CPPM 2010

UNIVERSITE DE LA MEDITERRANEE
AIX-MARSEILLE II
FACULTE DES SCIENCES DE LUMINY
163, avenue de Luminy
13288 MARSEILLE cedex 09
FRANCE

THESE DE DOCTORAT

spécialité : Physique des Particules et Astroparticules

présentée par

Betty CALPAS

en vue d'obtenir le grade de docteur de l'Université de la Méditerranée

Recherche du boson de Higgs du modèle standard dans le canal $ZH \to e^+e^-b\bar{b}$ avec le détecteur DØ auprès du Tevatron.

Soutenue le 11 juin 2010

devant le jury composé de :

M. Bernard	Ille	
M. Eric	Kajfasz	Président
M. Elemér	Nagy	Directeur de thèse
M. Gérard	Sajot	Rapporteur
Mme. Isabelle	Wingerter-Seez	Rappoteur
M. Takahiro	Yasuda	

Table des matières

Remerciements 5
Introduction 7

1 Cadre théorique **9**
1.1 Le Modèle Standard 9
1.1.1 Les quarks et les leptons 9
1.1.2 Les interactions fondamentales 11
1.2 Théorie électrofaible et mécanisme de Higgs 12
1.2.1 Théorie électrofaible 12
1.2.2 Mécanisme de Higgs 12
1.2.2.1 Le champ de Higgs 12
1.2.2.2 Le boson de Higgs 15
1.3 La recherche du Higgs au Tevatron 17
1.3.1 Contraintes théoriques 17
1.3.2 Contraintes expérimentales directes 17
1.3.3 Contraintes expérimentales indirectes 18
1.3.4 Production et désintégration du Higgs au Tevatron 20
1.3.5 La recherche du Higgs dans le canal $ZH \to e^+e^-b\bar{b}$ 20
1.4 Conclusion 20

2 Dispositif expérimental **24**
2.1 Le complexe d'accélération de Fermilab 24
2.1.1 Production des protons 26
2.1.2 Production des antiprotons 26
2.1.3 Le Tevatron 27
2.2 Le détecteur DØ 30
2.2.1 La zone des collisions 30
2.2.2 La luminosité 32
2.2.2.1 Calcul de la luminosité 32
2.2.2.2 Mesure de la luminosité 32
2.2.2.3 Les périodes de prise de données 33
2.2.3 Les sous-détecteurs 34
2.2.3.1 Le détecteur de traces 34
2.2.3.2 Les détecteurs de pieds de gerbes 40
2.2.3.3 Le calorimètre 41
2.2.3.4 Les détecteurs de muons 51
2.2.4 Le système de déclenchement 51

2.2.5 Reconstruction des évènements et format des données 53
2.3 Conclusion 54

3 Objets physiques utilisés 57
3.1 Les traces 57
3.2 Les vertex primaire et secondaire 57
3.3 Les objets électromagnétiques 58
3.4 Les jets hadroniques 59
3.5 L'énergie transverse manquante 64
3.6 Conclusion 64

4 Correction de l'énergie mesurée dans le calorimètre 65
4.1 Méthode des trois échantillons 65
4.2 Echantillons de données 67
4.3 Sélection des évènements 67
4.3.1 Détermination du piédestal 67
4.3.2 Sélection du signal 67
4.4 Détermination des facteurs de correction 71
4.5 Conclusion 71

5 Certification de l'identification des électrons dans le calorimètre 72
5.1 Méthode "Tag and Probe" 72
5.2 Echantillons de données 73
5.3 Sélection des évènements 74
5.4 Les évènements simulés 75
5.5 Le bruit de fond multijet 75
5.6 Identification des électrons 76
5.6.1 Critères de sélection 76
5.6.2 Efficacités d'identification 77
5.6.3 Correction des efficacités 77
5.7 Conclusion 78

6 Certification de l'identification des électrons dans l'ICR 84
6.1 Introduction 84
6.2 Les objets taus 85
6.3 Echantillon de données 86
6.4 Sélection des évènements 86
6.5 Les évènements simulés 87
6.6 Efficacités 87
6.6.1 Efficacité du trajectographe 88
6.6.2 Efficacité de reconstruction des objets taus 88
6.6.3 Efficacité du réseau de neurones tau 90
6.6.4 Efficacité totale 92
6.7 Bruit de fond multijet 92
6.8 Facteurs de correction 92
6.9 Conclusion 93

7 Recherche du boson de Higgs dans le canal $ZH \to e^+e^-b\bar{b}$ 94
7.1 Principe de l'analyse . . . 95
7.2 Les données . . . 98
7.2.1 Echantillons de données . . . 98
7.2.2 Qualité des données et luminosité . . . 98
7.2.3 Pré-sélection des évènements . . . 99
7.2.3.1 Vertex d'interaction principal . . . 99
7.2.3.2 Electrons ee . . . 100
7.2.3.3 Electrons ee^{icr} . . . 100
7.2.3.4 Bosons Z . . . 101
7.2.3.5 Jets . . . 101
7.3 Les évènements simulés . . . 102
7.3.1 Génération des évènements simulés . . . 106
7.3.2 Corrections des évènements simulés . . . 106
7.4 Détermination du bruit de fond multijet . . . 109
7.4.1 Les évènements multijets dans l'analyse ee . . . 110
7.4.2 Les évènements multijets dans l'analyse ee^{icr} . . . 111
7.5 Normalisation des bruits de fond . . . 111
7.6 Distributions de contrôle avant étiquetage des jets de b . . . 117
7.7 Etiquetage des jets de b . . . 122
7.7.1 Principe de fonctionnement du b-tagging . . . 122
7.7.2 Le b-tagging dans l'analyse ZH . . . 123
7.7.3 Correction de l'étiquetabilité des jets . . . 125
7.8 Distributions de contrôle après étiquetage des jets de b . . . 125
7.9 Ajustement cinématique . . . 129
7.9.1 Principe . . . 129
7.9.2 Résultats . . . 130
7.10 Analyse discriminante . . . 131
7.10.1 Construction des arbres . . . 131
7.10.2 Variables utilisées . . . 132
7.10.3 Variable discriminante finale . . . 138
7.11 Conclusion . . . 138

8 Limites sur la section efficace de production du boson de Higgs 141
8.1 La statistique de test . . . 141
8.1.1 La méthode CL_s . . . 141
8.1.2 Le niveau de confiance CL_s . . . 144
8.2 Les erreurs systématiques . . . 144
8.2.1 Erreurs systématiques dans l'analyse ZH . . . 144
8.2.2 Traitement des erreurs systématiques . . . 149
8.3 Calcul de la limite . . . 149
8.3.1 Implémentation de la limite . . . 149
8.3.2 Limite obtenue . . . 149
8.4 Conclusion . . . 150
Résumé . . . 155

Remerciements

Je remercie tous les membres du groupe DØ du CPPM de Marseille.
Je remercie mon directeur de thèse Elemér Nagy qui m'a soutenu pour obtenir cette thèse, et bien encadré durant ces trois années. Etant sa dernière étudiante, il a su me faire profiter de son expérience et de sa pédagogie acquises durant sa carrière. Il m'a conseillé et orienté avec pour seul objectif, je cite "il faut que tu réussisses ta thèse".
Je remercie Marie-Claude Cousinou qui m'a également soutenue durant ces trois années, et avec qui j'ai eu la chance de travailler. J'ai apprécié la méthodologie et la rigueur qu'elle emploie dans son travail.
Je remercie Eric Kajfasz, actuellement le directeur de notre laboratoire. Il m'a énormément soutenu pour obtenir une bourse de thèse. Je suis impressionné du travail qu'il a accompli dans DØ et de son parcours. J'apprécie son humilité et sa sympathie.
Je remercie Smaïn Kermiche avec qui j'ai partagé mon bureau. C'est une personne très agréable. J'ai aimé sa dynamique au travail, il est ordonné, efficace, et toujours disponible pour aider.
Je remercie Arnaud Duperrin pour ces messages d'encouragement et de soutien. C'est une personne ouverte, motivée, et qui n'hésite pas à mettre en pratique ses nombreuses idées. Sa dynamique sert de moteur au groupe.
Je remercie Steve Muanza pour ses encouragements. C'est une personne très souriante et chaleureuse. J'ai apprécié ses réponses à mes différentes questions, et nos discussions.
Je remercie David Jamin pour son soutien. C'est une personne sur qui on peut compter. J'ai apprécié nos discussions, il est rempli d'humour et est très agréable.
Je remercie également des anciens membres du groupe, Anne-Fleur Barfuss, Samuel Calvet et Aurélien Mendès pour leurs aides précieuses.

Je remercie les membres de groupe de DØ à Fermilab, je pense en particulier à mes collègues : D. Bandurin, X. Bu, K. Chan, S. Desai, S. Fu, W. Geist, M. Kirby, J. Kraus, M. Mulhearn, E. Strauss et T. Yasuda.

