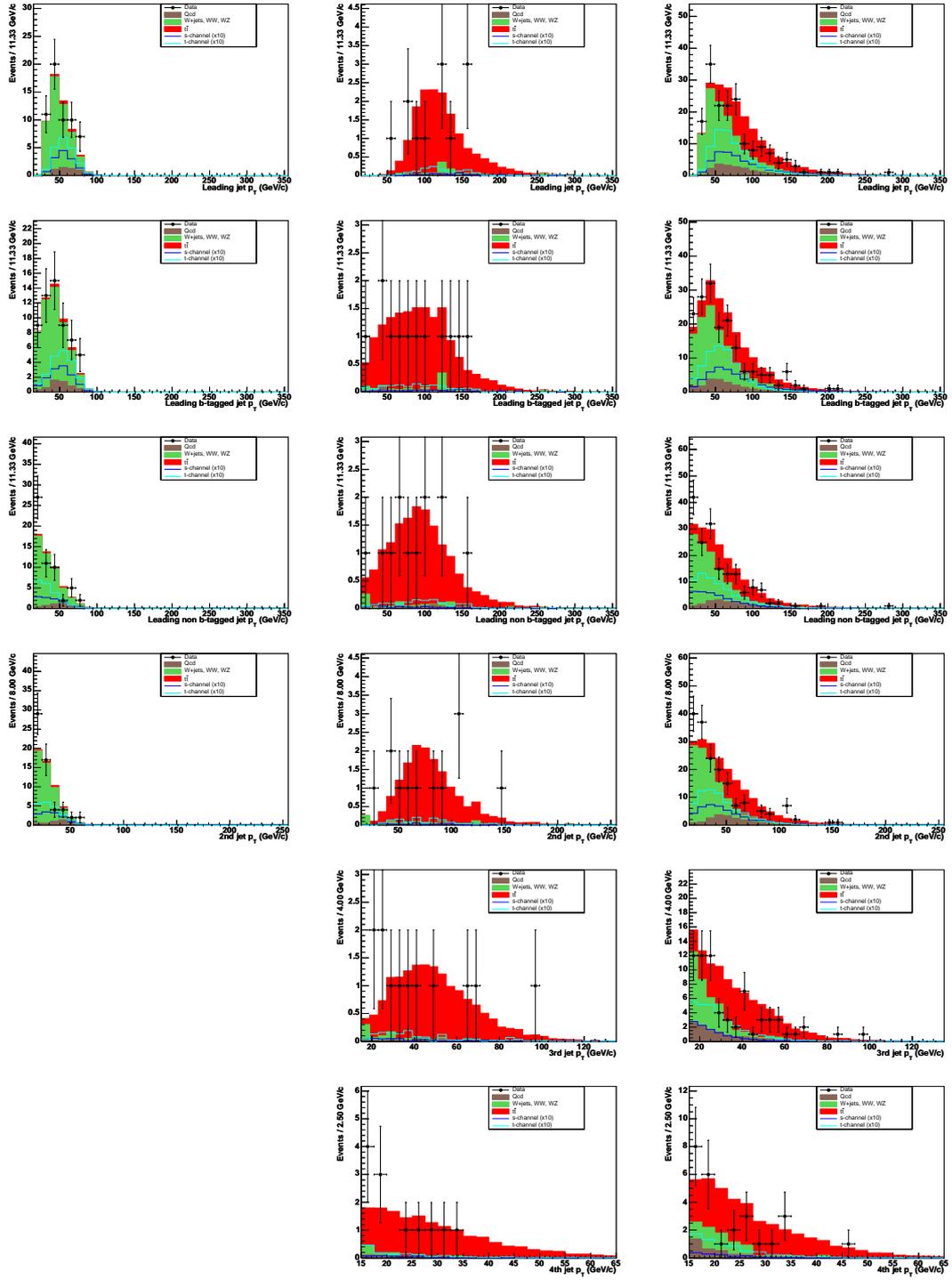


FIG. C.9: Histogrammes de contrôle, canal muon, événements simplement et doublement étiquetés combinés.

De haut en bas :  $p_T^{j1}$ ,  $p_T^b$ ,  $p_T^u$ ,  $p_T^{j2}$ ,  $p_T^{j3}$  et  $p_T^{j4}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

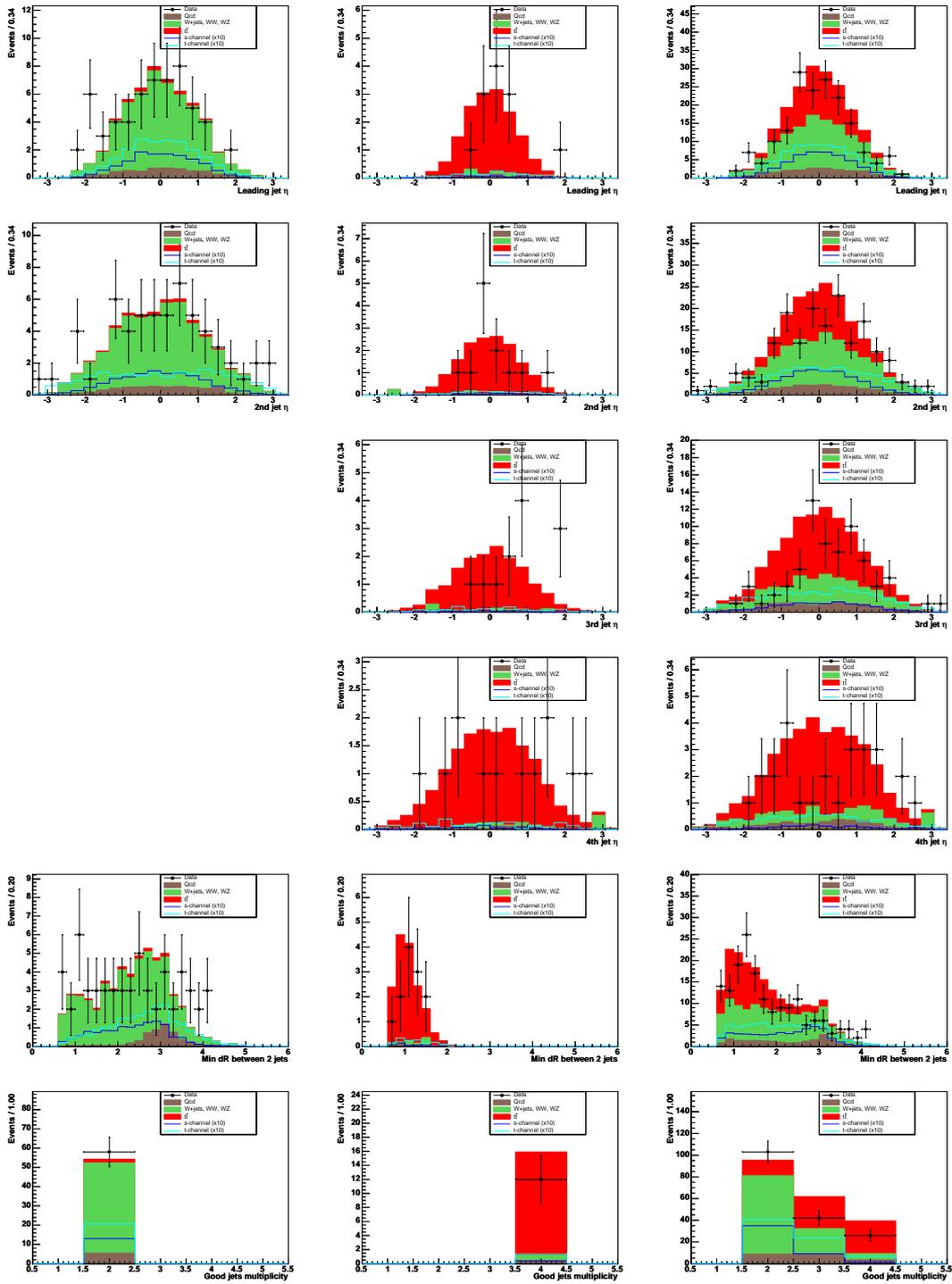


FIG. C.10: Histogrammes de contrôle, canal muon, événements simplement et doublement étiquetés combinés.

De haut en bas :  $\eta^{j1}, \eta^{j2}, \eta^{j3}, \eta^{j4}, \Delta R_{min}^{jj}, N_{jet}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

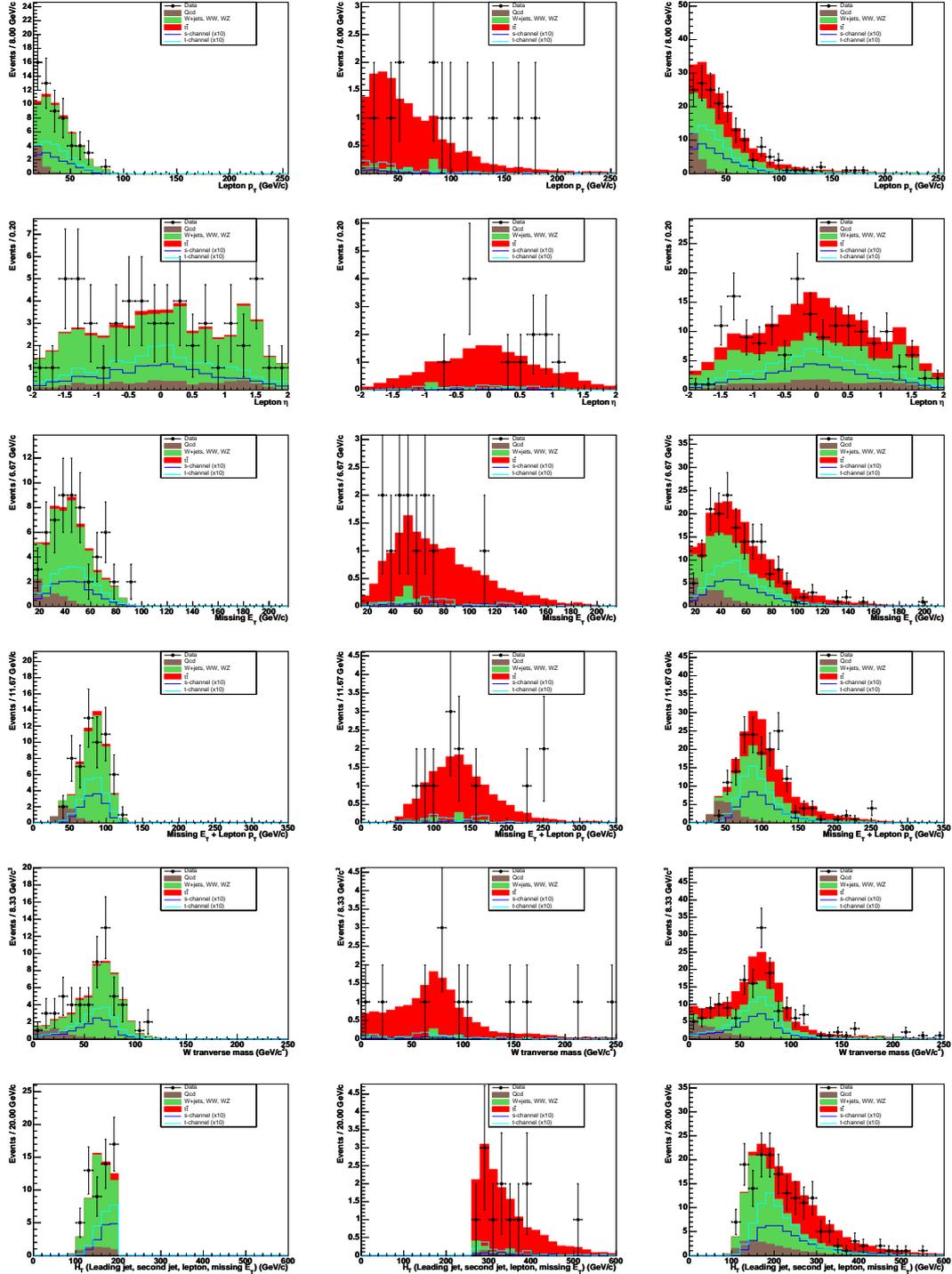


FIG. C.11: Histogrammes de contrôle, canal muon, événements simplement et doublement étiquetés combinés.

De haut en bas :  $p_T^\ell$ ,  $\eta^\ell$ ,  $\cancel{E}_T$ ,  $MET_{Lep}$ ,  $M_T^W$ ,  $H_T$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

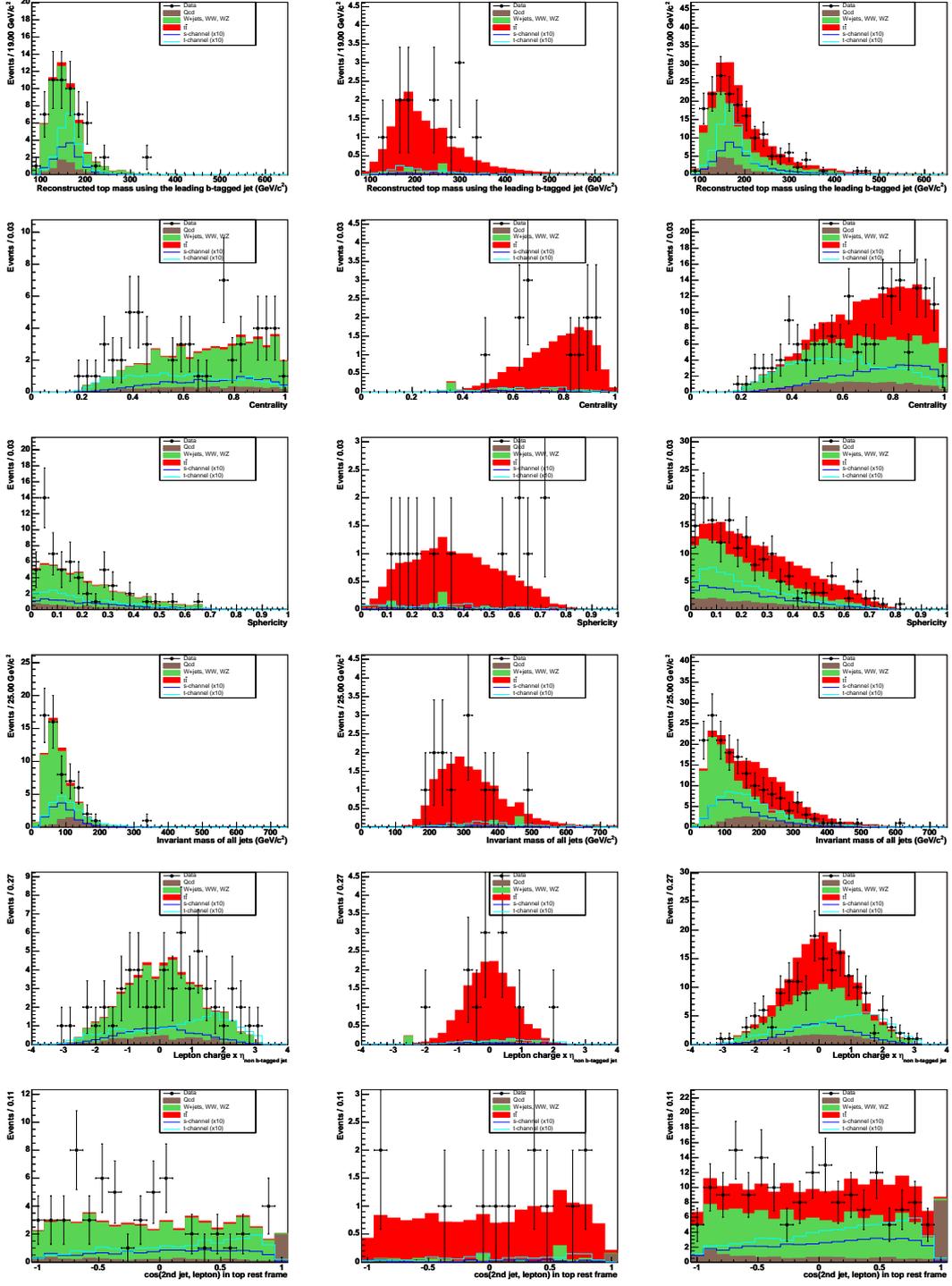


FIG. C.12: Histogrammes de contrôle, canal muon, événements simplement et doublement étiquetés combinés.

De haut en bas :  $M_{Wb}$ ,  $C$ ,  $S$ ,  $M_{jets}$ ,  $Q \times \eta$ ,  $\cos(j_2, \ell)_{top}$ .

De gauche à droite : lot  $W$ +jets, lot  $t\bar{t}$ , l'ensemble des données sélectionnées.