Je remercie les membres du jury pour le temps qu'ils m'ont accordé, et pour avoir accepté de juger mon travail. Je remercie Eric Kajfasz pour avoir accepté de présider. Je remercie particulièrement Takahiro Yasuda qui m'a encadré durant plusieurs mois, j'ai beaucoup apprécié travailler avec lui, il est infiniment patient et gentil, et il est également très efficace dans son travail. Je remercie mes deux rapporteurs : – Isabelle Wingerter-Seez pour son soutien. J'ai apprécié ces nombreuses remarques et questions qui témoignent de son engagement. Je l'ai aussi trouvé trés rigoureuse et perspicace. – Gérard Sajot pour son soutien. C'est une personne très souriante et gentille, ce qui m'a permis d'écrire ce manuscrit avec un peu moins de stress. Je remercie Bernard Ille pour s'être déplacé. Enfin je remercie Elemér Nagy pour m'avoir constitué ce jury.

Je remercie toute ma famille proche et moins proche pour leur soutien. Je remercie

ma mère Robert Jonas Euphrasie pour m'avoir soutenu durant toutes ces années. Maman radieuse, aimante, dévouée ..., les mots me manquent pour la décrire. Je remercie mon père pour son soutien.

Introduction

Le modèle standard de la physique des particules, élaboré au début des années 70, est un succès depuis de nombreuses années à l'exception du boson de Higgs qui n'a pas encore été découvert. Comme le modèle standard ne donne aucune précision sur sa masse, le boson de Higgs est l'objet de recherche de toute une génération d'expériences de physique des hautes énergies, et c'est également le sujet de recherche de cette thèse. Des théories ainsi que des expériences directs ou indirects, favorisent un Higgs de basse masse, c'est-à-dire inférieure à 135 GeV. Aux basses masses, étant donné le taux de production élevé des bruits de fond par rapport à notre signal, la recherche du Higgs au Tevatron se fait dans les canaux où il est produit en association avec un boson W ou Z.
Au cours de ma thèse, j'ai participé à la recherche du Higgs produit en association avec un boson Z se désintégrant en électrons (paire électron-positron). L'atout de ce canal, est que les électrons produisent un signal net et facile à identifier dans le calorimètre. J'ai par ailleurs effectué mes travaux de tâches de services sur la correction de la mesure de l'énergie dans le calorimètre, ainsi que sur l'identification des électrons dans le calorimètre.
Pour augmenter l'efficacité du signal, le cas particulier où un des électrons se trouve dans la région inter-cryostat du calorimètre a également été analysé séparément. Ma contribution a été plus importante dans ce dernier cas.
La recherche a été réalisée dans les données correspondant à une luminosité intégrée de 4.2 fb^{-1}, collectées entre avril 2002 et décembre 2008 avec le détecteur DØ auprès du Tevatron. Aucun excès d'évènements compatible avec la présence du boson de Higgs n'a été observé, cependant des limites supérieures sur la section efficace de production du boson de Higgs ont été établies sur une gamme de basses masses.

L'organisation de ce mémoire s'oriente suivant trois grands axes : le premier présente le cadre théorique et expérimental, ainsi que les objets utilisés lors de l'analyse ; le deuxième concerne les différents travaux de tâches de services que j'ai effectués sur le calorimètre ; et le troisième porte sur la recherche du boson de Higgs dans le canal $ZH \to e^+e^-b\bar{b}$. Ainsi, ce mémoire a la structure suivantes :

- le chapitre 1 est une introduction au modèle standard de la physique des particules et au mécanisme de Higgs ;
- le chapitre 2 est une vue globale du complexe d'accélération du Tevatron et du détecteur DØ à Fermilab ;
- le chapitre 3 est une introduction aux objets physiques utilisés dans cette thèse ;
- le chapitre 4 présente l'étude faite sur la correction de l'énergie mesurée dans le calorimètre ;
- le chapitre 5 décrit l'étude de la certification des électrons dans le calorimètre ;

- le chapitre 6 décrit l'étude de la certification des électrons dans la région inter-cryostat du calorimètre ;
- le chapitre 7 détaille l'analyse sur la recherche du Higgs dans le canal de production $ZH \rightarrow e^+e^-b\bar{b}$;
- le chapitre 8 présente les résultats finaux des calculs des limites supérieures sur la section efficace de production du boson de Higgs sur une gamme de basses masses.

Chapitre 1

Cadre théorique

Contents

1.1 Le Modèle Standard 9
1.1.1 Les quarks et les leptons 9
1.1.2 Les interactions fondamentales 11
1.2 Théorie électrofaible et mécanisme de Higgs 12
1.2.1 Théorie électrofaible 12
1.2.2 Mécanisme de Higgs 12
1.3 La recherche du Higgs au Tevatron 17
1.3.1 Contraintes théoriques 17
1.3.2 Contraintes expérimentales directes 17
1.3.3 Contraintes expérimentales indirectes 18
1.3.4 Production et désintégration du Higgs au Tevatron 20
1.3.5 La recherche du Higgs dans le canal $ZH \to e^+e^-b\bar{b}$ 20
1.4 Conclusion 20

Ce chapitre pose le cadre théorique du sujet de thèse étudié. D'abord le Modèle Standard de la physique des particules est décrit, suivit d'une introduction à la théorie électrofaible et au mécanisme de Higgs, enfin, la phénoménologie du Higgs au Tevatron est abordée dans la dernière partie.

1.1 Le Modèle Standard

Le Modèle Standard (noté SM, pour Standard Model en anglais) de la physique des particules est une théorie qui décrit les constituants élémentaires de la matière et leurs interactions. C'est une théorie quantique relativiste basée sur le groupe de symétrie $SU(3)_{couleur} \times SU(2)_L \times U(1)_Y$. Sa conception utilise le formalisme lagrangien et les théories de jauge [1] [2].

1.1.1 Les quarks et les leptons

Les particules élémentaires de matière du SM sont supposées ponctuelles, elles ont un spin demi-entier, ce sont donc des fermions et par conséquent elles respectent

	Première famille (masse)	Deuxième famille (masse)	Troisième famille (masse)	Charge électrique	Nombre leptonique	Nombre baryonique
Quarks	u (1.5-3.3 MeV)	c (1.27 GeV)	t (171.2 GeV)	2/3	0	1/3
	d (3-6 MeV)	s (70-130 MeV)	b (4.2-4.7 GeV)	$-1/3$	0	1/3
Leptons	e^- (0.5 MeV)	μ^- (105.7 MeV)	τ^- (1.777 GeV)	-1	1	0
	ν_e ($<$ 2 eV)	ν_μ ($<$ 0.19 MeV)	ν_τ ($<$ 18.2 MeV)	0	1	0

TAB. 1.1 – Les fermions du SM et leurs propriétés. Les charges et les nombres leptoniques et baryoniques sont opposés pour les anti-particules.

le principe d'exclusion de Pauli qui dit que deux fermions ne peuvent occuper le même état quantique.

Les fermions sont également répartis en deux groupes, les leptons et les quarks, sensibles à différentes interactions, suivant trois générations (ou familles) qui ne diffèrent que par l'échelle de masse, voir tableau 1.1. Les quarks u, d, et les leptons électron (e^-) et neutrino électronique (ν_e), appartiennent à la première génération. Les quarks c, s, et les leptons muon (μ^-) et neutrino muonique (ν_μ), font partie de la deuxième génération. Enfin la troisième génération comporte les quarks t, b, et les leptons tau (τ^-) et neutrino tau (ν_τ). Les particules de première génération sont les plus légères et elles sont stables, de ce fait elles forment la matière ordinaire. Les particules de deuxième et troisième génération sont plus lourdes et instables, elles se désintègrent donc rapidement en particules de première génération, les plus légères. Le neutrino associé au lepton de chaque génération, interagit très peu avec la matière. Il a été montré que les neutrinos oscillent, c'est-à-dire qu'ils peuvent passer d'une génération à une autre, et possèdent donc, bien que très faible, une masse qui n'a pas encore été mesurée précisément [3].

Pour caractériser l'état quantique des différentes particules, on leurs attribue des nombres en fonction de leur propriété. Ainsi les leptons possèdent un nombre leptonique (électronique, muonique et tauonique) qui permet de distinguer les générations, et un nombre leptonique global égale à 1. Les nombres leptoniques sont conservés lors des interactions à courant neutre (photon, gluon, Z^0). Les quarks possèdent un nombre baryonique égale à 1/3. Ce dernier et est toujours conservé lors des interactions entre particules.
Les particules soumises à l'interaction électromagnétique porte une charge électromagnétique, cette dernière fixée à -1 pour les leptons e^-, μ^- et τ^-, vaut $+2/3$ pour les quarks u, c et t, et -1/3 pour les quarks d, s et b. Les quarks, également soumis à l'interaction forte, possèdent un nombre supplémentaire, la couleur, qui peut prendre par convention les trois valeurs rouge bleu ou vert.