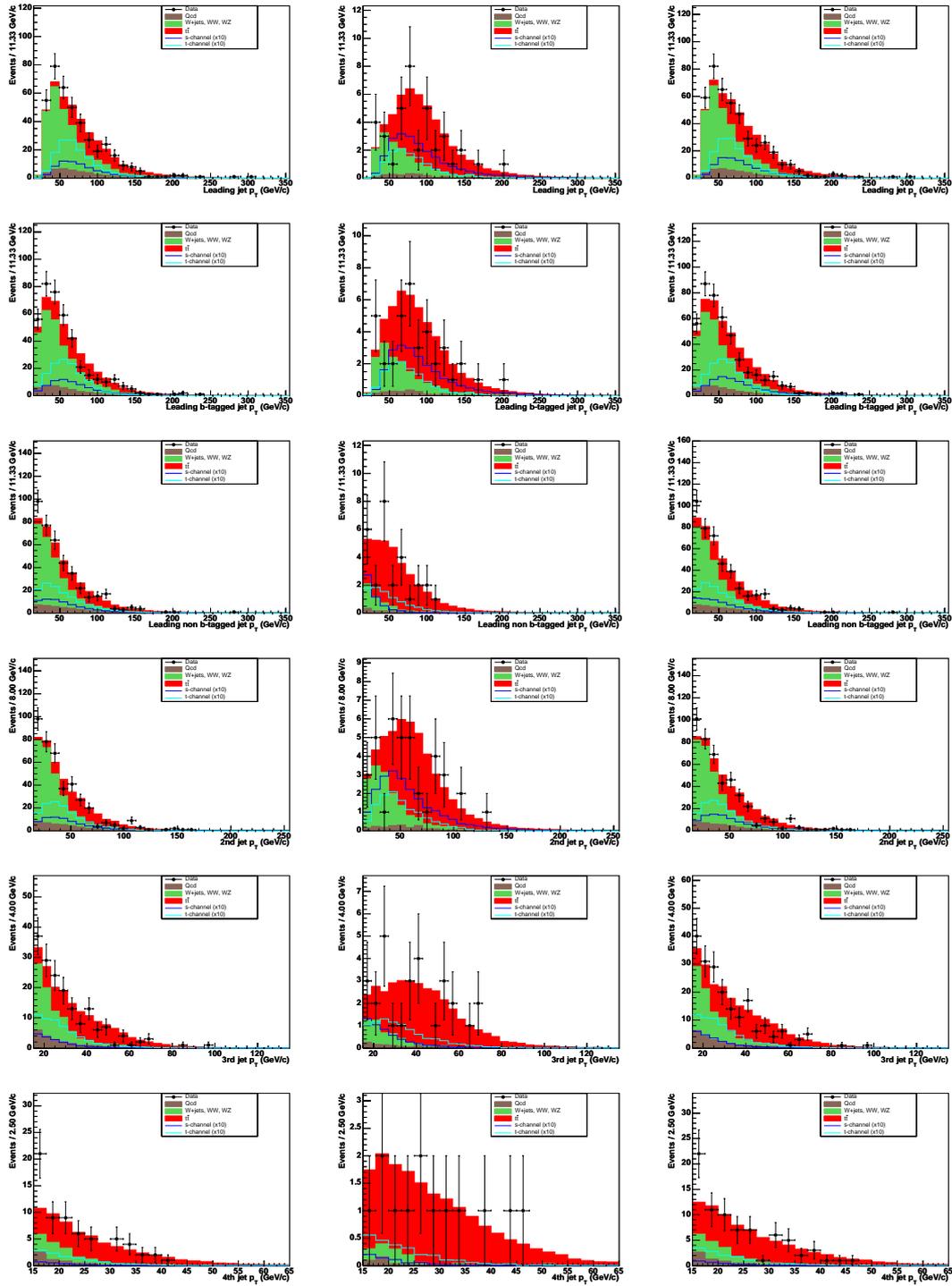


FIG. C.13: Histogrammes de contrôle, canaux électron et muon combinés.  
 De haut en bas :  $p_T^{j1}$ ,  $p_T^b$ ,  $p_T^u$ ,  $p_T^{j2}$ ,  $p_T^{j3}$  et  $p_T^{j4}$ .  
 De gauche à droite : schéma STag, schéma DTag, schémas combinés STag+DTag.

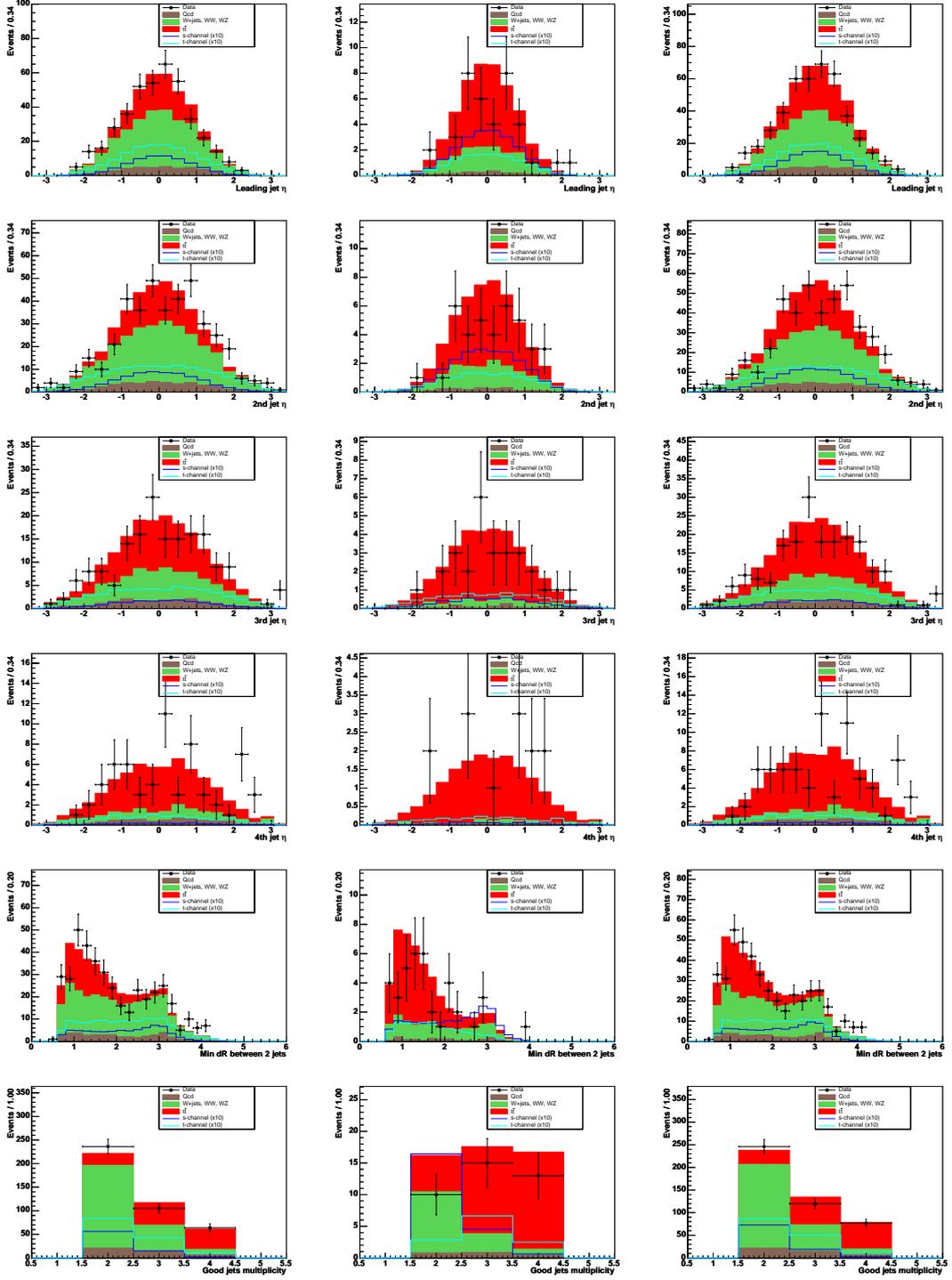


FIG. C.14: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $\eta^{j1}, \eta^{j2}, \eta^{j3}, \eta^{j4}, \Delta R_{min}^{jj}, N_{jet}$ .

De gauche à droite : schéma STag, schéma DTag, schémas combinés STag+DTag.

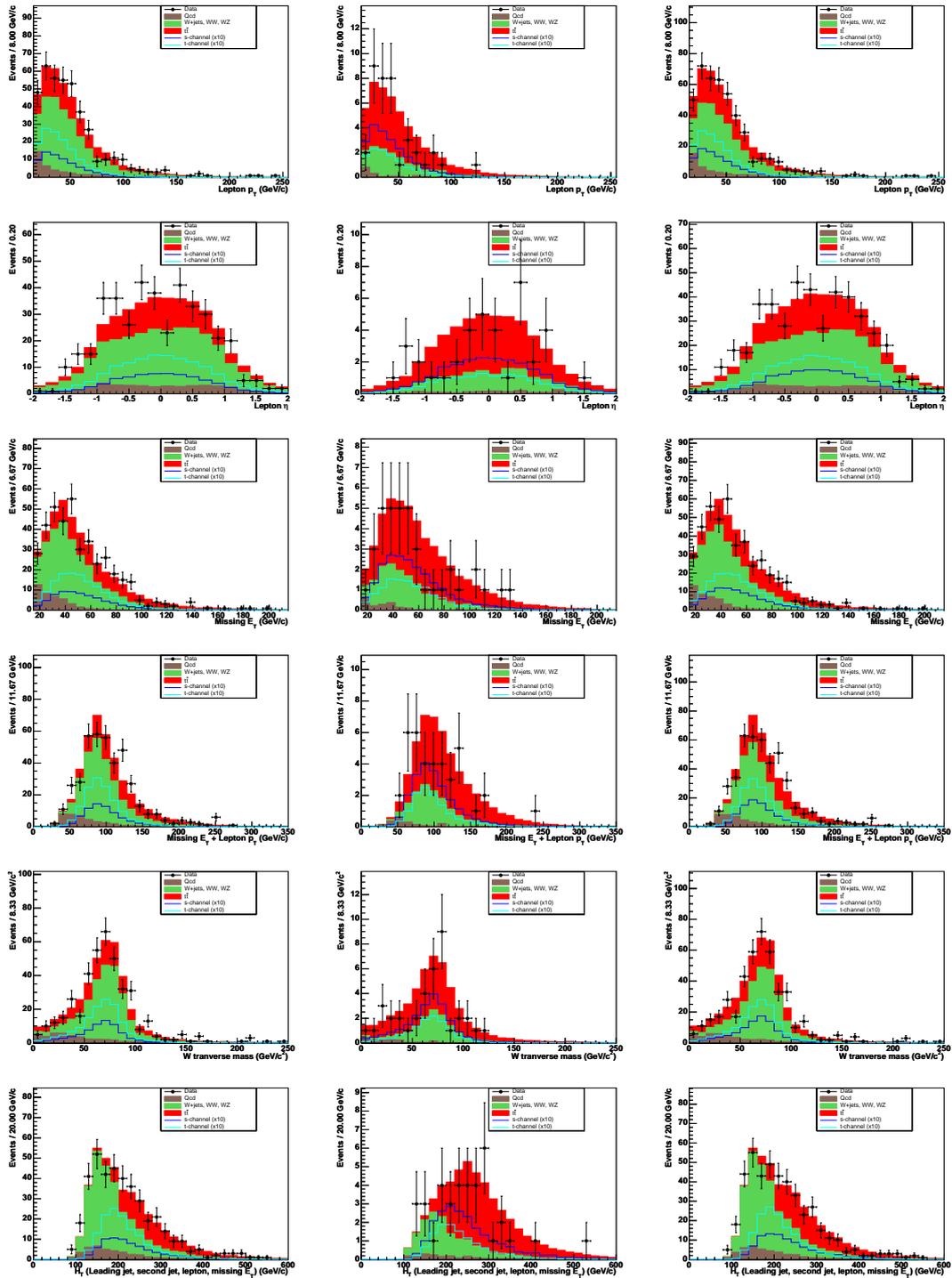


FIG. C.15: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $p_T^l$ ,  $\eta^l$ ,  $\cancel{E}_T$ ,  $M_{T}^{Lep}$ ,  $M_T^W$ ,  $H_T$ .

De gauche à droite : schéma STag, schéma DTag, schémas combinés STag+DTag.

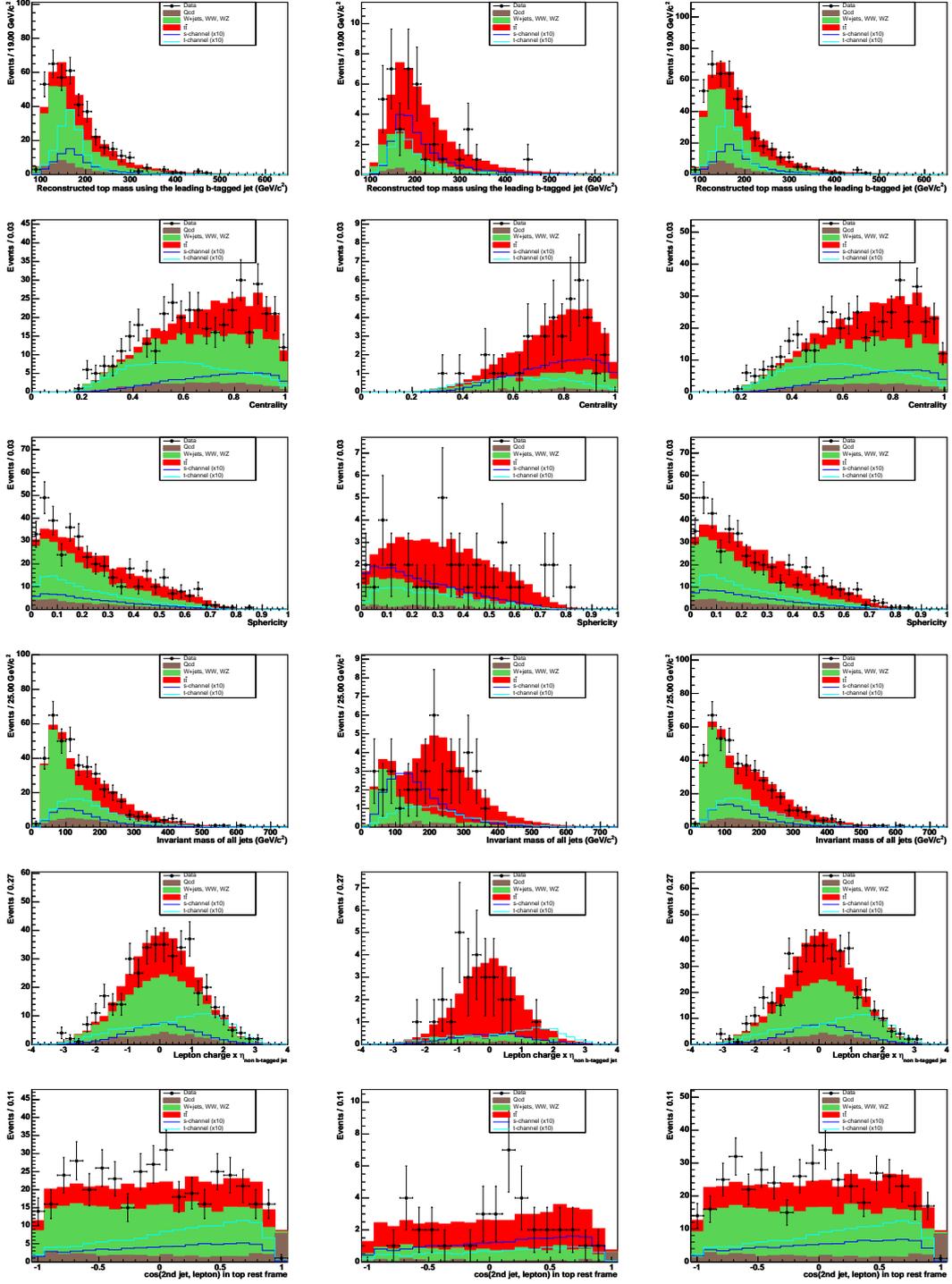


FIG. C.16: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $M_{Wb}$ ,  $C$ ,  $S$ ,  $M_{jets}$ ,  $Q \times \eta$ ,  $\cos(j_2, \ell)_{top}$ .

De gauche à droite : schéma STag, schéma DTag, schémas combinés STag+DTag.