Enfin, toutes les particules du MS ont une anti-particule, de même masse, spin et de durée de vie, mais avec des charges et des nombres leptoniques et baryoniques opposés.

1.1.2 Les interactions fondamentales

Une interaction entre deux particules de spin quelconque est interprétée comme l'échange d'une particule de spin entier, appelée boson. Ce dernier obéit à la statistique de Bose-Einstein, qui stipule que contrairement aux fermions, les bosons ne sont pas soumis au principe d'exclusion de Pauli et peuvent donc se regrouper dans le même état quantique. Il existe quatre interactions fondamentales listées ci-dessous :

- l'interaction gravitationnelle à l'origine de l'attraction réciproque des corps massifs entre eux, sous l'effet de leur masse. Elle serait véhiculée par le graviton. Il n'y a aucune preuve expérimentale de l'existence de ce dernier, mais la présence des ondes gravitationnelles serait une preuve indirect [4].

- l'interaction faible qui agit sur les particules et qui est responsable de la désintégration β. Elle est véhiculée par les bosons $W^{\pm}$ et Z^0.

- l'interaction électromagnétique qui agit sur les particules possédants une charge électrique, c'est elle qui permet le maintient les électrons autour du noyau des atomes. Son boson associé est le photon (γ).

- l'interaction forte qui agit sur les particules qui possèdent une charge de couleur, autrement dit les quarks et les gluons.
  La chromodynamique quantique (ou QCD, de l'anglais Quantum Chromo-Dynamics) décrit l'interaction forte. Elle postule l'existence de trois nombres quantiques appelés couleur (rouge, bleu et vert), mais qui ne sont pas observables expérimentalement. La symétrie associé à la rotation de ces charges est $SU(3)_{couleur}$. Or ce groupe de symétrie possède huit générateurs, l'interaction forte est donc véhiculée par huit bosons appelés gluons (g) et supposés de masse nulle, chacun portant une charge de couleur et une charge d'anti-couleur, donnant globalement un état sans couleur.
  L'interaction forte présente les deux grandes caractéristiques suivantes : la liberté asymptotique, c'est à dire que son intensité décroît quand la distance entre quarks diminue ; et le confinement des partons[1], en effet ces derniers ne sont pas observable à l'état libre, le parton isolé va donc se recombiner avec d'autres quarks pour former un hadron, un composite neutre de couleur. On dit qu'il s'hadronise.

Le SM décrit les interactions faibles, électromagnétique et forte mais pas l'interaction gravitationnelle. Le tableau 1.2 résume ces interactions et les caractéristiques de leur(s) boson(s) médiateur(s).

[1]Les partons sont les quarks et les gluons.

Interaction	Groupe de symétrie	Boson(s) associé(s)	Masse GeV	Charge électro-magnétique
Gravitationnelle		graviton		
Faible	SU(2)	$W^{\pm}, Z^0$	80.4, 91.2	± 1, 0
Electro-magnétique	U(1)	γ	0	0
Forte	SU(3)	g_i (i=1 à 8)	0	0

TAB. 1.2 – Les interactions fondamentales décrites par le SM, le groupe de symétrie auquel elles appartiennent, leur(s) boson(s) associé(s) avec leur masse et leur charge électromagnétique.

1.2 Théorie électrofaible et mécanisme de Higgs

1.2.1 Théorie électrofaible

La Théorie électrofaible [5] [6] unifie les interactions électromagnétique et faible. Elle est fondée sur les théories de jauge locales. Celles-ci décrivent l'origine des interactions fondamentales comme conséquence de l'invariance des phénomènes physiques dans des transformations de jauge locales. La Théorie électrofaible est basée sur le groupe de symétrie $SU(2)_L \times U(1)_Y$. Les transformations sous $SU(2)_L$ et $U(1)_Y$ sont respectivement décrites par l'isospin faible et l'hypercharge. L'invariance des phénomènes physiques suivant les transformations de ces deux groupes implique l'apparition de quatre bosons de jauge, mais la théorie électrofaible prédit une masse nulle pour ces bosons, ce qui est en contradiction avec les observations expérimentales. Effectivement, la masse du boson Z (resp. $W^{\pm}$) est proche de 91 GeV (resp. 80 GeV). Le mécanisme de Higgs, décrit dans la prochaine section, permet de lever cette contradiction.

1.2.2 Mécanisme de Higgs

Le mécanisme de Higgs [7] [8] [9] [10] [11], proposé dans les années 60 par Higgs, Brout, Englert, Guralnik, Hagen et Kibble a été incorporé à la théorie basée sur la symétrie du groupe $SU(2)_L \times U(1)_Y$ pour former la théorie électrofaible (voir section 1.2.1). Il permet, en ajoutant un champ scalaire, de provoquer la brisure de symétrie pour les basses énergies et ainsi de générer une masse pour les bosons $W^{\pm}$ et Z^0, tout en conservant l'invariance de jauge du lagrangien. Ce mécanisme explique également l'origine de la masse des fermions comme une interaction avec ce nouveau champ. Le boson de Higgs, est une conséquence direct du mécanisme de Higgs.

1.2.2.1 Le champ de Higgs

Introduction d'un champ scalaire

Le mécanisme de Higgs consiste à introduire un champ scalaire, le champ de Higgs, qui interagit avec les particules qui ont des interactions électrofaibles. Ainsi le champ de Higgs doit avoir un isospin faible et une hypercharge non nulle. Le choix porte sur un doublet de champs scalaires et complexes, appartenant à $SU(2)_L$. L'un doit être chargé électriquement et l'autre doit être neutre. Il a donc la forme suivante,

avec 4 degrés de libertés :

$$\Phi \;=\; \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} \Phi_1 \;-\; i\Phi_2 \\ \Phi_3 \;-\; i\Phi_4 \end{pmatrix}. \tag{1.1}$$

Le lagrangien du champ scalaire a la forme suivante :

$$\mathcal{L}_\Phi \;=\; |\mathcal{D}_\mu \Phi|^2 \;-\; V(\Phi). \tag{1.2}$$

– Le premier terme du lagrangien est le terme cinétique, il comprend les termes de masse des bosons de jauge ainsi que des termes d'interactions entre les champs de jauge et le champ scalaire. $\mathcal{D}_\mu$ est la dérivée covariante, introduit afin de respecter l'invariance de jauge. Elle est définie par :

$$\mathcal{D}_\mu \;=\; \partial_\mu \;+\; ig\frac{\tau^a}{2}W_{\mu a} \;+\; ig'\frac{Y}{2}B_\mu, \quad (a = 1, 2, 3) \tag{1.3}$$

g et g' sont respectivement les constantes de couplage des groupes $SU(2)_L$ et $U(1)_Y$, τ^a sont les matrices de Pauli, Y est l'hypercharge faible, $W_{\mu a}$ sont les trois champs du groupe $SU(2)_L$ et B_μ est le champ du groupe $U(1)_Y$.

– Le deuxième terme du lagrangien est le potentiel scalaire $V(\Phi)$, il implique, comme on le verra plus loin, une masse pour le boson de Higgs et des interactions du champ de Higgs avec lui-même. On choisit le potentiel renormalisable le plus simple qui brise la symétrie :

$$V(\Phi) \;=\; \mu^2|\Phi|^2 \;+\; \lambda(|\Phi|^2)^2, \tag{1.4}$$

μ est homogène à une masse, et λ est un paramètre d'auto-couplage sans dimension et il est supérieur à 0 pour que $V(\Phi)$ soit borné inférieurement.

On voit donc que le lagrangien du champ scalaire est invariant sous transformation de jauge locale de $SU(2)_L \times U(1)_Y$.