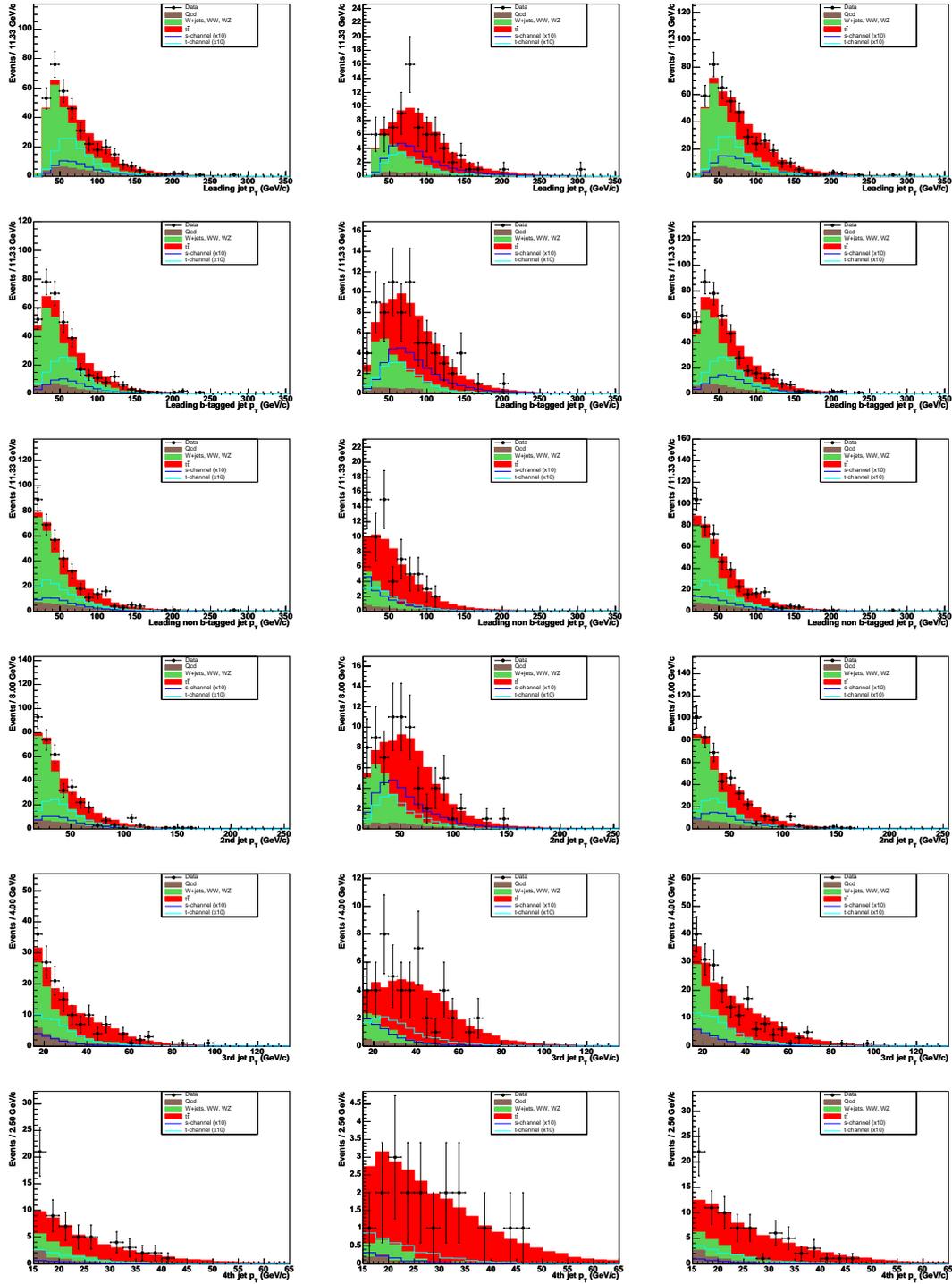


FIG. C.17: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $p_T^{J1}$ ,  $p_T^b$ ,  $p_T^u$ ,  $p_T^{J2}$ ,  $p_T^{J3}$  et  $p_T^{J4}$ .

De gauche à droite : schéma STagVeto, schéma DTagAsym, schémas combinés STagVeto+DTagAsym.

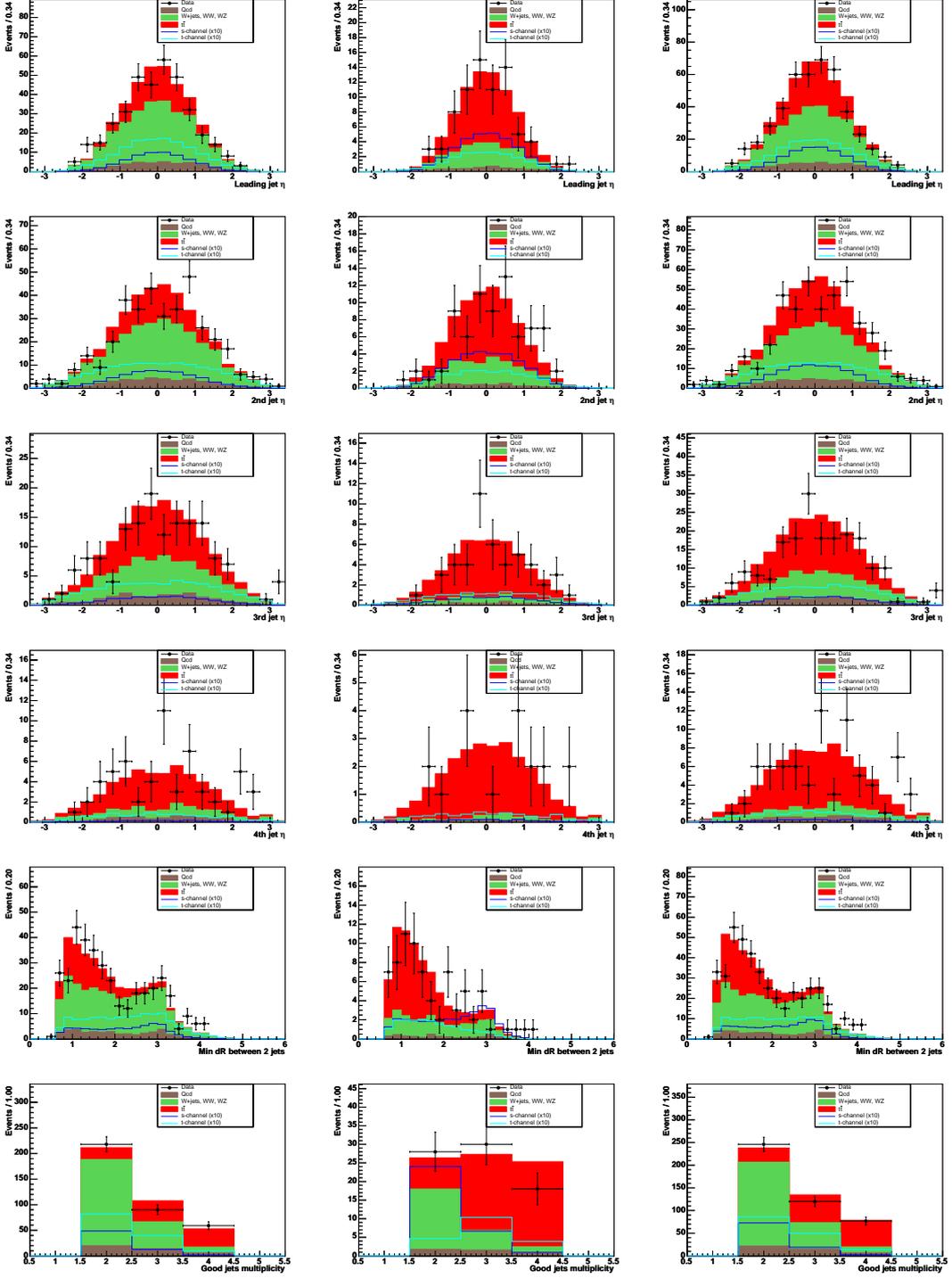


FIG. C.18: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $\eta^{j1}, \eta^{j2}, \eta^{j3}, \eta^{j4}, \Delta R_{min}^{jj}, N_{jet}$ .

De gauche à droite : schéma STagVeto, schéma DTagAsym, schémas combinés STagVeto+DTagAsym.

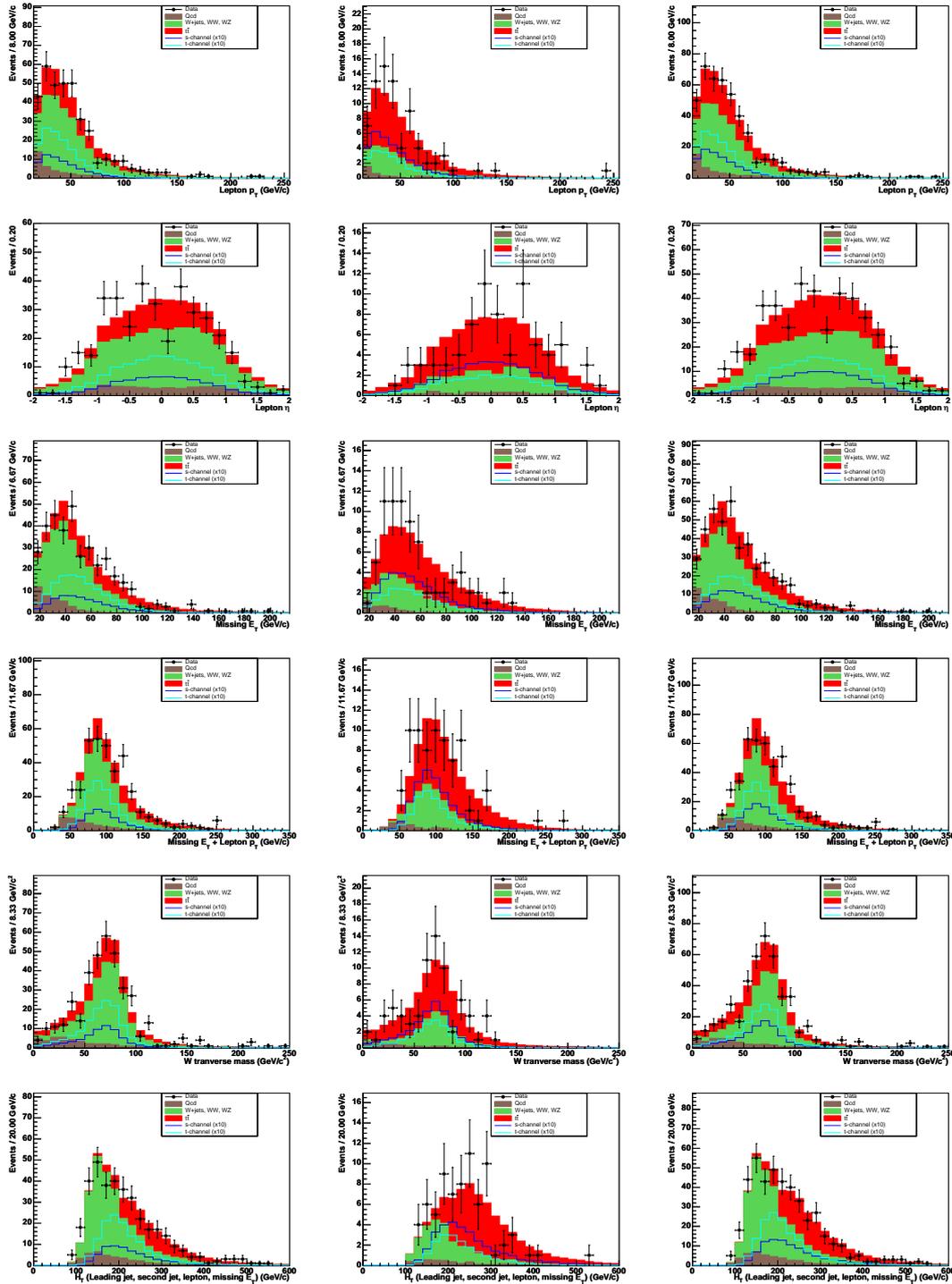


FIG. C.19: Histogrammes de contrôle, canaux électron et muon combinés.

De haut en bas :  $p_T^l$ ,  $\eta^l$ ,  $E_T^{\text{miss}}$ ,  $MET^{\text{Lep}}$ ,  $M_T^W$ ,  $H_T$ .

De gauche à droite : schéma STagVeto, schéma DTagAsym, schémas combinés STagVeto+DTagAsym.

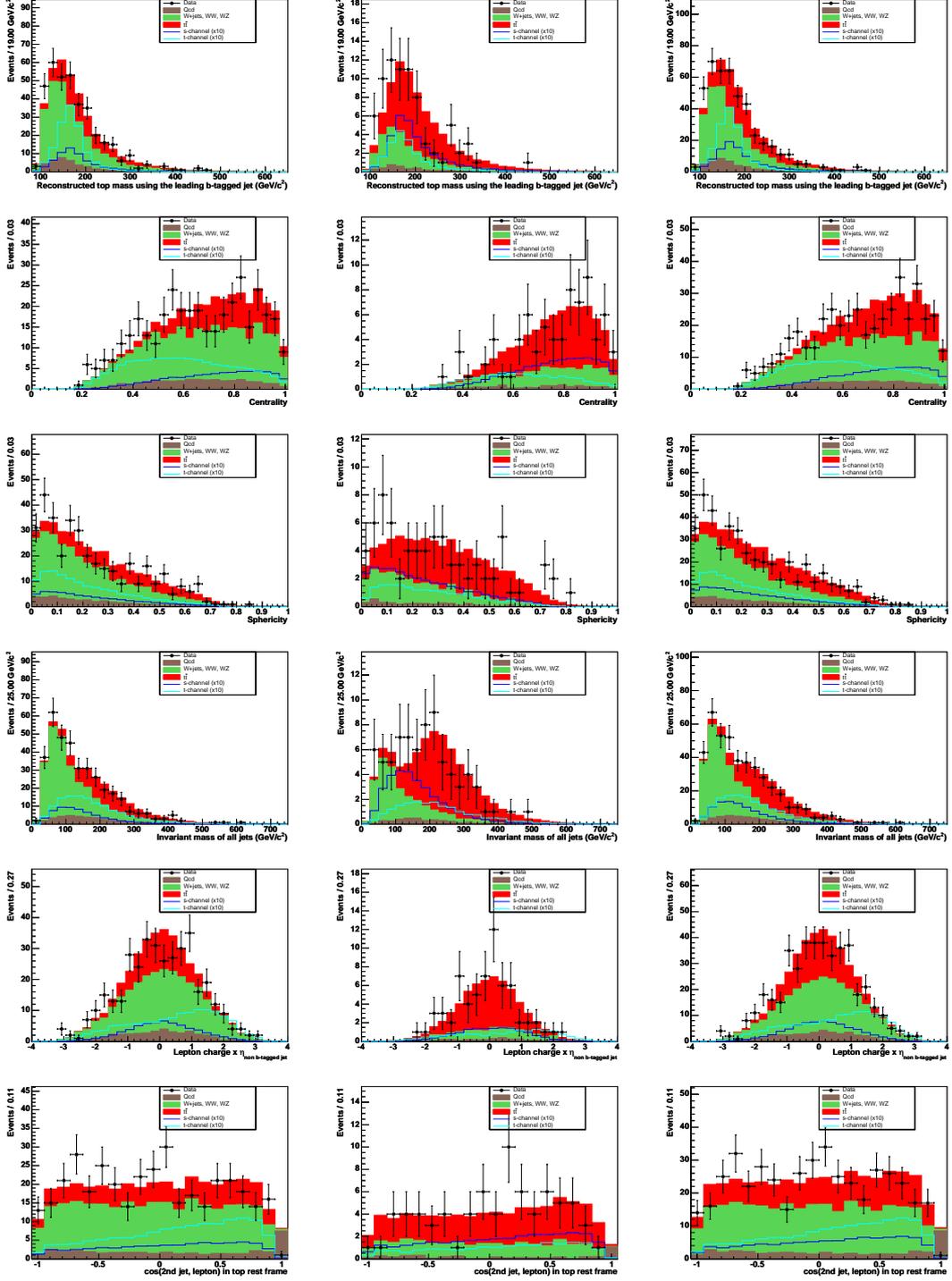


FIG. C.20: Histogrammes de contrôle, canaux électron et muon combinés.  
 De haut en bas :  $M_{Wb}$ ,  $C$ ,  $S$ ,  $M_{jets}$ ,  $Q \times \eta$ ,  $\cos(j_2, \ell)_{top}$ .  
 De gauche à droite : schéma STagVeto, schéma DTagAsym, schémas combinés STagVeto+DTagAsym.



# D

## Densités de probabilité

Les densités de probabilité pour les 12 variables discriminantes utilisées dans la construction des fonctions de vraisemblance sont données sur :

- les figures D.1 à D.3, canal électron, STag.
- les figures D.4 à D.6, canal muon, STag.
- les figures D.7 à D.9, canal électron, STagVeto.
- les figures D.10 à D.12, canal muon, STagVeto.
- les figures D.13 à D.15, canal électron, DTag.
- les figures D.16 à D.18, canal muon, DTag.
- les figures D.19 à D.21, canal électron, DTagAsym.
- les figures D.22 à D.24, canal muon, DTagAsym.

La légende de l'ensemble des graphes est la suivante :

- Signal *single top* voie  $s$  : bleu
- Signal *single top* voie  $t$  : cyan
- Fond  $t\bar{t}$  : rouge
- Fond  $W$ +jets : vert. Pour les schémas STag et STagVeto, ce fond inclut les processus  $Wjj$ ,  $Wbb$  et multijet. Pour les schémas DTag et DTagAsym seul le processus  $Wbb$  est inclus.

## Canal électron, STag

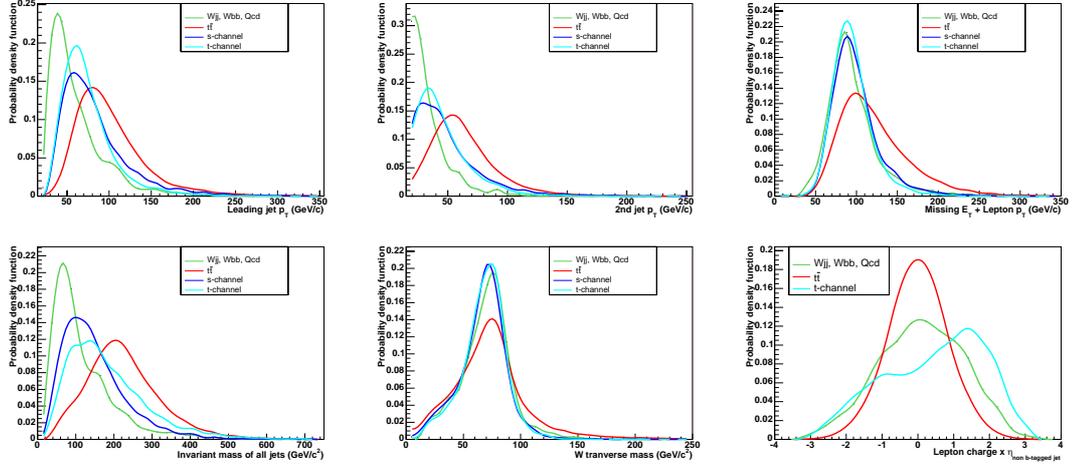


FIG. D.1: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

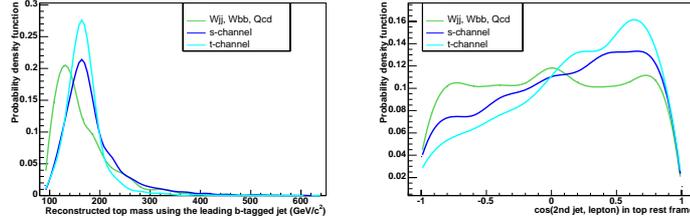


FIG. D.2: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(j_2, \ell)_{top}$ .

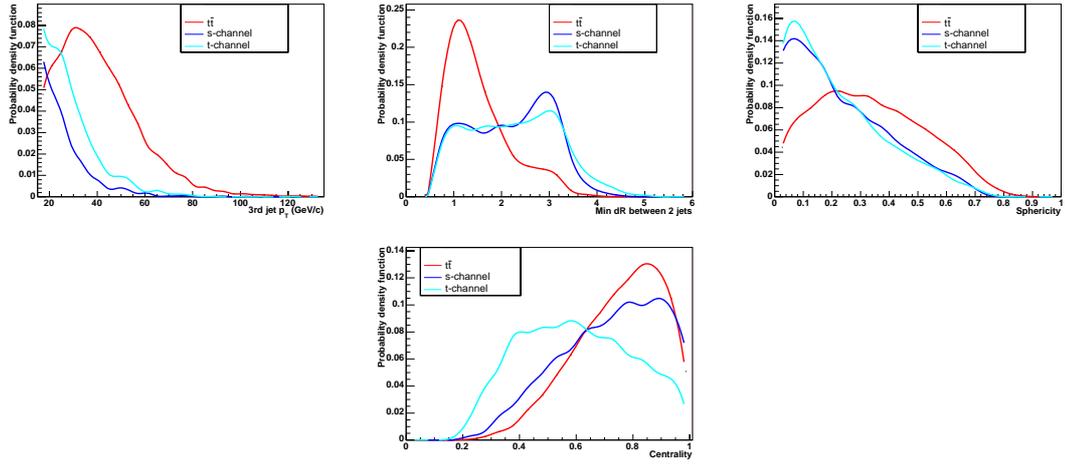


FIG. D.3: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphéricité$  et  $Centralité$ .

## Canal muon, STag

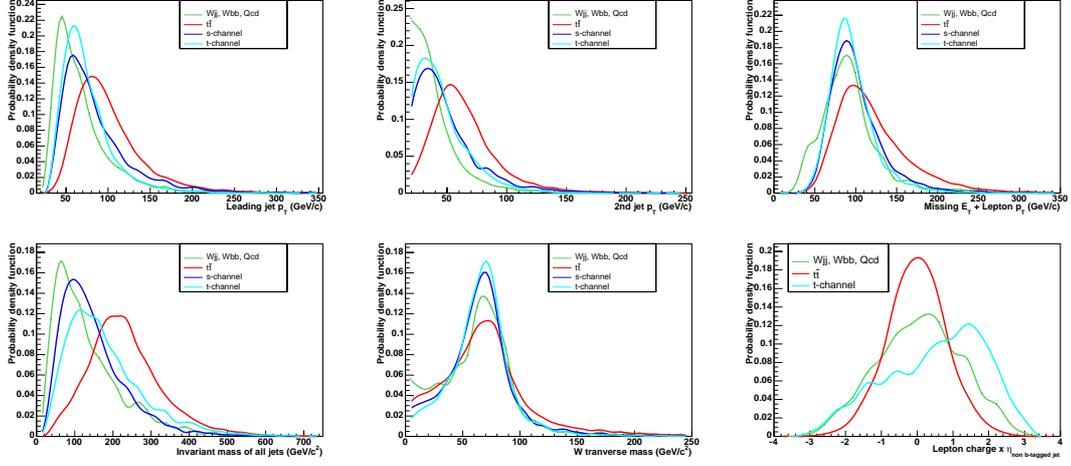


FIG. D.4: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $MET_{Lep}$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

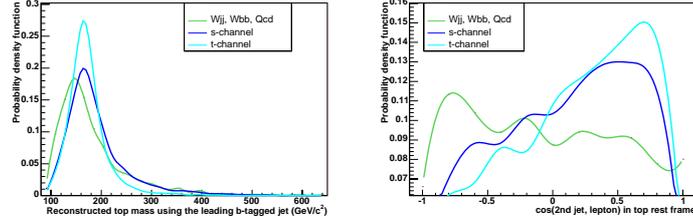


FIG. D.5: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(j_2, \ell)_{top}$ .

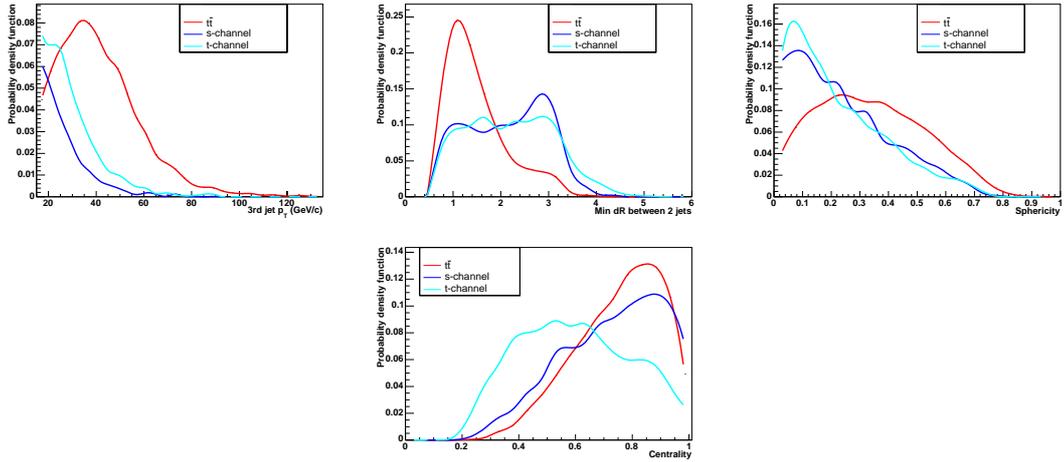


FIG. D.6: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphéricité$  et  $Centralité$ .

## Canal électron, STagVeto

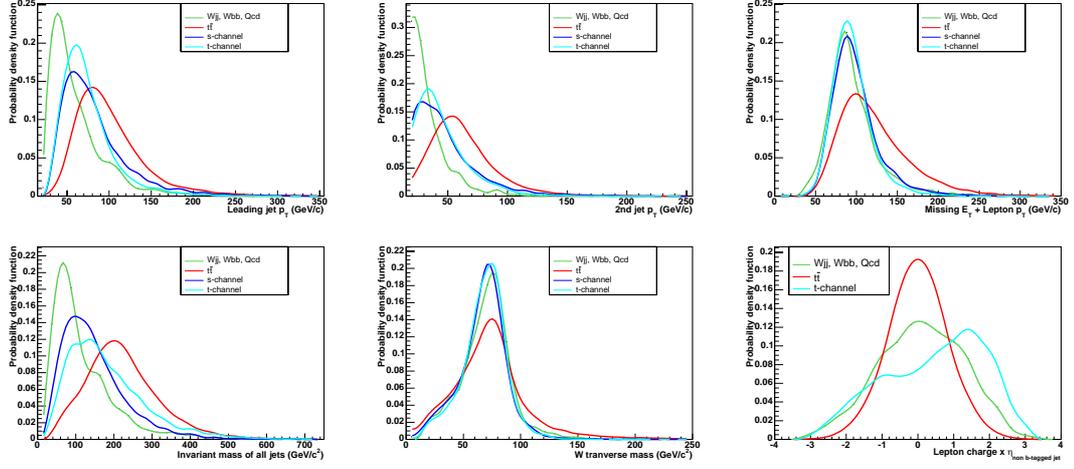


FIG. D.7: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $MET Lep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

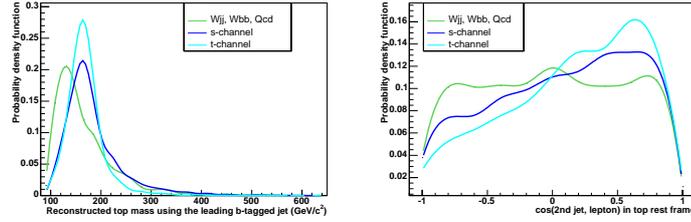


FIG. D.8: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(j_2, \ell)_{top}$ .

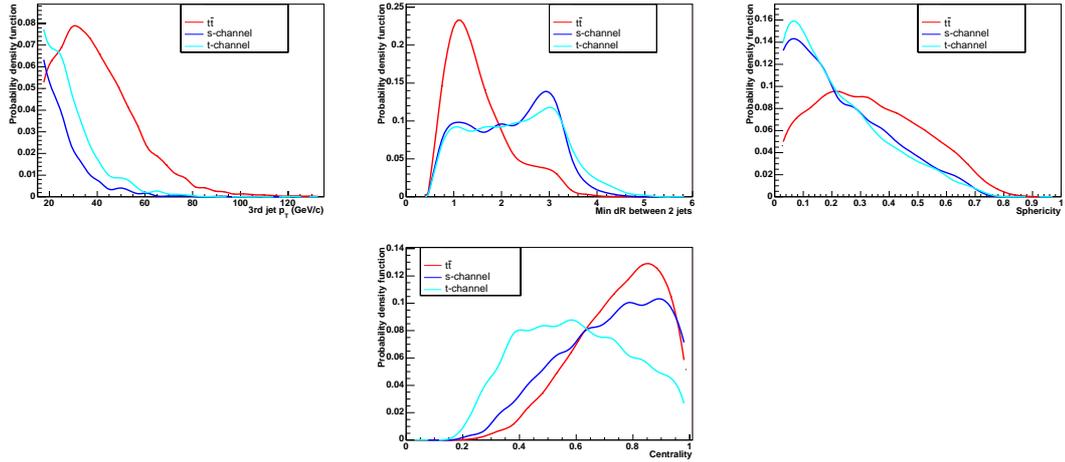


FIG. D.9: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ , *Sphéricité* et *Centralité*.

## Canal muon, STagVeto

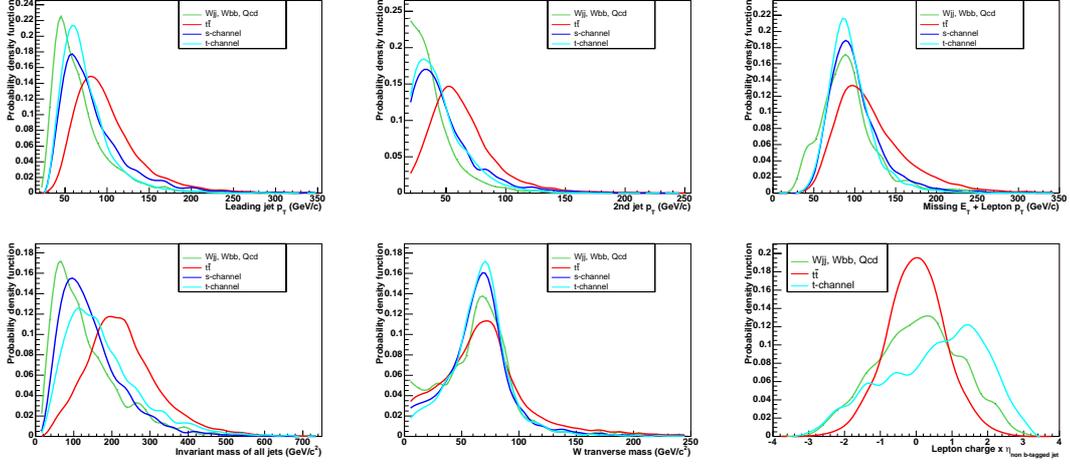


FIG. D.10: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

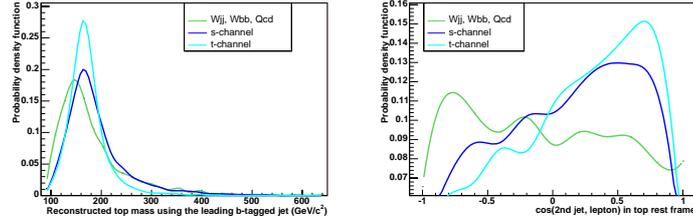


FIG. D.11: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(\widehat{j_2, \ell})_{top}$ .

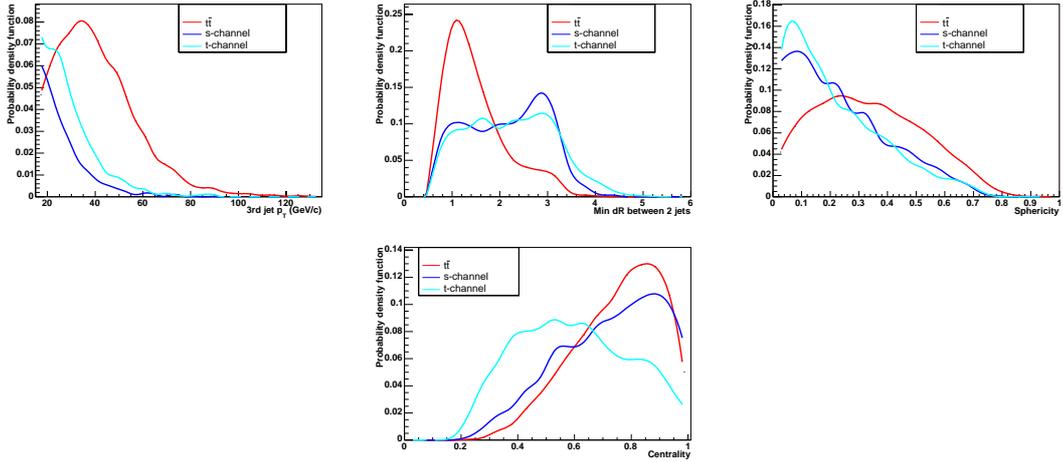


FIG. D.12: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphé$ ricité et  $Centralité$ .