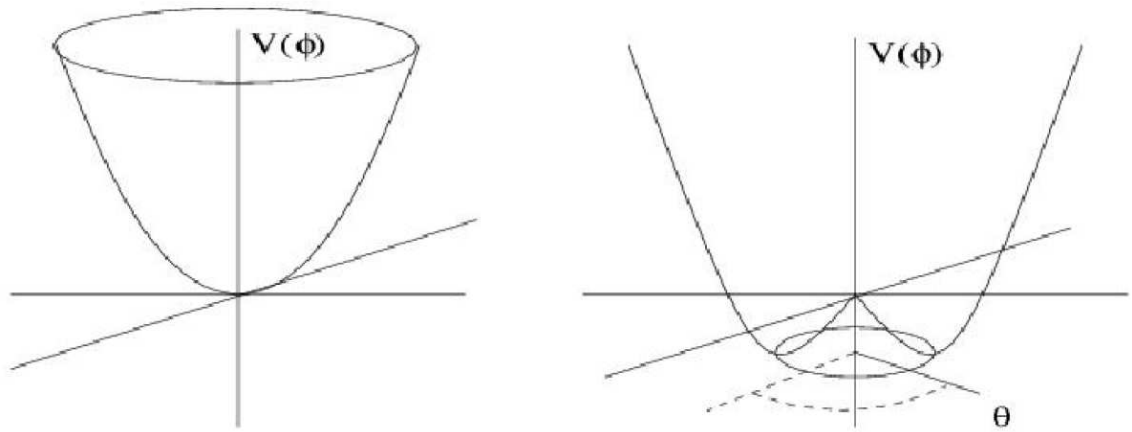
Valeur de l'état de vide

L'état fondamental d'un champ, qui correspond à l'état de vide, est défini comme étant la valeur du champ, Φ_{VEV}, qui correspond au minimum du potentiel. Le potentiel de l'Eq. (1.4) peut prendre deux formes différentes suivant le signe de μ^2 :

– si $\mu^2 > 0$, le potentiel n'a qu'un minimum à $\Phi_{VEV} = 0$, (voir figure 1.1 a), on dit alors qu'il n'y a qu'un VEV (vacuum expectation value) et on voit que le potentiel est symétrique. Ce cas ne permet pas de briser la symétrie ;

– si $\mu^2 < 0$, le VEV a une valeur non nulle et le potentiel prend la forme d'un "chapeau mexicain", (voir figure 1.1 b). Il y a alors un continuum de valeurs possibles pour Φ_{VEV} et le choix d'un minimum, arbitraire, brise spontanément la symétrie[2].

[2] On parle de brisure de symétrie spontanée quand les équations du mouvement qui gouvernent l'évolution d'un système sont invariantes sous une transformation, mais qu'une solution particulière ne l'est pas.



(a) $\mu^2 \geq 0$ (b) $\mu^2 < 0$

FIG. 1.1 – Forme du potentiel du champ de Higgs : (a) $\mu^2 \geq 0$; (b) $\mu^2 < 0$.

On prend en général la solution particulière suivante, qui n'est pas symétrique dans $SU(2)_L \times U(1)_Y$:

$$\Phi_{VEV} = \begin{pmatrix} 0 \\ \frac{\nu}{\sqrt{2}} \end{pmatrix}, \tag{1.5}$$

ν est la VEV, elle s'exprime par $\nu = \sqrt{\frac{-\mu^2}{\lambda}}$, et vaut 246 GeV.

On peut illustrer le phénomène de "brisure de symétrie spontanée" avec le cas du ferromagnétisme. Pour des températures assez élevées, les dipôles magnétiques ne sont pas alignés, ils s'orientent dans des directions aléatoires, on dit que le système est invariant sous rotation. Par contre, en dessous d'une certaine température, les dipôles s'alignent selon une certaine direction, ce qui brise spontanément la symétrie de rotation. De la même façon, notre mécanisme implique une brisure de symétrie spontanée pour les énergies faibles.

Génération de la masse des bosons de jauge

En développant Φ_{VEV} autour du minimum du potentiel, on a :

$$\Phi_{VEV} \simeq \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H \end{pmatrix}, \tag{1.6}$$

où H est le champ du boson de Higgs. En remplaçant Φ_{VEV} par sa valeur dans l'Eq. 1.2, des termes de masses pour les bosons de jauge apparaissent dans le lagrangien :

$$|\mathcal{D}_\mu \Phi|^2 = \frac{g^2\nu^2}{4} W^+W^- + \frac{g^2 + g'^2\nu^2}{8} \left(\frac{gW_3 - g'B}{\sqrt{g^2 + g'^2}} \right)^2, \tag{1.7}$$

W^+ et W^- représentent les deux premières coordonnées du champ W_μ, et W_3 la troisième. Ainsi on obtient l'expression de la masse des bosons $W^\pm$ et Z^0 :

$$M_{W^\pm} = \frac{g\nu}{2}, \quad M_{Z^0} = \frac{\sqrt{g^2+g'^2}}{2}\nu, \quad M_\gamma = 0 \tag{1.8}$$

On voit donc que le photon reste sans masse. Si on définit θ_W, l'angle de mélange des champs W et B (ou angle de Weinberg) par $\cos\theta_W = \frac{g}{\sqrt{g^2+g'2}}$, alors on voit que :

$$\frac{M_{W^\pm}}{M_{Z^0}} = cos\theta_W. \tag{1.9}$$

Génération de la masse des fermions

De même que pour la génération de masse des bosons de jauge, on génère la masse des fermions grâce au potentiel scalaire de Higgs. Pour cela on pose un lagrangien invariant (il ne doit donc pas contenir de terme de masse explicite pour les fermions), puis on exige son invariance sous $SU(2)_L \times U(1)_Y$, ce qui donne naissance à de nouveaux termes d'interaction entre le champ scalaire et les champs de matière. Ces interactions ont des constantes de couplage proportionnelles à la masse des fermions associés. Par ce procédé, la masse des fermions s'exprime comme suit :

$$M_{fermion} = \frac{\lambda_f \nu}{\sqrt{2}}, \tag{1.10}$$

où λ_f représente le couplage du fermion au boson de Higgs.

Brisure de la symétrie électrofaible

Le mécanisme de Higgs permet donc de briser spontanément la symétrie de jauge locale du groupe $SU(2)_L \times U(1)_Y$ pour donner des masses aux bosons $W^\pm$ et Z^0 et aux fermions. Comme l'ajout de ce champ de Higgs laisse le champ associé au photon sans masse, la symétrie du groupe $U(1)_{em}$ est conservée. En effet, la charge électrique annihile l'état fondamental, et il n'y a donc pas de brisure de symétrie. Le mécanisme de Higgs permet ainsi de faire passer la symétrie de jauge locale de $SU(2)_L \times U(1)_Y$, à une symétrie de $U(1)_{em}$.

Cette brisure spontanée de la symétrie n'existe que pour des basses énergies. A ces énergies, l'interaction faible est bien décrite par la théorie de Fermi et elle apparaît totalement différente de l'interaction électromagnétique. Cependant, à très haute énergie (de l'ordre de 100 GeV), la "bosse" dans la forme du potentiel est négligeable et la symétrie de jauge est restaurée. A ces échelles d'énergies, les forces électromagnétiques et faibles sont deux aspects de la même force, la force électrofaible, voir figure 1.2.

1.2.2.2 Le boson de Higgs

le boson de Higgs est une conséquence du mécanisme de Higgs. Observons les degrés de liberté avant et après la brisure de symétrie. Au départ, les trois champs

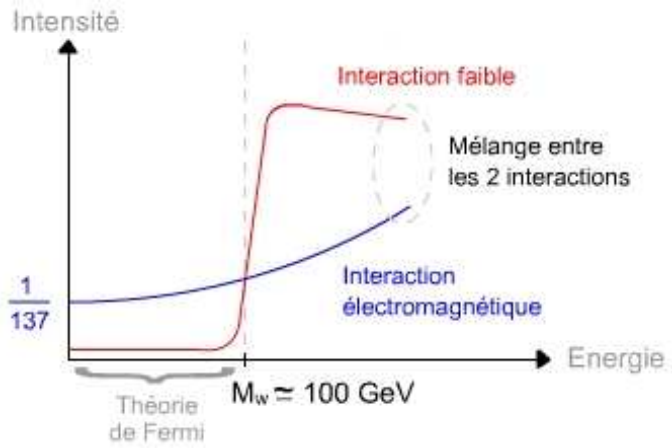


FIG. 1.2 – Représentation de l'unification des forces électromagnétique et faible.

de jauge sans masse de $SU(2)_L$ ont tous 2 degrés de liberté, le champ non massif de $U(1)_Y$ en à 2, et le champ scalaire complexe introduit en possède 4, c'est un total de 12 degrés de liberté. Après la brisure spontanée de symétrie, il reste les 3 bosons $W^{\pm}$ et Z^0, qui ont 3 degrés de liberté chacun. En effet, les 3 bosons ont en quelque sorte "pris" trois des degrés de liberté du champ scalaire dans le processus qui leur a donné une masse. Le photon non massique conserve ses 2 degrés de liberté. Cela donne donc 11 degrés de liberté. Il en reste un douzième, qu'on note H ; c'est ce degré de liberté qui correspond au boson de Higgs, une particule physique massive. Notons enfin que le champ de Higgs est un champ scalaire, ce qui implique que le Higgs est un boson de spin nul.
En remplaçant 1.6 dans 1.2, le lagrangien associé au champ devient :

$$\mathcal{L}_\Phi = \left| \left(\partial_\mu + ig\frac{\tau^a}{2}W_{\mu a} + ig'\frac{Y}{2}B_\mu \right) \frac{(\nu + H)}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|^2 - \mu^2 \frac{(\nu + H)^2}{2} - \lambda \frac{(\nu + H)^4}{4}$$

La première partie de ce lagrangien donne les termes de masse aux champs de jauge ; la deuxième partie contient des termes qui impliquent seulement le champ scalaire H. En développant ce dernier, on obtient les termes suivants :

$$- \frac{1}{2}(-2\mu^2)H^2 + \frac{\mu^2}{\upsilon}H^3 + \frac{\mu^2}{4\upsilon^2}H^4 - \frac{1}{4}\mu^2\upsilon^2, \tag{1.11}$$

les deuxième et troisième termes sont associés à des interactions du champ de Higgs avec lui-même. Le premier terme est un terme de masse pour le Higgs avec $M_H = \sqrt{-2\mu^2}$. Comme vu précédemment, μ^2 doit être négatif pour que la brisure de la symétrie ait lieu, mais sa valeur reste indéterminé.
La masse du Higgs demeure donc inconnue, cependant des recherches théoriques et expérimentaux permettent de définir des gammes de masses accessibles au bosons de Higgs pour guider les recherches.