## Canal électron, DTag

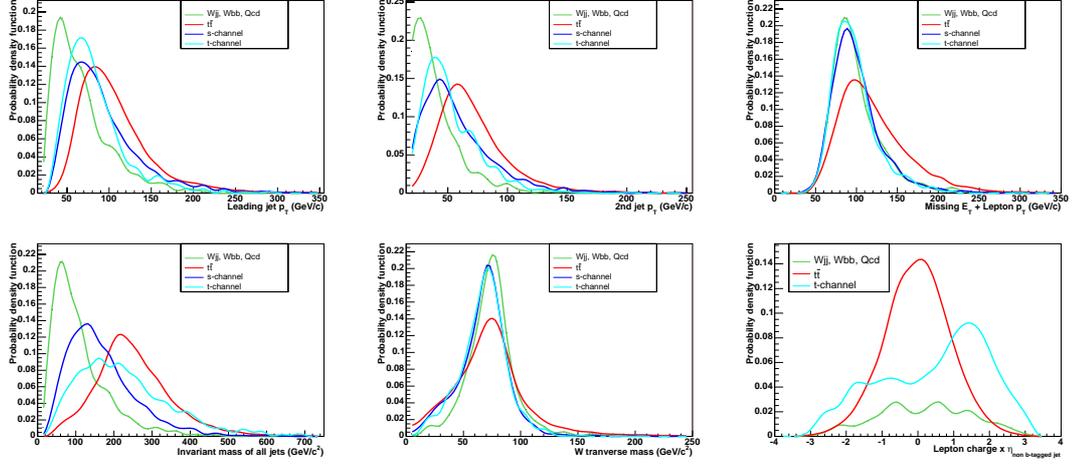


FIG. D.13: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

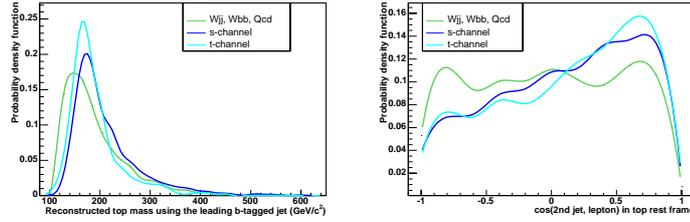


FIG. D.14: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(\widehat{j_2}, \ell)_{top}$ .

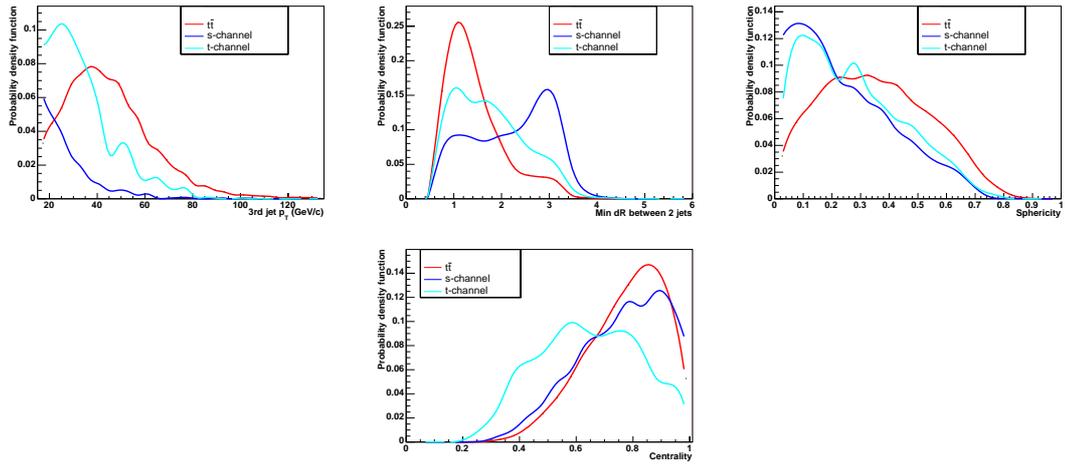


FIG. D.15: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphé$ ricité et  $Centralité$ .

## Canal muon, DTag

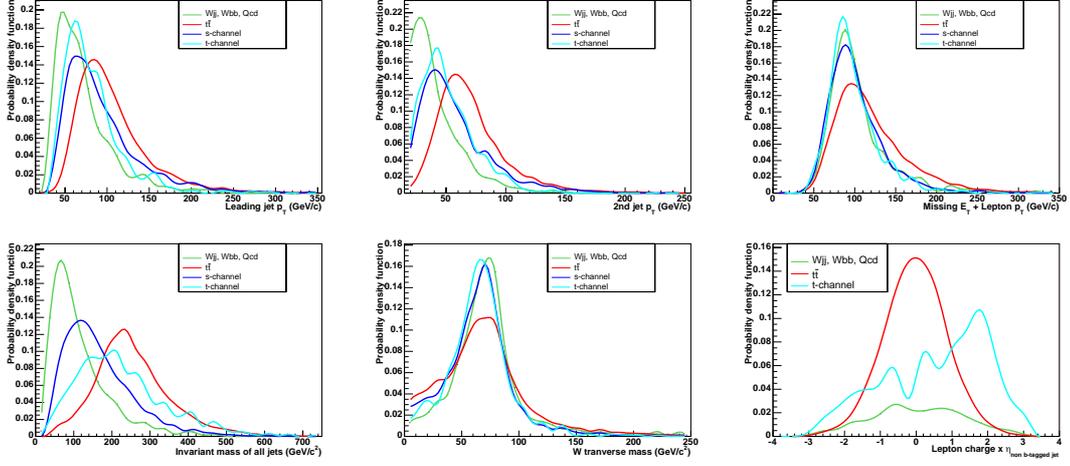


FIG. D.16: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

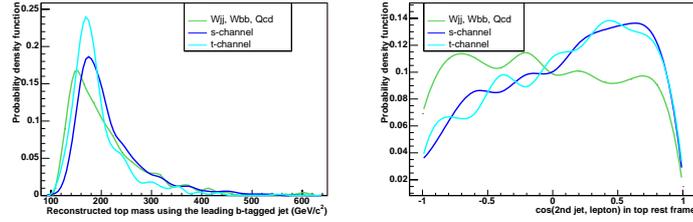


FIG. D.17: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(j_2, \ell)_{top}$ .

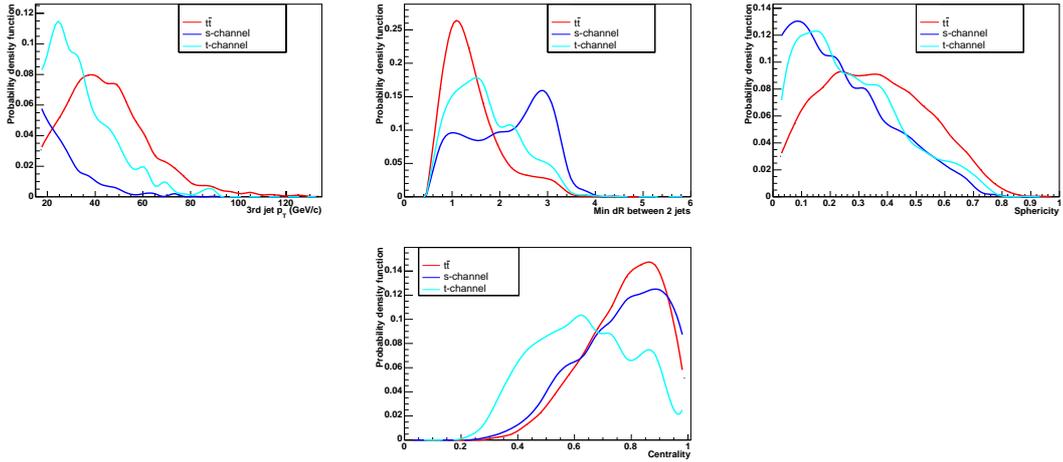


FIG. D.18: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphé$ ricité et  $Centralité$ .

## Canal électron, DTagAsym

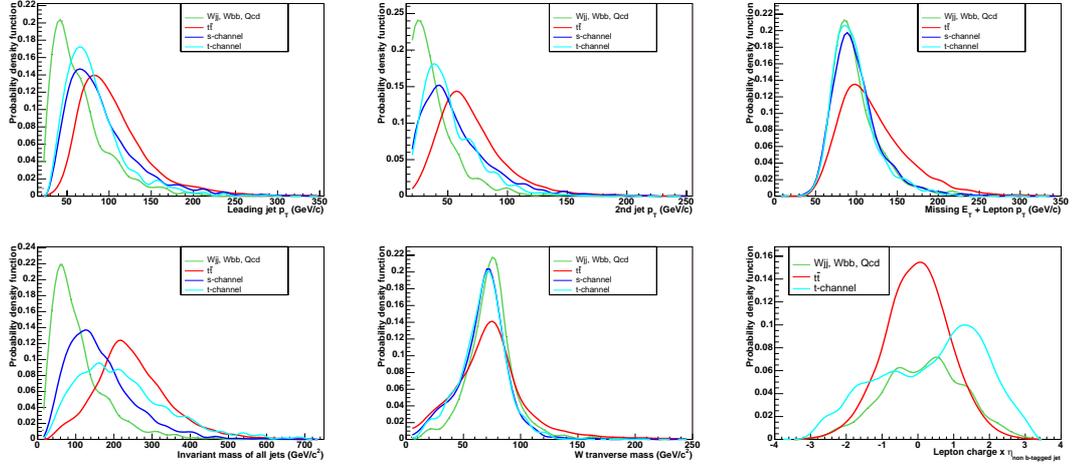


FIG. D.19: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

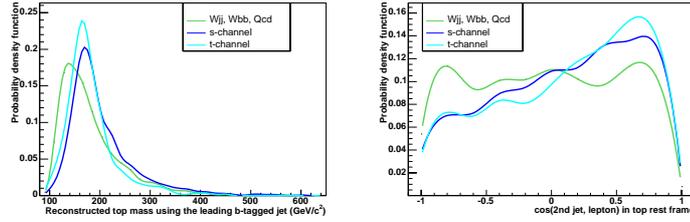


FIG. D.20: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(\widehat{j_2}, \ell)_{top}$ .

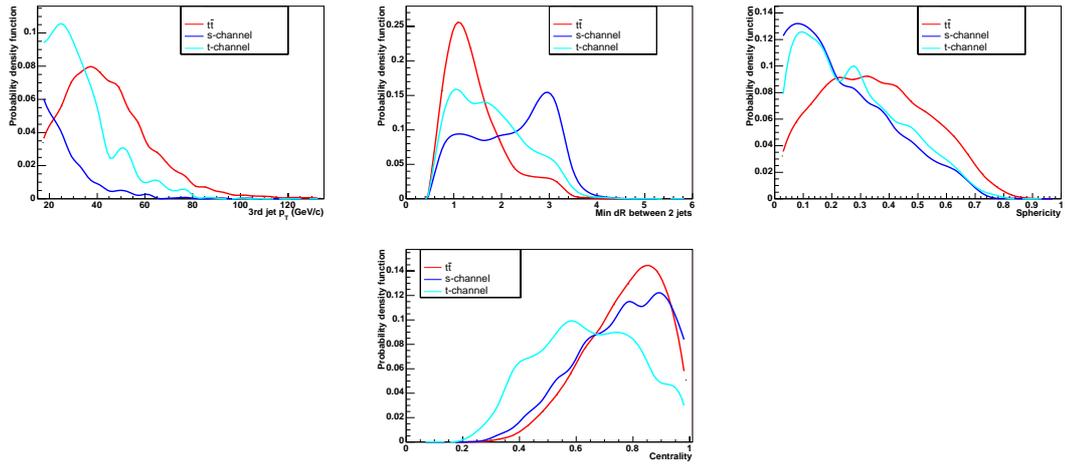


FIG. D.21: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphé$ ricité et  $Centralité$ .

## Canal muon, DTagAsym

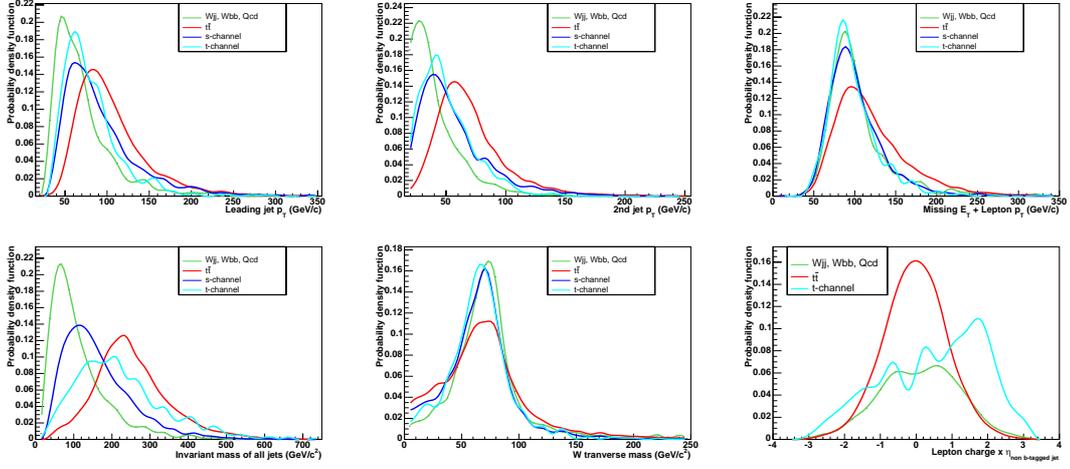


FIG. D.22: Densités de probabilité utilisées dans les filtres  $W$ +jets et  $t\bar{t}$  :  $p_T^{j1}$ ,  $p_T^{j2}$ ,  $METLep$ ,  $M_{jets}$ ,  $M_T^W$  et  $Q \times \eta$ . La variable  $Q \times \eta$  ne sert que pour la voie  $t$ .

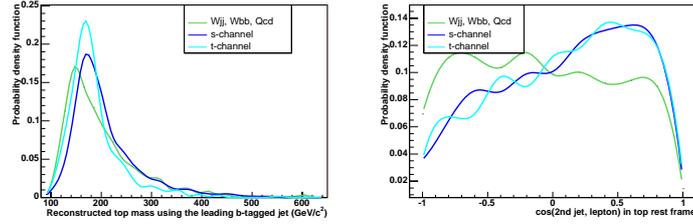


FIG. D.23: Densités de probabilité utilisées dans les filtres  $W$ +jet uniquement :  $M_{Wb}$  et  $\cos(j_2, \ell)_{top}$ .

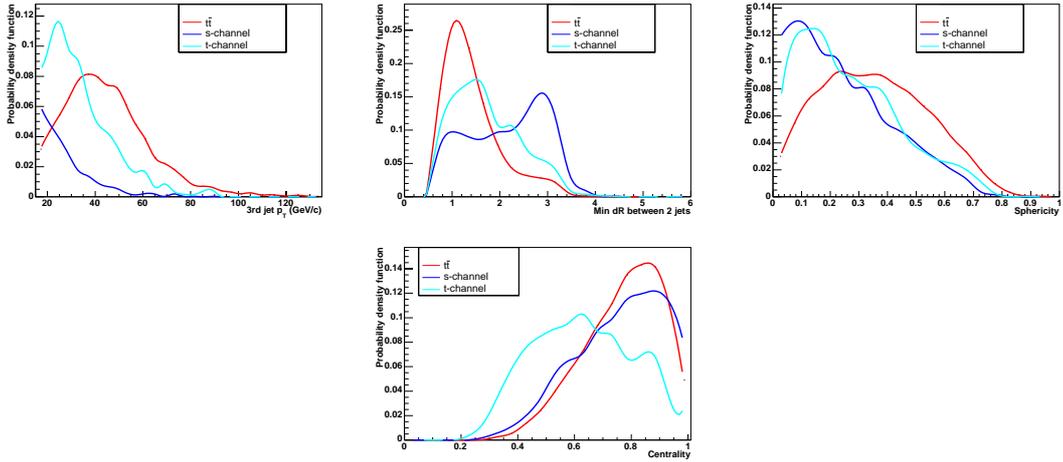


FIG. D.24: Densités de probabilité utilisées dans les filtres  $t\bar{t}$  uniquement :  $p_T^{j3}$ ,  $\Delta R_{min}^{jj}$ ,  $Sphé$ ricité et  $Centralité$ .



# E

## Limite bayésienne

### E.1 La méthode bayésienne

Il existe deux approches conceptuellement différentes du traitement de données statistiques : la méthode fréquentiste et la méthode bayésienne. Ces deux approches sont discutées dans la référence [103]. Suivant les recommandations du groupe de travail sur les intervalles de confiance du Run I de DØ [111], notre analyse utilise une méthode entièrement bayésienne pour déterminer les niveaux de confiance sur les sections efficaces de production *single top*. Le théorème de Bayes permet d'exprimer une probabilité conditionnelle d'une hypothèse  $A$  sachant que l'hypothèse  $B$  est vérifiée par  $P(A|B)$  en fonction de la probabilité de l'hypothèse  $B$  sachant que  $A$  est vérifiée. Ce théorème se déduit en écrivant de deux manières la probabilité que les deux hypothèses  $A$  et  $B$  soient vérifiées :

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) \quad (\text{E.1})$$

$$\Rightarrow P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (\text{E.2})$$

Le dénominateur peut être déterminé par la condition de normalisation

$$\sum_{A_i} P(A_i|B) = 1 \quad (\text{E.3})$$

où l'ensemble des  $A_i$  constitue une partition de l'univers  $\Omega$ .

Pour la mesure d'une section efficace par comptage, l'objectif est de déterminer la valeur de section efficace la plus probable ou la valeur limite de la section efficace au delà

de laquelle le signal recherché peut être exclu avec un niveau de confiance donné. La probabilité d'observer  $\mathcal{D}$  événements pour  $\mathcal{Y}(\sigma, \mathbf{x})$  événements de signal et de bruit de fond attendus peut être décrite par une distribution de Poisson :

$$P(\mathcal{D}|\mathcal{Y}(\sigma, \mathbf{x}), I) = P(\mathcal{D}|\sigma, \mathbf{x}, I) = \frac{e^{-\mathcal{Y}(\sigma, \mathbf{x})} \mathcal{Y}(\sigma, \mathbf{x})^{\mathcal{D}}}{\mathcal{D}!} \quad (\text{E.4})$$

Le nombre d'événements attendus dépend de la section efficace recherchée  $\sigma$  et d'un ensemble de paramètres  $\mathbf{x}$  dits paramètres de nuisance. La condition  $I$  regroupe l'ensemble des hypothèses émises pour déterminer les différents paramètres  $\mathbf{x}$ . Pour alléger les notations, cette condition n'apparaîtra plus explicitement par la suite. La probabilité d'avoir  $\mathcal{Y}(\sigma, \mathbf{x})$  événements connaissant  $\mathcal{D}$ , peut alors être déduite de la probabilité précédente par le théorème de Bayes :

$$P(\sigma, \mathbf{x}|\mathcal{D}) = \frac{P(\mathcal{D}|\sigma, \mathbf{x}) \cdot P(\sigma, \mathbf{x})}{P(\mathcal{D})} \quad (\text{E.5})$$

La probabilité  $P(\sigma, \mathbf{x}) = \pi(\sigma, \mathbf{x}) d\sigma d\mathbf{x}$  est la probabilité a priori de la section efficace  $\sigma$ . Pour utiliser la méthode bayésienne il est nécessaire de faire une hypothèse a priori sur la section efficace et les paramètres de nuisance. En supposant que les paramètres de nuisance sont indépendants de la section efficace, la densité de probabilité a priori  $\pi(\sigma, \mathbf{x})$  peut se factoriser en :

$$\pi(\sigma, \mathbf{x}) = \pi(\sigma)\pi(\mathbf{x}) \quad (\text{E.6})$$

La probabilité a priori sur les paramètres de nuisance permet de prendre en compte les incertitudes systématiques. On a alors :

$$P(\sigma, \mathbf{x}|\mathcal{D}) \propto \frac{e^{-\mathcal{Y}(\sigma, \mathbf{x})} \mathcal{Y}(\sigma, \mathbf{x})^{\mathcal{D}}}{\mathcal{D}!} \pi(\sigma)\pi(\mathbf{x}) d\sigma d\mathbf{x} \quad (\text{E.7})$$

La constante de normalisation est imposée par la condition d'unitarité :

$$\int_{\sigma} \int_{\mathbf{x}} f(\sigma, \mathbf{x}|\mathcal{D}) d\sigma d\mathbf{x} = 1 \quad (\text{E.8})$$

$$\text{avec } P(\sigma, \mathbf{x}|\mathcal{D}) = f(\sigma, \mathbf{x}|\mathcal{D}) d\sigma d\mathbf{x} \quad (\text{E.9})$$

La probabilité a posteriori sur la section efficace, s'obtient alors en intégrant sur les paramètres de nuisance :

$$P(\sigma|\mathcal{D}) = f(\sigma|\mathcal{D}) d\sigma = d\sigma \int_{\mathbf{x}} f(\sigma, \mathbf{x}|\mathcal{D}) d\mathbf{x} \quad (\text{E.10})$$

La section efficace mesurée (valeur la plus probable) est donnée par la valeur maximale de la densité de probabilité a posteriori  $f(\sigma|\mathcal{D})$ . La limite supérieure à 95% de niveau de confiance,  $\Sigma$  vérifie alors :

$$\int_0^{\Sigma} f(\sigma|\mathcal{D}) d\sigma = 0.95 \quad (\text{E.11})$$

## E.2 Application à la recherche du *single top*

Pour un canal d'analyse et un schéma d'étiquetage, les cellules de la distribution à deux dimensions des fonctions de vraisemblance sont traitées comme des canaux statistiquement décorrélés. Les canaux d'analyse (électron et muon) et les schémas d'étiquetage sont également statistiquement indépendants. La probabilité d'observer un ensemble de distributions  $\mathcal{D}$  dans les canaux combinés pour un ensemble  $\mathcal{Y}(\sigma, \mathbf{x})$  de distributions attendues est donnée par le produit sur toutes les cellules et tous les canaux d'analyse des probabilités pour chaque cellule (représentée par une distribution de Poisson) :

$$P(\mathcal{D}|\mathcal{Y}(\sigma, \mathbf{x})) = P(\mathcal{D}|\sigma, \mathbf{x}) = \prod_{c=e,\mu}^{\text{canaux}} \prod_{s=\text{STag, DTag}}^{\text{schémas}} \prod_k^{\text{cellules}} \frac{e^{-y_k^{c,s}} y_k^{c,s} d_k^{c,s}}{d_k^{c,s}!} \quad (\text{E.12})$$

Les grandeurs  $d_k^{c,s}$  et  $y_k^{c,s}$  représentent respectivement les nombres d'événements observés et attendus dans la  $k^{\text{ème}}$  cellule de la distribution des fonctions de vraisemblance dans le canal  $c$  et le schéma d'étiquetage  $s$ . Le nombre d'événements attendus est alors :

$$y_k^{c,s} = a_k^{c,s} \sigma + \sum_{i=1}^N b_k^{c,s} i \quad (\text{E.13})$$

où :

- $\sigma$  est la section efficace à déterminer,
- $a_k^{c,s} = \mathcal{L}^c \alpha_k^{c,s}$  est la luminosité effective, c'est-à-dire le produit de la luminosité intégrée  $\mathcal{L}^c$  et de l'acceptance du signal  $\alpha_k^{c,s}$ ,
- $N$  le nombre total de bruits de fond,
- $b_k^{c,s} i$  le nombre d'événements attendus pour le fond  $i$ .

Le terme  $\mathbf{x}$  dans l'équation E.12 représente l'ensemble des paramètres de nuisance  $\alpha_k^{c,s}$ ,  $\mathcal{L}^c$  et  $b_k^{c,s} i$ .

La densité de probabilité a posteriori de la section efficace s'écrit alors :

$$f(\sigma|\mathcal{D}) = \frac{L(\mathcal{D}|\sigma)\pi(\sigma)}{\int_{\sigma'} L(\mathcal{D}|\sigma')\pi(\sigma')d\sigma'} \quad (\text{E.14})$$

où :

$$L(\mathcal{D}|\sigma) = \int_{\mathbf{x}} P(\mathcal{D}|\sigma, \mathbf{x})\pi(\mathbf{x})d\mathbf{x} \quad (\text{E.15})$$

La densité de probabilité a priori pour la section efficace est choisie constante jusqu'à une certaine valeur  $\sigma_{max}$ , raisonnablement grande :

$$\pi(\sigma) = \begin{cases} \frac{1}{\sigma_{max}} & \text{pour } \sigma \leq \sigma_{max} \\ 0 & \text{pour } \sigma > \sigma_{max} \end{cases} \quad (\text{E.16})$$

La densité de probabilité a priori pour les paramètres de nuisance  $\pi(\mathbf{x})$  est déterminée à partir de l'estimation de chacun des paramètres et de l'erreur associée. Elle est modélisée

par un nuage de points ( $\mathbf{x}_n$ ). Chaque point peut s'écrire  $\mathbf{x}_n = \mathbf{x} + \delta\mathbf{x}_n$  où  $\delta\mathbf{x}_n$  est une variable aléatoire suivant une distribution de probabilité reflétant l'erreur sur  $\mathbf{x}$ . Cette variation se compose de deux termes,  $\delta\mathbf{x}_n = \delta\mathbf{x}_n^{\text{normalisation}} + \delta\mathbf{x}_n^{\text{forme}}$ . Le terme de normalisation  $\delta\mathbf{x}_n^{\text{normalisation}}$  est commun à toutes les cellules d'une distribution. Le terme de forme  $\delta\mathbf{x}_n^{\text{forme}}$  varie entre cellules et résulte des erreurs systématiques, d'identification et de calibration des jets, de modélisation du déclenchement et des poids d'étiquetage des jets  $b$  (voir § 4.9) :

- Le terme de normalisation suit une loi de probabilité gaussienne multivariée décrite par une matrice de covariance qui décrit les corrélations d'erreur entre les différents paramètres de nuisance :

$$c_{ij} = \sum_s^{\text{sources}} \Delta x_i^s \Delta x_j^s \quad (\text{E.17})$$

où  $\Delta x_i^s$  est l'erreur absolue sur le paramètre de nuisance  $i$  pour la source d'erreur  $s$ .

- Le terme de forme est estimé en déterminant un jeu de distributions des fonctions de vraisemblance déformées correspondant à des variations de  $\pm 1$  écart standard de chacune des sources d'erreur de forme. Ces distributions permettent également de déterminer une matrice de covariance qui décrit les corrélations entre chaque cellule et permet de générer les variations  $\delta\mathbf{x}_n^{\text{forme}}$ .

L'intégrale de l'équation E.15 peut être discrétisée en :

$$L(\mathcal{D}|\sigma) = \frac{1}{N} \sum_{n=1}^N (P(D|\sigma, \mathbf{x}_n)) \quad (\text{E.18})$$

où  $N$  est la dimension du nuage de points.  $N$  doit être suffisamment grand pour pouvoir approximer correctement l'intégrale. La valeur  $N = 2000$  utilisée pour l'analyse est suffisante. La limite  $\Sigma$  à 95% de niveau de confiance est finalement déterminée en résolvant l'équation E.11. Cette équation peut se réécrire dans notre cas :

$$0.95 = \int_0^{\Sigma} f(\sigma|\mathcal{D}) d\sigma \quad (\text{E.19})$$

$$= \int_0^{\Sigma} \frac{L(\mathcal{D}|\sigma)\pi(\sigma)}{\int_{\sigma'} L(\mathcal{D}|\sigma')\pi(\sigma') d\sigma'} d\sigma \quad (\text{E.20})$$

$$= \frac{\int_0^{\Sigma} L(\mathcal{D}|\sigma)\pi(\sigma) d\sigma}{\int_0^{\sigma_{max}} L(\mathcal{D}|\sigma')\pi(\sigma') d\sigma'} \quad (\text{E.21})$$

L'équation E.21 montre que la valeur maximale de la section efficace dans la densité de probabilité a priori doit être choisie suffisamment grande pour que  $L(\mathcal{D}|\sigma)$  pour  $\sigma > \sigma_{max}$  soit négligeable. La limite trouvée doit satisfaire  $\Sigma \ll \sigma_{max}$ . Pour le calcul des limites combinées de l'analyse, nous utilisons  $\sigma_{max} = 40$ .

# Table des figures

1.1	Vertex QED . . . . .	6
1.2	Corrections radiatives QED . . . . .	7
1.3	Constantes de couplage mobiles . . . . .	8
1.4	Couplages à 3 et 4 gluons . . . . .	10
1.5	Fonctions de densité de partons CTEQ5 . . . . .	12
1.6	Incertitude sur les <i>PDF</i> CTEQ5 . . . . .	12
1.7	Diagrammes de production de paire de top . . . . .	13
1.8	Vertex électrofaibles . . . . .	16
1.9	Diagrammes de production du Single Top . . . . .	16
1.10	<i>Single top</i> voie t au Tevatron . . . . .	17
1.11	Mesures de $\sigma_{t\bar{t}}$ à DØ . . . . .	19
1.12	Potentiel de Higgs . . . . .	20
1.13	Corrections radiatives à la masse du $W$ : top . . . . .	22
1.14	Corrections radiatives à la masse du $W$ : Higgs . . . . .	22
1.15	Mesures de la masse du quark top . . . . .	23
1.16	Contraintes sur la masse du boson de Higgs . . . . .	24
1.17	Distribution angulaire pour la désintégration de $W$ . . . . .	25
1.18	Déviations par rapport au Modèle Standard des mesures électrofaibles . . . . .	27
1.19	Constantes de couplages mobiles dans le MSSM . . . . .	29
2.1	Fermilab et le Tevatron . . . . .	32
2.2	Fermilab et le Tevatron . . . . .	33
2.3	Production d'antiprotons . . . . .	35
2.4	Le détecteur DØ . . . . .	37
2.5	Schéma du détecteur de traces . . . . .	38
2.6	Schéma du <i>SMT</i> . . . . .	39
2.7	Un tonneau du <i>SMT</i> . . . . .	39
2.8	Une section de cylindre du <i>CFT</i> . . . . .	40
2.9	Carte du champ magnétique du solénoïde . . . . .	41
2.10	Efficacité de reconstruction des traces . . . . .	42
2.11	Résolution en impulsion des traces . . . . .	42
2.12	Quelques caractéristiques des vertex reconstruits . . . . .	44

2.13	Tomographie du SMT . . . . .	44
2.14	Struture du détecteur de pieds de gerbe . . . . .	45
2.15	Le calorimètre de DØ . . . . .	46
2.16	Coupe longitudinale du calorimètre . . . . .	47
2.17	Correction d'énergie des jets . . . . .	51
2.18	Le spectromètre à muons . . . . .	54
2.19	Le luminomètre . . . . .	55
2.20	La chaîne de déclenchement . . . . .	56
3.1	$p_{Trel}$ du muon dans un jet . . . . .	60
3.2	Paramètre d'impact signé . . . . .	62
3.3	Significance signée pour différents saveurs . . . . .	63
3.4	Efficacités de <i>taggabilité</i> . . . . .	67
3.5	Reconstruction des $V^0$ . . . . .	69
3.6	Fractions de $V^0$ dans les données . . . . .	70
3.7	Ajustements des distributions de significance . . . . .	71
3.8	Corrections de l'erreur sur la significance . . . . .	72
3.9	Résolution corrigée sur le paramètre d'impact . . . . .	73
3.10	Fonctions de résolution . . . . .	73
3.11	Probabilité de temps de vol par trace . . . . .	75
3.12	Efficacité de <i>JLIP</i> en fonction de $\kappa_{cl}^{(B,J)}$ . . . . .	79
3.13	Efficacités de <i>JLIP</i> dans les données réelles . . . . .	81
3.14	Facteur de corrélation $\kappa_b^{(B,J)}$ . . . . .	83
3.15	Facteur de corrélation $\kappa_b^{(J,M)}$ . . . . .	84
3.16	Test de <i>System8</i> sur des événements simulés . . . . .	85
3.17	Efficacité dans la simulation . . . . .	88
3.18	Rapports d'efficacité entre différents processus dans la simulation . . . . .	89
3.19	Facteur correctif $SF_b$ . . . . .	91
3.20	Efficacité d'identification des jets $b$ . . . . .	92
3.21	Efficacité d'identification des jets $c$ . . . . .	93
3.22	Taux de mauvaise identification pour <i>JLIP Medium</i> . . . . .	95
3.23	Différence d'efficacités d'étiquetage négatif entre QCD et EM1TRK . . . . .	96
3.24	Topologie des lots QCD et EM1TRK . . . . .	97
3.25	Taux de mauvaise identification . . . . .	99
3.26	Histogrammes de contrôle des <i>TRF</i> . . . . .	101
3.27	Test de la factorisation des <i>TRF</i> . . . . .	103
3.28	Erreur totale sur les <i>TRF</i> . . . . .	104
3.29	Performances de <i>JLIP</i> et <i>NN</i> . . . . .	106
3.30	Performances de <i>JLIP</i> . . . . .	107
4.1	Diagrammes de production du Single Top . . . . .	110
4.2	Distributions en $p_T$ et $\eta$ au niveau partonique . . . . .	112