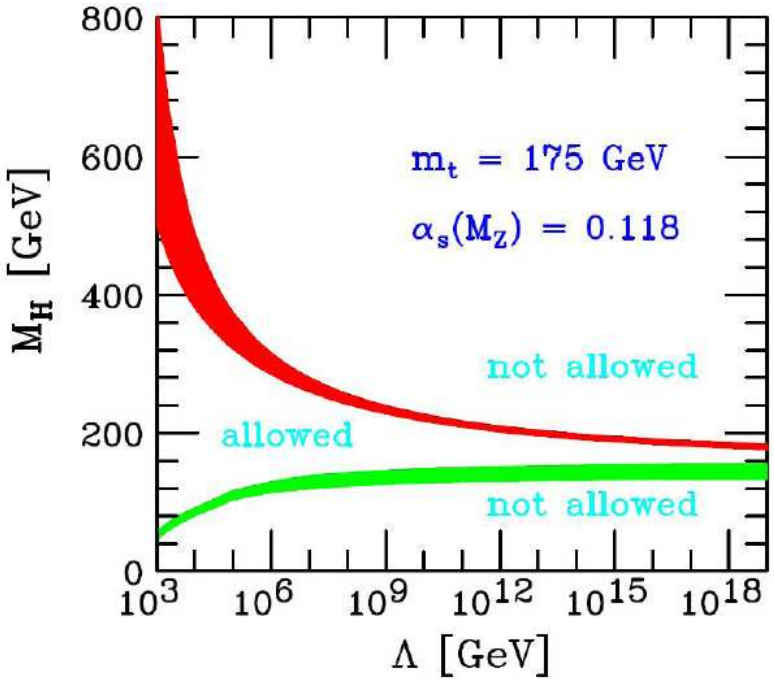


FIG. 1.3 – Contraintes théoriques établies en fonction de l'échelle d'énergie [12] [13] [14]. La bande supérieure rouge (resp. inférieure verte) représente la frontière fixée par le régime perturbatif (resp. la stabilité du vide) ; leur largeur dépendent des incertitudes. La gamme de masse accessible au boson de Higgs se trouve entre les deux bandes.

1.3 La recherche du Higgs au Tevatron

Le SM ne donne aucune indication sur la masse du boson de Higgs, mais des recherches théoriques ainsi que les recherches expérimentales directes et indirectes ont permis de définir des gammes de masses accessibles au bosons de Higgs. Les recherches expérimentales ont d'abord été réalisées au LEP un collisionneur e^+e^- avec une énergie au centre de masse de 200 GeV, puis au Tevatron un collisionneur $p\bar{p}$ avec une énergie au centre de masse de 1.98 TeV.

1.3.1 Contraintes théoriques

Des calcules détaillés basés sur l'équation du groupe de renormalisation pour l'auto-couplage du Higgs [12] [13] [14], montrent qu'un boson de Higgs avec une masse trop grande rend le modèle non-perturbatif (les calculs divergent), et un Higgs de masse trop faible déstabilise l'état fondamental, voir fig. 1.3.

1.3.2 Contraintes expérimentales directes

Les limites les plus contraignantes ont été établies par les expériences du LEP [15] (Large Electron-Positron collider), collisionneur e^+e^- du CERN qui a fonctionné

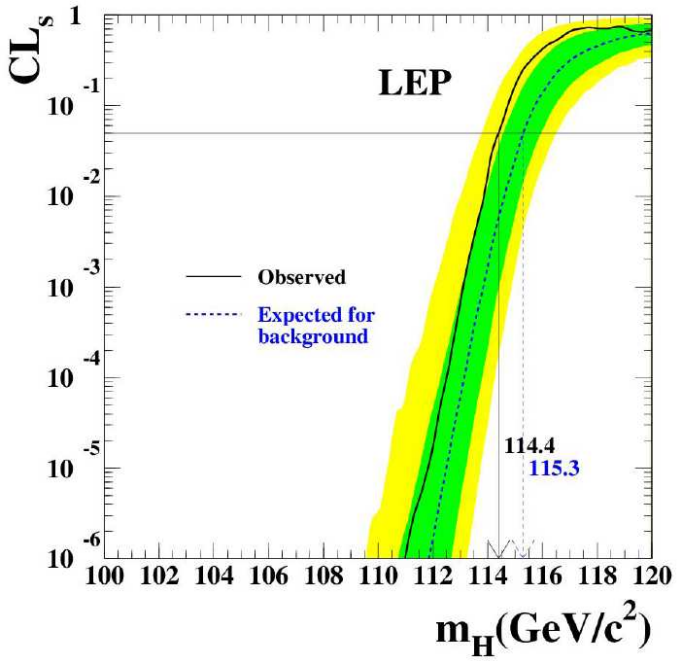


FIG. 1.4 – Niveau de confiance de la présence d'un boson de Higgs en fonction de sa masse, pour les expériences du LEP. La ligne pleine noire (resp. pointillée bleue) représente la limite observée (resp. attendue). Leur intersection avec la ligne horizontale noire indique le seuil de 95% CL, attendue pour une masse de 114,4 GeV observée et 115,3 GeV pour une masse attendue. Les bandes intérieure verte et extérieure jaune représentent respectivement un écart d'1 et 2 σ autour de la valeur centrale attendue.

de 1989 à 2000, et cumulé une luminosité intégrée de 2.5 pb^{-1}. Les données recueillis ont permit d'exclure un Higgs de masse inférieure à 114.4 GeV à 95% CL, et la limite attendue était de 115.3 GeV [16], voir figure 1.4.

En 2009 la combinaison des expériences DØ et CDF au Tevatron, incluant 28 analyses différentes et une luminosité intégrée variant entre 2.1 et 5.4 fb^{-1} selon le canal considéré, à permit d'exclure un Higgs de masse comprise entre 163 GeV $<$ m$_H$ $<$ 166 GeV à 95% CL, les limites attendues sont comprises entre 159 GeV $<$ M$_H$ $<$ 168 GeV, voir fig. 1.5 [17] [18].

1.3.3 Contraintes expérimentales indirectes

Les recherches directes peuvent être complétées efficacement avec les recherches indirects. Ces dernières basées utilisent des paramètres électro-faibles reliés à la masse du Higgs, comme la masse du boson W ou celle du quark top. Les mesures de précisions électro-faibles resserrent de plus en plus les contraintes sur la masse du Higgs, et permettent donc de guider les recherches directes, voir figure 1.6.

La combinaison de ces mesures avec celles d'autres paramètres électro-faibles comme

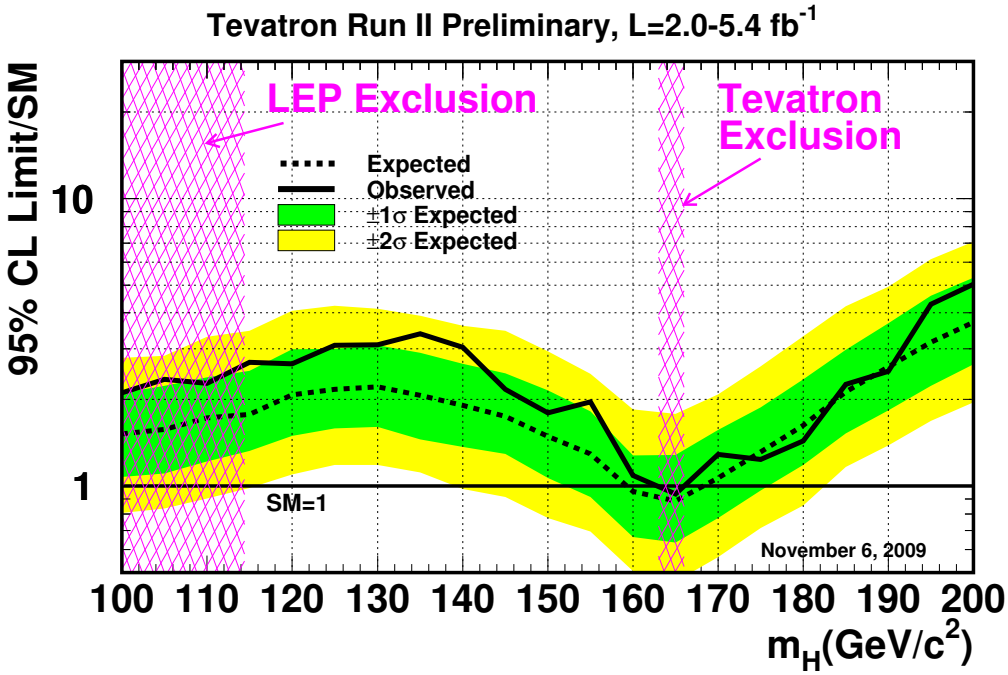


FIG. 1.5 – Limites sur la section efficace de production du boson de Higgs obtenues par la combinaison des expériences du Tevatron, un collisionneur $p\bar{p}$ avec une énergie au centre de masse de 1.98 TeV. La ligne trait plein (pointillée) noire, représente la limite observée (attendue) en unité de la section efficace du SM. Les bandes verte et jaune représentent un écart de 1 et 2 σ autour de la valeur nominale de la limite attendue. La partie des lignes en trait plein (pointillé) noire sous la valeur 1, c'est-à-dire entre 163 GeV et 166 GeV (159 GeV et 168 GeV) détermine la région de masse du Higgs exclue observée (attendue) à 95% CL.