4.3	Exemples de bruits de fond . . . . .	113
4.4	Structure de l'analyse . . . . .	115
4.5	Coupures de sélection : vertex primaire . . . . .	119
4.6	Coupures de sélection : qualité de l'électron . . . . .	120
4.7	Coupures de sélection : isolation du muon . . . . .	121
4.8	Coupures de sélection : jets . . . . .	123
4.9	Coupures de sélection : énergie manquante . . . . .	124
4.10	Coupures triangulaires, canal électron . . . . .	125
4.11	Coupures triangulaires, canal muon . . . . .	126
4.12	Biais électromagnétique des jets . . . . .	131
4.13	Efficacité de déclenchement en fonction $\eta$ . . . . .	132
4.14	Efficacités de CJT(1,5) . . . . .	133
4.15	Efficacités de JET(1,10) . . . . .	133
4.16	Efficacités de JET(1,15,3) . . . . .	134
4.17	Efficacités de JET(1,15,3) . . . . .	134
4.18	<i>Taggabilité</i> dans le canal électron . . . . .	135
4.19	<i>Taggabilité</i> dans le canal muon . . . . .	135
4.20	Validation des paramétrisations de la <i>taggabilité</i> , canal électron . . . . .	137
4.21	Validation des paramétrisations de la <i>taggabilité</i> , canal muon . . . . .	138
4.22	Taux d'identification d'un faux muon isolé . . . . .	143
4.23	Décomposition de la masse transverse du boson $W$ . . . . .	145
4.24	Validation de la <i>TRF</i> inclusive au Run I . . . . .	146
4.25	<i>TRF</i> inclusive pour <i>JLIP Tight</i> . . . . .	147
4.26	Histogrammes de contrôle pour la <i>TRF</i> inclusive <i>JLIP Tight</i> . . . . .	147
4.27	Coupures de sélection du lot multijet . . . . .	150
4.28	Variable discriminantes : $p_T^{j_1}, p_T^{j_2}, p_T^{j_3}$ et $METLep$ . . . . .	154
4.29	Variables discriminantes : $M_{jets}, M_T^W, M_{Wb}$ et $\Delta R_{min}^{jj}$ . . . . .	154
4.30	Variables discriminantes : $\cos(\widehat{j_2, \ell})_{top}, \mathcal{S}, \mathcal{C}$ et $Q \times \eta$ . . . . .	155
4.31	Courbes de performances : filtre $t-t\bar{t}$ . . . . .	158
4.32	Biais des fonctions de vraisemblance : voie $t$ . . . . .	159
4.33	Filtres voie $s$ , STag et DTag . . . . .	160
4.34	Filtres voie $t$ , STag et DTag . . . . .	161
4.35	Filtres voie $s$ , STagVeto et DTagAsym . . . . .	162
4.36	Filtres voie $t$ , STagVeto et DTagAsym . . . . .	163
4.37	Densités de probabilité a posteriori : étiquetage standard, voie $s$ . . . . .	174
4.38	Densités de probabilité a posteriori : étiquetage standard, voie $t$ . . . . .	175
4.39	Densités de probabilité a posteriori : étiquetage asymétrique, voie $s$ . . . . .	176
4.40	Densités de probabilité a posteriori : étiquetage asymétrique, voie $t$ . . . . .	177
4.41	Limites à 2D . . . . .	179
4.42	Densités de probabilité a posteriori pour les limites inclusives . . . . .	181
4.43	Densités de probabilité : limites finales . . . . .	183

4.44	Projection de la signficance des analyses <i>single top</i> . . . . .	184
B.1	Sélection du $p_Z$ du neutrino . . . . .	192
B.2	Distributions de contrôle . . . . .	193
C.1	Histogrammes de contrôle : avant étiquetage (1) . . . . .	198
C.2	Histogrammes de contrôle : avant étiquetage (2) . . . . .	199
C.3	Histogrammes de contrôle : avant étiquetage (3) . . . . .	200
C.4	Histogrammes de contrôle : avant étiquetage (4) . . . . .	201
C.5	Histogrammes de contrôle : canal électron (1) . . . . .	202
C.6	Histogrammes de contrôle : canal électron (2) . . . . .	203
C.7	Histogrammes de contrôle : canal électron (3) . . . . .	204
C.8	Histogrammes de contrôle : canal électron (4) . . . . .	205
C.9	Histogrammes de contrôle : canal muon (1) . . . . .	206
C.10	Histogrammes de contrôle : canal muon (2) . . . . .	207
C.11	Histogrammes de contrôle : canal muon (3) . . . . .	208
C.12	Histogrammes de contrôle : canal muon (4) . . . . .	209
C.13	Histogrammes de contrôle : étiquetage standard (1) . . . . .	210
C.14	Histogrammes de contrôle : étiquetage standard (2) . . . . .	211
C.15	Histogrammes de contrôle : étiquetage standard (3) . . . . .	212
C.16	Histogrammes de contrôle : étiquetage standard (4) . . . . .	213
C.17	Histogrammes de contrôle : étiquetage asymétrique (1) . . . . .	214
C.18	Histogrammes de contrôle : étiquetage asymétrique (2) . . . . .	215
C.19	Histogrammes de contrôle : étiquetage asymétrique (3) . . . . .	216
C.20	Histogrammes de contrôle : étiquetage asymétrique (4) . . . . .	217
D.1	Densités de probabilité : électron, STag, filtres $W$ +jets et $t\bar{t}$ . . . . .	220
D.2	Densités de probabilité : électron, STag, filtres $W$ +jets . . . . .	220
D.3	Densités de probabilité : électron, STag, filtres $t\bar{t}$ . . . . .	220
D.4	Densités de probabilité : muon, STag, filtres $W$ +jets et $t\bar{t}$ . . . . .	221
D.5	Densités de probabilité : muon, STag, filtres $W$ +jets . . . . .	221
D.6	Densités de probabilité : muon, STag, filtres $t\bar{t}$ . . . . .	221
D.7	Densités de probabilité : électron, STagVeto, filtres $W$ +jets et $t\bar{t}$ . . . . .	222
D.8	Densités de probabilité : électron, STagVeto, filtres $W$ +jets . . . . .	222
D.9	Densités de probabilité : électron, STagVeto, filtres $t\bar{t}$ . . . . .	222
D.10	Densités de probabilité : muon, STagVeto, filtres $W$ +jets et $t\bar{t}$ . . . . .	223
D.11	Densités de probabilité : muon, STagVeto, filtres $W$ +jets . . . . .	223
D.12	Densités de probabilité : muon, STagVeto, filtres $t\bar{t}$ . . . . .	223
D.13	Densités de probabilité : électron, DTag, filtres $W$ +jets et $t\bar{t}$ . . . . .	224
D.14	Densités de probabilité : électron, DTag, filtres $W$ +jets . . . . .	224
D.15	Densités de probabilité : électron, DTag, filtres $t\bar{t}$ . . . . .	224
D.16	Densités de probabilité : muon, DTag, filtres $W$ +jets et $t\bar{t}$ . . . . .	225

---

D.17 Densités de probabilité : muon, DTag, filtres $W$ +jets . . . . .	225
D.18 Densités de probabilité : muon, DTag, filtres $t\bar{t}$ . . . . .	225
D.19 Densités de probabilité : électron, DTagAsym, filtres $W$ +jets et $t\bar{t}$ . . . . .	226
D.20 Densités de probabilité : électron, DTagAsym, filtres $W$ +jets . . . . .	226
D.21 Densités de probabilité : électron, DTagAsym, filtres $t\bar{t}$ . . . . .	226
D.22 Densités de probabilité : muon, DTagAsym, filtres $W$ +jets et $t\bar{t}$ . . . . .	227
D.23 Densités de probabilité : muon, DTagAsym, filtres $W$ +jets . . . . .	227
D.24 Densités de probabilité : muon, DTagAsym, filtres $t\bar{t}$ . . . . .	227



# Liste des tableaux

1.1	Fermions élémentaires . . . . .	4
1.2	Principaux lagrangiens . . . . .	5
1.3	Nombres quantiques de $SU(2)_L \times U(1)_Y$ . . . . .	14
1.4	Sections efficaces de production de <i>single top</i> . . . . .	17
2.1	Caractéristiques du Tevatron . . . . .	35
3.1	Les 29 catégories de traces . . . . .	68
3.2	Application numérique de <i>System8</i> . . . . .	80
3.3	Erreurs systématiques de <i>System8</i> . . . . .	86
3.4	Erreurs systématiques relatives sur $\varepsilon_b^{JLIP}$ et $\varepsilon_c^{JLIP}$ . . . . .	90
3.5	Erreurs systématiques relatives sur le taux de mauvaise identification. . . . .	98
3.6	Erreurs systématiques de factorisation. . . . .	102
3.7	Erreurs systématiques sur les <i>TRF</i> . . . . .	102
4.1	Sections efficaces des principaux bruits de fonds . . . . .	114
4.2	Lots de données simulées . . . . .	117
4.3	Critères de déclenchement . . . . .	118
4.4	Efficacités de sélection . . . . .	124
4.5	Acceptances Monte Carlo dans le canal électron . . . . .	141
4.6	Acceptances Monte Carlo dans le canal muon . . . . .	141
4.7	Efficacité d'identification d'un vrai lepton isolé . . . . .	143
4.8	Taux d'identification d'un faux lepton isolé . . . . .	143
4.9	Résultats de la méthode de la matrice . . . . .	144
4.10	Estimations du nombre d'événements multijet . . . . .	149
4.11	Nombres d'événements attendus dans le canal électron . . . . .	152
4.12	Nombres d'événements attendus dans le canal muon . . . . .	152
4.13	Jeux de variables des fonctions de vraisemblance . . . . .	157
4.14	Erreurs systématiques relatives, canal électron, STag . . . . .	166
4.15	Erreurs systématiques relatives, canal muon, STag . . . . .	166
4.16	Erreurs systématiques relatives, canal électron, DTag . . . . .	167
4.17	Erreurs systématiques relatives, canal muon, DTag . . . . .	167
4.18	Erreurs systématiques relatives, canal électron, STagVeto . . . . .	168

---

4.19	Erreurs systématiques relatives, canal muon, STagVeto . . . . .	168
4.20	Erreurs systématiques relatives, canal électron, DTagAsym . . . . .	169
4.21	Erreurs systématiques relatives, canal muon, DTagAsym . . . . .	169
4.22	Nombres d'événements attendus avec erreurs systématiques . . . . .	171
4.23	Limites par comptage simple . . . . .	172
4.24	Limites avec fonctions de vraisemblance . . . . .	173
4.25	Optimisation de l'analyse . . . . .	178
4.26	Limites inclusives . . . . .	181
4.27	Différences entre les analyses NN et LH . . . . .	182
4.28	Comparaison des limites pour $230 \text{ pb}^{-1}$ . . . . .	182

# Bibliographie

## Modèle Standard et physique du quark top

- [1] W. Greiner et B. Müller, “Quantum Mechanics, Symmetries, second edition”, Springer-Verlag (2004) ;  
W. Greiner et J. Reinhardt, “Quantum Electrodynamics, third edition”, Springer-Verlag (2002) ;  
W. Greiner et B. Müller, “Gauge Theory of Weak Interactions, third edition”, Springer-Verlag (2004) ;  
W. Greiner, S. Schramm et E. Stein, “Quantum Chromodynamics, second edition”, Springer-Verlag (2002).
- [2] G. Chanfray et G. Smadja, “Les Particules et leurs Symétries”, Masson (1997).
- [3] J.-P. Derendinger, “Théorie Quantique des Champs”, Presses polytechniques et universitaires romandes (2001).
- [4] S.L. Glashow, Nucl. Phys. **22**, 579 (1961) ;  
A. Salam et J.C. Ward, Phys. Rev. Lett. **13**, 168 (1964) ;  
S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).
- [5] S. Eidelman *et al.*, Phys. Lett. B **592**, 1 (2004).
- [6] The CDF Collaboration, the DØ Collaboration and the Tevatron Electroweak Working Group, hep-ex/0507091 (2005).
- [7] S. Willenbrock, “The standard model and the top quark”, Cours donné à la *12th Advanced Study Institute on Techniques and Concepts of High-Energy Physics, St. Croix, U.S. Virgin Islands, 13-24 Jun 2002*, hep-ph/0211067 (2002).
- [8] F. Abe *et al.*, (CDF Collaboration), “Observation of Top Quark Production in  $p\bar{p}$  Collisions,” Phys. Rev. Lett. **74**, 2626 (1995) ;  
S. Abachi *et al.*, (DØ Collaboration), “Observation of the Top Quark,” Phys. Rev. Lett. **74**, 2632 (1995).
- [9] M. Gell-Mann, Phys. Lett. **8**, 214 (1964).

- [10] V. E. Barnes *et al.*, Phys. Rev. Lett. **12**, 204 (1964).
- [11] S.D. Drell et T.M. Yan, Ann. Phys. **66** 578 (1971).
- [12] H. L. Lai *et al.* (CTEQ Collaboration), “Global QCD analysis of parton structure of the nucleon : CTEQ5 parton distributions,” Eur. Phys. J. C **12** 375 (2000).
- [13] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, “Parton distributions : A new global analysis,” Eur. Phys. J. C **4** 463 (1998).
- [14] A. Aktas *et al.* (H1 Collaboration), hep-ex/0507081, Eur. Phys. J. C **45** 23 (2006) ;  
A. Aktas *et al.* (H1 Collaboration), hep-ex/0411046, Eur. Phys. J. C **40** 349 (2005).
- [15] T. Sjöstrand *et al.*, “PYTHIA 6.2 : Physics and Manual,” hep-ph/0108264 (2001).
- [16] G. Corcella *et al.*, “HERWIG 6 : An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes),” JHEP **01** 010, hep-ph/0011363 (2001).
- [17] N. Kidonakis and R. Vogt, “Next-to-Next-to-Leading Order Soft-Gluon Corrections in Top Quark Hadroproduction,” Phys. Rev. D **68**, 114014 (2003).
- [18] N. Kidonakis, communication privée, (Février 2004).
- [19] C.S. Wu et al : Phys. Rev. **105** 1413 (1957).
- [20] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **122**, 103 (1983) ;  
M. Banner *et al.* (UA2 Collaboration), Phys. Lett. B **122**, 476 (1983).
- [21] G. Arnison *et al.* (UA1 Collaboration), Phys. Lett. B **126**, 398 (1983) ;  
P. Bagnaia *et al.* (UA2 Collaboration), Phys. Lett. B **129**, 130 (1983).
- [22] M.C. Smith et S. Willenbrock, “QCD and Yukawa Corrections to Single-Top-Quark Production via  $q\bar{q}$ ,” Phys. Rev. D **54**, 6696 (1996) ;  
T. Stelzer, Z. Sullivan et S. Willenbrock, “Single-Top-Quark Production via W-gluon Fusion at Next-to-Leading Order,” Phys. Rev. D **56**, 5919 (1997) ;  
B.W. Harris, E. Laenen, L. Phaf, Z. Sullivan, et S. Weinzierl, “Fully Differential Single-Top-Quark Cross Section in Next-to-Leading Order QCD,” Phys. Rev. D **66**, 054024 (2002) ;  
Z. Sullivan, “Understanding single-top-quark production and jets at hadron colliders,” Phys. Rev. D **70** 114012, hep-ph/0408049 (2004) ;  
Q.-H. Cao, R. Schwienhorst, et C.-P. Yuan, “Next-to-Leading Order Corrections to Single Top Quark Production and Decay at the Tevatron : 1. s-channel process,” Phys. Rev. D **71** 054023, hep-ph/0409040 (2005) ;  
Q.-H. Cao, R. Schwienhorst, J.A. Benitez, R. Brock et C.-P. Yuan, “Next-to-Leading Order Corrections to Single Top Quark Production and Decay at the Tevatron : 2. t-channel process,” Phys. Rev. D **72** 094027, hep-ph/0504230 (2005).