les mesures de la masse et de la largeur du boson Z^0 par exemple permettent d'augmenter les contraintes. Pour cela on ajuste des valeurs mesurées et théoriques dépendant de la masse du Higgs, et on cherche le meilleur ajustement, c'est-à-dire celui qui donne le plus petit χ^2, voir fig. 1.7. Le minimum de cette fonction (la masse la plus probable du boson de Higgs) se trouve à 76^{+33}_{-24} GeV à 68% CL. De plus la limite supérieure à 95% de CL calculée pour $\chi^2 = 2.7$ prevoit un Higgs de masse $m_H \lesssim$ 190 GeV, le Higgs léger est donc favorisé. Notons que ces limite ont été établies dans le cadre stricte du SM. Cette gamme de masses relativement basses (100 GeV - 190 GeV) est accessible au Tevatron, comme vu précédemment, et on comprend donc l'importance de la recherche du Higgs léger dans ces experiences.

1.3.4 Production et désintégration du Higgs au Tevatron

La production du Higgs à basse masse (c'est-à-dire $M_H < 135$ GeV) au Tevatron résulte de deux principaux mécanismes : la production direct par fusion de gluons, et la production associée à un boson W ou Z (ou Higgsstrahlung), voir figure 1.8. La production direct par fusion de gluons domine, mais ce canal est difficile à maîtriser à basse masse car le bruit de fond QCD est trop important. A basse masse, la recherche du Higgs au Tevatron se fait donc essentiellement dans les canaux où le Higgs est produit en association avec un boson W ou Z. La figure 1.9 montre les diagrammes de Feynman des trois principaux canaux étudiés au Tevatron à basse masse. L'avantage du canal $ZH \rightarrow l^+l^-b\bar{b}$ par rapport au canaux $ZH \rightarrow \nu\nu b\bar{b}$ et $WH \rightarrow l\nu b\bar{b}$, qui contiennent un ou plusieurs neutrino(s) difficilement détectable(s), est que le boson Z se désintègre en deux leptons (électrons ou muons) qui produisent un signal clair et facile à identifier.

1.3.5 La recherche du Higgs dans le canal $ZH \rightarrow e^+e^-b\bar{b}$

Comme vu précédemment, le canal ZH est l'un des canaux le plus sensible pour la recherche du Higgs à basse masse ($M_H < 135$ GeV), et par conséquent le canal particulier $ZH \rightarrow e^+e^-b\bar{b}$, qui est le sujet de recherche de cette thèses, est donc très important pour la recherche du Higgs au Tevatron.
De nombreux processus physiques présentent la même topologie finale que notre signal, soit deux électrons et deux jets de quark b, et constituent un immense bruit de fond pour notre analyse. Ces bruits de fond sont de réels défits pour la recherche du Higgs dans les collisions hadroniques. Les bruits de fond dominant proviennent de différents processus suivant le niveau de l'analyse, (cf. 7.3).

1.4 Conclusion

Le Modèle Standard de la physique des particules qui permet de décrire les fermions ainsi que leurs interactions est un succès depuis de nombreuses années, cependant, il n'explique pas l'origine de la masse des particules. Le mécanisme de Higgs permet de résoudre ce problème en introduisant un nouveau champ qui génère la masse des bosons et des fermions. Comme à tout champ, on lui associe une particule, le boson de Higgs, qui n'a pas encore été découvert à ce jour. Le SM ne fournissant aucune indication sur sa masse, des expériences directes et indirectes permettent cependant de réduire les domaines de recherches et favorisent un Higgs de masse léger. Enfin, le canal $ZH \rightarrow e^+e^-b\bar{b}$ est très important pour la recherche du Higgs à basse masse au Tevatron.

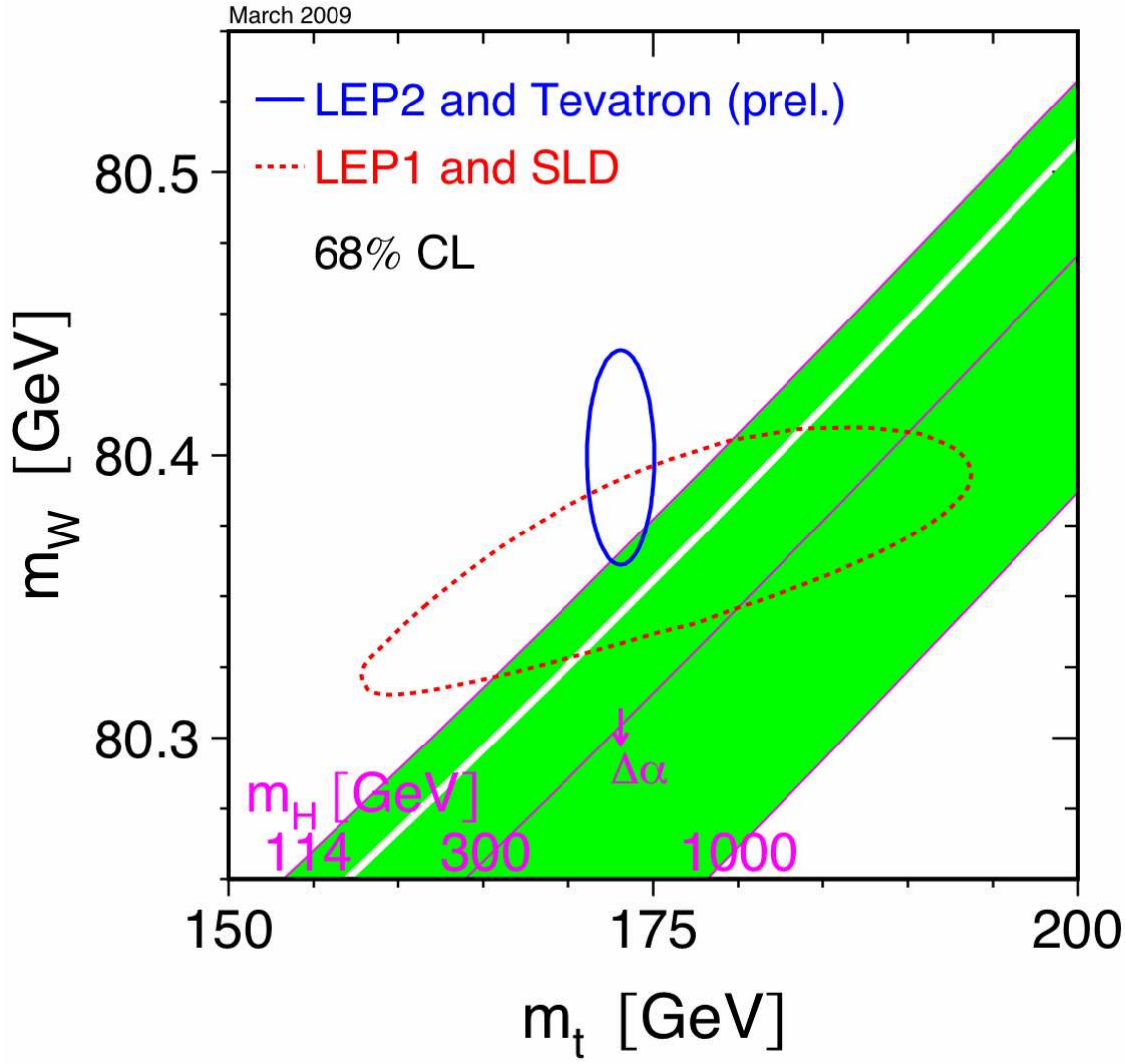


FIG. 1.6 – Masse du boson W en fonction de la masse du quark top. Le contour en pointillé est la contrainte posée par des mesures directes au LEP et au Tevatron ; le contour en trait plein représente les contraintes posées par des mesures indirectes réalisées par les collaborations LEP et SLD. Les deux contours sont établis à 68% de niveau de confiance. La bande et les chiffres surimposés indique la relation de ces deux masses à celle du boson de Higgs.

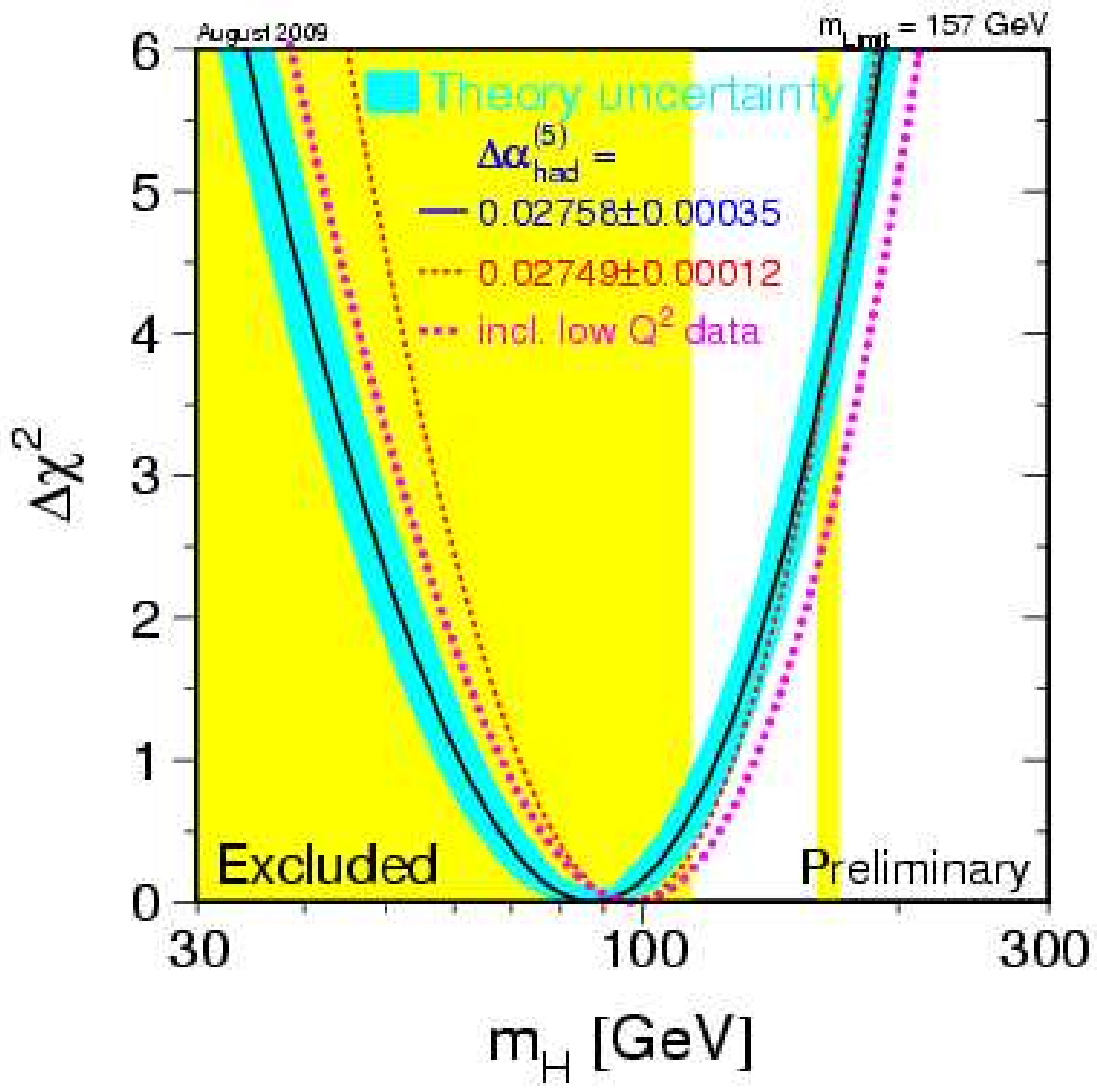
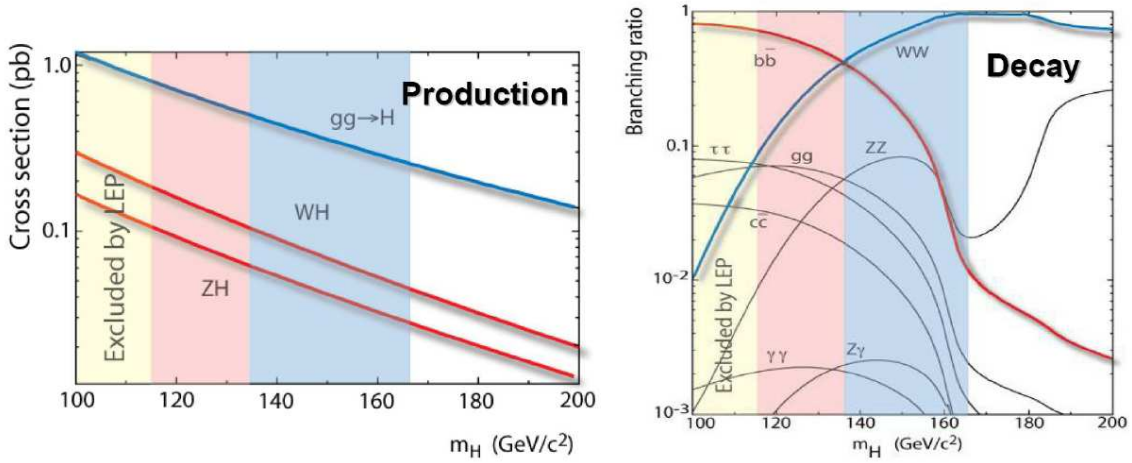


FIG. 1.7 – Courbe de $\Delta\chi^2$, représente la différence entre le χ^2 minimal et le χ^2 obtenu en faisant varier la masse du Higgs comme un paramètre. Le minimum indique la masse la plus probable, soit 76 GeV. La bande bleue qui entoure la ligne noire est l'erreur théorique. La grande bande jaune horizontal à gauche et la petite à droite (excluded) représente la zone de masse exclue à 95% CL par le LEP et le Tevatron respectivement.



(a) Modes de production

(b) Canaux de désintégration

FIG. 1.8 – Section efficace de production du Higgs au Tevatron pour différents modes de production (a), ainsi que le rapport de branchement des canaux de désintégration (b), en fonction de sa masse. La bande jaune à gauche indique la zone de masse exclue par le LEP ; la bande du milieu (rouge), et celle de droite (bleue), indiquent les zones dans lesquelles la production associée et la production directe respectivement, apportent la majorité de la sensibilité. On voit également que le mode de désintégration du Higgs en $b\bar{b}$ domine jusqu'à 135 GeV.

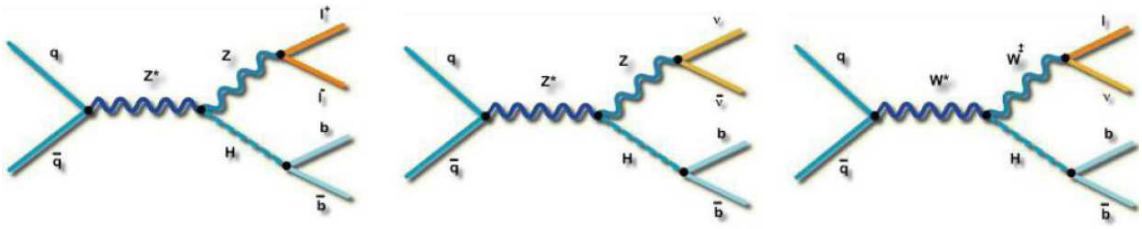


FIG. 1.9 – Diagrammes de Feynman des trois principaux canaux analysés au Tevatron à basse masse, M_H<135 GeV : à gauche $ZH \rightarrow l^+l^-b\bar{b}$, au milieu $ZH \rightarrow \nu\nu b\bar{b}$, et à droite $WH \rightarrow l\nu b\bar{b}$.

Chapitre 2

Dispositif expérimental

Contents

2.1 Le complexe d'accélération de Fermilab 24
2.1.1 Production des protons 26
2.1.2 Production des antiprotons 26
2.1.3 Le Tevatron 27
2.2 Le détecteur DØ 30
2.2.1 La zone des collisions 30
2.2.2 La luminosité 32
2.2.3 Les sous-détecteurs 34
2.2.4 Le système de déclenchement 51
2.2.5 Reconstruction des évènements et format des données . . 53
2.3 Conclusion 54

Ce chapitre est consacré au dispositif expérimental du Tevatron. Le complexe d'accélération au Run II est présenté dans la première section, le détecteur DØ avec son système de déclenchement ainsi que l'étape de la reconstructions des évènements sont décris dans la deuxième section.