- [23] A.P. Heinson, A.S. Belyaev and E.E. Boos, "Single Top Quarks at the Fermilab Tevatron," *Phys. Rev. D* **56**, 3114 (1997).
- [24] Z. Sullivan, communication privée, (Février 2004).
- [25] The DØ Collaboration, "Measurement of the  $t\bar{t}$  Production Cross Section in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV in the all hadronic final state", DØ Note 4879-Conf, (Septembre 2005).
- [26] P.W. Higgs, *Phys. Lett.* **12** 132 (1964)  
P.W. Higgs, *Phys. Rev. Lett.* **13** 508 (1964)  
P.W. Higgs, *Phys. Rev.* **145** 1156 (1966)
- [27] F. Englert et R. Brout, *Phys. Rev. Lett.* **13** 321 (1964)
- [28] D. Bardin *et al.*, "Reports of the Working Group on Precision Calculations for the Z Resonance", CERN 95-03 (1995).
- [29] The LEP Electroweak Working Group, "A combination of preliminary electroweak measurements and constraints on the standard model," hep-ex/0511027 (2005).
- [30] A. Abulencia *et al.* (CDF Collaboration), "Measurement of the Helicity of W Bosons in Top-Quark Decays", Soumis à *Phys. Rev. Lett.*, hep-ex/0511023 (Novembre 2005).
- [31] The DØ Collaboration, "Search for Right-handed W Bosons in Dilepton Top Quark Pair Candidates", DØ Note 4839-Conf, (Juillet 2005).
- [32] N. Cabibbo, *Phys. Rev. Lett.* **10** 531 (1963);  
M. Kobayashi, T. Maskawa, *Prog. Theor. Phys.* **49** 652 (1973).
- [33] B. Pontecorvo, *Zh. Eksp. Teor. Fiz.* **33** 549(1957);  
Z. Maki, M. Nakagawa et S. Sakata, *Prog. Theor. Phys.* **28** 870 (1962).
- [34] N. Ellis et J. March-Russel (ed.), 2001 European School of High-Energy Physics, Proceedings, p. 213 (2001)

## Le Tevatron et DØ

- [35] <http://www.fnal.gov>
- [36] <http://www-d0.fnal.gov>
- [37] <http://www-cdf.fnal.gov>

- [38] “Operations Rookie Books”  
[http://www-bdnew.fnal.gov/operations/rookie\\_books/rbooks.html](http://www-bdnew.fnal.gov/operations/rookie_books/rbooks.html)  
– Concepts Rookie Book,  
– Linac Rookie Book,  
– Booster Rookie Book,  
– Antiproton Source Rookie Book,  
– Tevatron Rookie Book,  
– Main Injector Rookie Book.
- [39] S. Abachi *et al.* (DØ Collaboration), “The D0 Detector” Nucl. Instr. and Methods, A338, 185 (1994).
- [40] V.M. Abazov *et al.* (DØ Collaboration), “The Upgraded D0 Detector”, hep-ph/0507191, soumis à Nucl. Instr. and Methods (2005).
- [41] H. Greenlee, “Motion of a Charged Particle in a Magnetic Field”, DØ Note 4180 (2003).
- [42] H. Greenlee, “The D0 Kalman Track Fit”, DØ Note 4303 (2004).
- [43] A. García-Bellido *et al.*, “Primary Vertex certification in p14”, DØ Note 4320 (2004).
- [44] M. Narain et A. Schwartzman “Probabilistic Primary Vertex Selection”, DØ Note 4042 (2002).
- [45] S. Crépé-Renaudin, “Energy corrections for geometry effects for electrons in Run II”, DØ Note 4023 (2002);  
J. Zhu, “Determination of Electron Energy Scale and Energy Resolution using p14 Zee Data”, DØ Note 4323 (2004).
- [46] J.-L. Agram, “Mesure de la section efficace inclusive de production de jets en fonction de leur impulsion transverse dans l’expérience DØ au Fermilab.”, Thèse de doctorat, Université de Haute-Alsace, n<sup>o</sup> d’ordre 04MULH0777 (2004).
- [47] F. Beaudette et J.-F. Grivaz “The Road Method (an algorithm for the identification of electrons in jets)”, DØ Note 3976 (2002).
- [48] [http://www-d0.fnal.gov/phys\\_id/emid/d0\\_private/emid.html](http://www-d0.fnal.gov/phys_id/emid/d0_private/emid.html)
- [49] J. Gardner *et al.*, “Electron Likelihood in p14”, DØ Note 4449 (2004).
- [50] E. Busato, B. Andrieu “Jet Algorithms in DØ Run II Software : Description and User’s Guide”, DØ Note 4457 (2004).

- [51] G. Bernardi *et al.*, “NADA : A New Event by Event Hot Cell Killer”, DØ Note 3687 (1999);  
G. Bernardi *et al.*, “Improvement of the NADA Algorithm : Hot Cell Killing in DØ Run II Data”, DØ Note 4457 (2002).
- [52] G. Bernardi *et al.*, “Improvements from the T42 Algorithm on Calorimeter Objects Reconstruction”, DØ Note 4335 (2004);  
J.-R. Vlimant *et al.*, “Technical description of the T42 algorithm for the calorimeter noise suppression”, DØ Note 4146 (2003).
- [53] J.-L. Agram *et al.*, “Jet Energy Scale at DØ Run II”, DØ Note 4720 (2005).
- [54] The DØ Collaboration, “Top Quark Mass Measurement with the Matrix Element Method in the Lepton+Jets Final State at DØ Run II”, DØ Note 4874-Conf (2005).
- [55] N. Makovec et J.-F. Grivaz, “The Relative Data - Monte Carlo Jet Energy Scale”, DØ Note 480705).
- [56] S. Calvet, P. Verdier et E. Kajfasz “Towards MissingET Certification and Unclustered Energy Studies”, DØ Note 4927 (2005).
- [57] V.M. Abazov *et al* (The DØ Muon Group), “The Muon System of the Run II DØ Detector”, Nucl.Instrum.Meth. **A552** 372-398 (2005).
- [58] [http://www-d0.fnal.gov/computing/algorithms/muon/muon\\_algo.html](http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html)
- [59] T. Edwards *et al.* (DØ Collaboration), “Determination of the Effective Inelastic  $p\bar{p}$  Cross-Section for the DØ Run II Luminosity Measurement,” FERMILAB-TM-2278-E (2004).
- [60] T Edwards *et al.*, “The Updated DØ Luminosity Determination – Short Summary”, DØ Note 4328, (2004).
- [61] <http://www-d0.fnal.gov/computing/algorithms/status/index.html>
- [62] S. Baffioni, E. Nagy et S. Protopopescu, “ThumbNail : a compact data format”, DØ Note 3979 (2002).
- [63] <http://root.cern.ch>
- [64] S. Baffioni *et al.*, “The tmb\_tree package”, DØ Note 3978 (2002).
- [65] Markus Klute, Lukas Phaf et Daniel Whiteson, “TopAnalyze - A Framework Analyze Package For Top Group Analyses”, DØ Note 4122 (2003).
- [66] <http://www-d0.fnal.gov/computing/MonteCarlo/verification.htm>

- [67] R. Brun *et al.*, “GEANT - Detector Description and Simulation,” CERN Program Library Long Writeup W 5013 (1994).
- [68] <http://www-d0.fnal.gov/computing/trigsim/trigsim.html>
- [69] S. Jain, “Scale and Over-smearing for MC Electron”, DØ Note 4402 (2004).
- [70] E. Busato *et al.*, “d0correct v8”, DØ Note 4646 (2004).

## Etiquetage des jets de $b$

- [71] L. Feligioni *et al.*, “Update on  $b$ -quark jet identification with secondary Vertex reconstruction using D0RECO version p14”, DØ Note 4414, (2004).
- [72] ALEPH Collab., Phys. Lett. **B313** 535 (1993);  
G. Borisov and C. Mariotti,, Nucl. Instr. & Meth. **A372**) 181 (1996);  
DELPHI Collab., E. Phys. J. **C10** 415 (1999).
- [73] [http://www-d0.fnal.gov/Run2Physics/top/d0\\_private/wg/top\\_analyze/Ipanema\\_Updated/Ipanema.html](http://www-d0.fnal.gov/Run2Physics/top/d0_private/wg/top_analyze/Ipanema_Updated/Ipanema.html)
- [74] DØ Common Sample Group,  
<http://www-d0.fnal.gov/Run2Physics/cs/index.html>
- [75] L. Feligioni *et al.*, “Update on p14 Certification for the *CSIP* and *SVT*  $b$ -quark Jet Identification Algorithms”, DØ Note 4838 (2005).
- [76] R. Demina *et al.*, “Measurement of  $b$ -Tagging Efficiency and Mistagging Rates with the Counting Signed Impact Parameter Method”, DØ Note 4432 (2004).
- [77] M. Anastasoae et T. Scanlon, “Neural Network  $b$ -tagging Tool on Pass 2 Data”, DØ Note 4890 (2005).
- [78] S. Greder, Ph.D. thesis, Université Louis Pasteur, Strasbourg, IRES 05-006 N° d’ordre ULP 4652 (2005).
- [79] V.M. Abazov *et al.*, (DØ Collaboration), “A Search for  $Wb\bar{b}$  and  $WH$  Production in  $p$  Collisions at  $\sqrt{s} = 1.96$  TeV,” Phys. Rev. Lett. **94** 091802, hep-ex/0410062 (2005).
- [80] P. Bargassa, “Search of the lightest scalar top  $\tilde{t}_1$  in  $\tilde{t}_1\tilde{t}_1^*\bar{b}\mu^+\mu^-\tilde{\nu}\tilde{\nu}$  Decays at DØ ”,  
DØ Note 4851 (2005);  
The DØ Collaboration, DØ Note 4866-Conf (2005).
- [81] A. Duperrin et E. Kajfasz, “Search for Direct Production of Scalar Bottom Quark”,  
DØ Note 4805 (2005);  
The DØ Collaboration, DØ Note 4832-Conf (2005).

## Recherche du *single top*

- [82] B. Abbott *et al.*, (DØ Collaboration), “Search for Electroweak Production of Single Top Quarks in  $p\bar{p}$  Collisions,” Phys. Rev. D **63**, 031101 (2001);  
V.M. Abazov *et al.*, (DØ Collaboration), “Search for Single Top Quark Production at DØ Using Neural Networks”, Phys. Lett. B **517**, 282 (2001).
- [83] V.M. Abazov *et al.*, (DØ Collaboration), “Search for single top quark production in pbarp collisions at sqrt(s)=1.96 TeV”, Phys. Lett. B **622**, 265, hep-ex/0505063 (2005).
- [84] D. Acosta *et al.*, (CDF Collaboration), “Search for Single Top Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.8$  TeV”, Phys. Rev. D **65**, 091102 (2002).
- [85] D. Acosta *et al.*, (CDF Collaboration), “Search for Electroweak Single-Top-Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV”, Phys. Rev. D **71**, 012005 (2005).
- [86] V.M. Abazov *et al.*, (DØ Collaboration), “Measurement of  $\mathcal{B}(t \rightarrow Wb)/\mathcal{B}(t \rightarrow Wq)$  at  $\sqrt{s} = 1.96$  TeV”, submitted to Phys. Rev. Lett., hep-ex :0603002 (2006).
- [87] E. Boos *et al.*, (CompHEP Collaboration), “CompHEP 4.4 : Automatic Computations from Lagrangians to Events,” Nucl. Instrum. Meth. A **534**, 250 (2004);  
E.E. Boos, L.V. Dudko, and V.I. Savrin, “SingleTop’ — an Event Generator for the Single Top Quark Production at the LHC,” CMSNote 2000/065 (2000).
- [88] <http://mcfm.fnal.gov>
- [89] <http://www.slac.stanford.edu/~lange/EvtGen>
- [90] S. Jadach *et al.*, “Tauola 2.5”, CERN-TH-6793 (1992).
- [91] M.L. Mangano *et al.*, “ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions,” J. High Energy Physics **0307**, 001 (2003).
- [92] K. Ranjan *et al.* “Calorimeter Event Quality Using Level 1 Confirmation”, DØ Note 4554 (2005).
- [93] <http://www-d0.fnal.gov/Run2Physics/top/index.html>
- [94] Top Production group, Ipanema analyses webpage  
[http://www-d0.fnal.gov/Run2Physics/top/d0\\_private/wg/production/Ipanema\\_analyses/](http://www-d0.fnal.gov/Run2Physics/top/d0_private/wg/production/Ipanema_analyses/)
- [95] S. Lager, “Primary Vertex Cut Efficiencies and Scale Factors,” DØ Note xxxx, (2004),  
[http://www-clued0.fnal.gov/~lager/pv\\_eff.ps](http://www-clued0.fnal.gov/~lager/pv_eff.ps)

- [96] D0 Top Physics Working Group, "D0 Top Analyses and Data Sample for the Winter Conference 2004", DØ Note 4419 (2004).
- [97] R. Schwienhorst, "Top Trigger Selection and Application of Turn On Curves to the Monte Carlo", DØ Note 4508 (2004).
- [98] M. Agelou *et al.*, "Top Trigger Efficiency Measurements and the top\_trigger package", DØ Note 4512 (2004).
- [99] M. Agelou *et al.*, "Search for Single Top Quark Production at DØ in RunII", DØ Note 4398, (2004).
- [100] P. Tamburello, "Measurement of the Top Quark Production Cross Section in Proton Antiproton Collisions at 1.8 TeV Center of Mass Energy," Ph.D. dissertation, University of Maryland, (1997).
- [101] M. Agelou *et al.*, "Improved Search for Single Top Quark Production", DØ Note 4670 (2005).
- [102] V.M. Abazov *et al.*, (DØ Collaboration), "Measurement of  $t\bar{t}$  cross section using l+jets kinematic characteristics", Phys. Lett. B **626**, 45, hep-ex/0504043 (2005).
- [103] J. Linnemann, M. Paterno, et H.B. Prosper, "Calculating Confidence Limits", DØ Note 4491 (2004).
- [104] T. Tait et C. P. Yuan, "Single Top Quark Production as a Window to Physics Beyond the Standard Model," Phys. Rev. D **63**, 014018 (2001).
- [105] L. Montanet *et al.*, Phys. Rev. D **50**, 1173 (1994).
- [106] The CDF and DØ Collaboration, "The Case for Run II", document soumis au *Particule Physics Project Prioritization Panel* (2005)  
[http://www-cdf.fnal.gov/physics/projections/RunIIPhysicsCase\\_V8.pdf](http://www-cdf.fnal.gov/physics/projections/RunIIPhysicsCase_V8.pdf)

## Annexes

- [107] E. Barberis *et al.*, "The Matrix Method and its Error Calculation", DØ Note 4564 (2004).
- [108] R. Schwienhorst et E. Perez, "Uncertainty Calculation for Combined  $W$ +jets and QCD Backgrounds in Analyses with  $W$ +jets Signatures", DØ Note 4597 (2004).
- [109] S. Snyder, "Measurement of the Top Quark Mass at DØ ", Ph.D. dissertation, University of New York at Stony Brook (1997).

- 
- [110] B. Clément et D. Gelé, “Search for Single Top Quark Production using Likelihood Discriminants”, DØ Note 4825 (2005).
- [111] I. Bertram *et al.* (DØ Collaboration), “A Recipe for the Construction of Confidence Limits,” Fermilab-TM-2104 (2000).



---

## Résumé

Le travail présenté dans cette thèse porte sur l'étude de la production électrofaible du quark top dans les collisions proton-antiproton à  $\sqrt{s} = 1.96$  TeV. Ce mode de production du quark top n'a jamais été observé. Les données analysées ont été collectées au cours du Run II de l'expérience DØ auprès du collisionneur Tevatron du Fermilab et représentent une luminosité intégrée d'environ  $370 \text{ pb}^{-1}$ .

La désintégration du quark top dans le Modèle Standard produit systématiquement un quark beau de grande impulsion. L'identification des jets issus de l'hadronisation d'un quark beau joue ainsi un rôle central dans l'analyse. L'important temps de vol des hadrons beaux qui se traduit par des particules chargées de grand paramètre d'impact relativement au vertex primaire de la collision. Les paramètres d'impact des traces associées à un jet sont convertis en une probabilité pour le jet de provenir du vertex primaire. Cet algorithme atteint une efficacité d'étiquetage des jets beaux de 45% pour environ un taux de mauvaise identification de 0.5%. Au Tevatron, la production électrofaible du quark top est dominée par deux processus (voies  $s$  et  $t$ ), avec des états finals légèrement différents. La signature recherchée consiste en 2 à 4 jets, dont au moins un jet de quark beau, un lepton chargé (électron ou muon) et de l'énergie manquante provenant d'un neutrino. Cet état final est dominé par les deux principaux bruits de fond : la production associée d'un boson  $W$  et de jets et la production d'une paire de quarks top. Une analyse multivariable est mise en œuvre pour séparer le signal de ces fonds. La sensibilité de l'analyse ne permet pas encore la mise en évidence du signal et des limites supérieures à 95% de niveau de confiance sur les sections efficaces de production de 5 pb (voie  $s$ ) et 4.3 pb (voie  $t$ ) ont été déterminées.

**Mots-clés** Tevatron, Dzero, Modèle Standard, quark beau, étiquetage, quark top, électrofaible, analyse multivariable

## Abstract

The work exposed in this thesis deals with the search for electroweak production of top quark (single top) in proton-antiproton collisions at  $\sqrt{s} = 1.96$  TeV. This production mode has not been observed yet. Analyzed data have been collected during the Run II of the DØ experiment at the Fermilab Tevatron collider. These data correspond to an integrated luminosity of  $370 \text{ pb}^{-1}$ .

In the Standard Model, the decay of a top quark always produce a high momentum bottom quark. Therefore bottom quark jets identification plays a major role in this analysis. The large lifetime of  $b$  hadrons and the subsequent large impact parameters relative to the interaction vertex of charged particle tracks are used to tag bottom quark jets. Impact parameters of tracks attached to a jet are converted into the probability for the jet to originate from the primary vertex. This algorithm has a 45% tagging efficiency for a 0.5% mistag rate.

Two processes ( $s$  and  $t$  channels) dominate single top production with slightly different final states. The searched signature consists in 2 to 4 jets with at least one bottom quark jet, one charged lepton (electron or muon) and missing energy accounting for a neutrino. This final state is background dominated and multivariate techniques are needed to separate the signal from the two main backgrounds : associated production of a  $W$  boson and jets and top quarks pair production. The achieved sensitivity is not enough to reach observation and we computed upper limits at the 95% confidence level at 5 pb ( $s$ -channel) and 4.3 pb ( $t$ -channel) on single top production cross-sections.

**Keywords** Tevatron, Dzero, Standard Model, bottom quark, tagging, top quark, electroweak, multivariate analysis