2.1 Le complexe d'accélération de Fermilab

Le complexe d'accélération de Fermilab ou FNAL (Fermi National Accelerator Laboratory) [19], situé dans la banlieue Ouest de Chicago aux USA, se compose d'un ensemble d'accélérateurs linéaire et circulaire (voir fig. 2.1). Le Tevatron est le dernier maillon de la chaîne, avec une circonférence de 6.3 km, il collisionne des protons et des antiprotons au centre des détecteurs DØ et CDF, avec une énergie au centre de masse de 1.96 TeV. Le Tevatron est ainsi, après le LHC (Large Hadron Collider) du CERN [20] situé à Genève, le deuxième plus puissant collisionneur au monde.

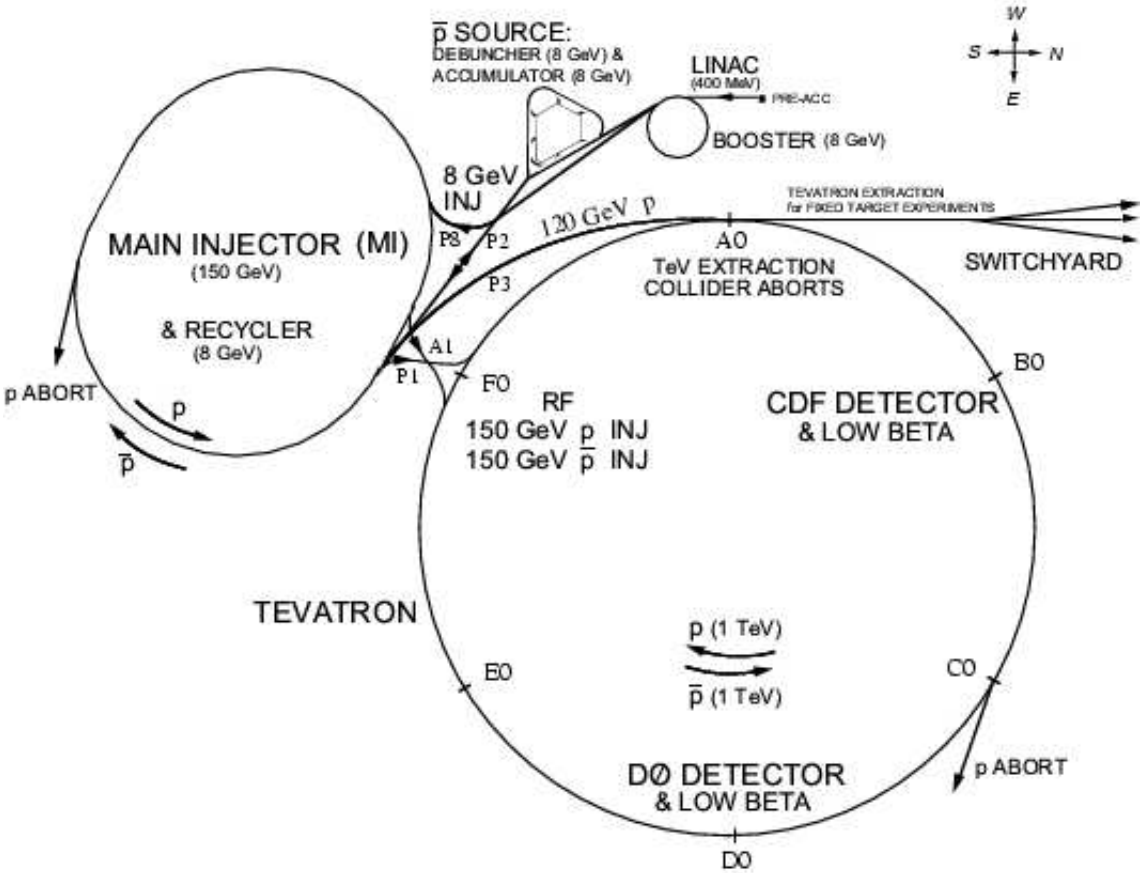


FIG. 2.1 – Les étapes successives de production et d'accélération des faisceaux au Tevatron pour le Run II. Pour les protons : le pré-accélérateur Cockroft-Walton, le Linac, le Booster puis l'injecteur principal et le Tevatron. Pour les antiprotons : la source d'antiprotons avec une station-cible, les anneaux debuncher et accumulateur, puis l'injecteur principal et le Tevatron.

2.1.1 Production des protons

La production et l'accélération d'un faisceau de protons nécessite plusieurs étapes décrites ci-dessous :

- l'accélérateur électrostatique **Cockcroft-Walton** est la première étape d'accélération du faisceau d'ions hydrogène négatifs H^-, à partir duquel les protons sont produits. D'abord des atomes d'hydrogène H sont plongés dans un champ électrique qui va les ioniser, cf. fig. 2.2. Les protons créés vont s'agglomérer sur une électrode négative en césium, d'où ils seront enlevés par les protons incidents suivants. Comme l'énergie de liaison des électrons du césium est faible, le proton repart avec deux électrons sous forme d'ion H^-. Ces derniers migrent dans une colonne de Cockcroft-Walton pour être accélérés jusqu'à une énergie de 750 KeV, à l'aide d'un système de condensateurs, cf. fig. 2.3. Enfin, les ions H^- de 750 KeV sont dirigés vers l'accélérateur linéaire, le Linac.

- le **Linac** est un accélérateur linéaire de 130 m, il reçoit le faisceau d'ions H^- issu de l'accélérateur Cockcroft-Walton, et l'accélère jusqu'à une énergie de 400 MeV avant de le diriger vers le booster.

- le **Booster** est un synchrotron d'une circonférence de 475 m. A la sortie du Linac, le faisceau d'ions H^- traverse une feuille de carbone de faible épaisseur, ce qui débarrasse les ions H^- de leurs électrons. Le faisceau de protons créé est récupéré par le Booster qui les accélère jusqu'à une énergie de 8 GeV. Le faisceau de protons de 8 GeV est ensuite stocké dans l'injecteur principal.

- l'**injecteur principal** est un synchrotron d'un diamètre de 2 km : il récupère une partie des protons provenant du Booster et les accélère jusqu'à 150 GeV ; il accélère l'autre partie des protons jusqu'à 120 GeV. Cette dernière servant à la fabrication des antiprotons, décrit ci-dessous, grâce à l'accumulateur, puis il accélère le faisceau d'antiproton jusqu'à 150 GeV. Finalement l'injecteur principal envoie les faisceaux de proton et d'antiproton de 150 GeV dans le Tevatron.

2.1.2 Production des antiprotons

La production des antiprotons se fait à partir des protons provenant de l'injecteur principal, et nécessite les étapes successives suivantes :

- **la station cible** permet de créer des antiprotons à partir du faisceau de protons initialement accéléré de 120 GeV dans l'injecteur principale. Il est donc récupéré et collisionné avec une cible de Nikel. Les particules créées sont ensuite focalisées avec une lentille au lithium, puis elles passent dans un champ magnétique qui sépare alors les antiprotons des autres particules, voir fig. 2.4. Il faut près d'un million de proton pour créer un antiproton. A ce stade, les antiprotons sont très dispersés en énergie et direction, et ils seront homogénéisés à l'aide du debuncher.

- **le debuncheur** permet d'homogénéiser le faisceau d'antiproton issu de la

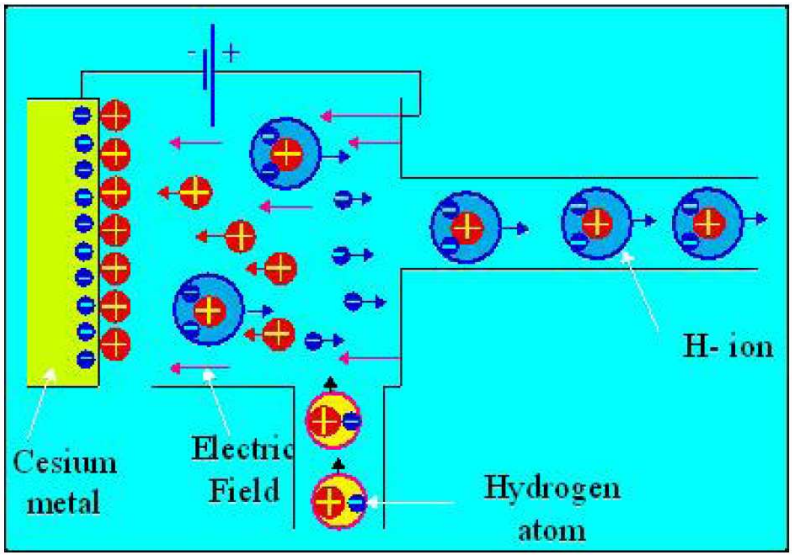


FIG. 2.2 – Schéma de la création des ions H^-.

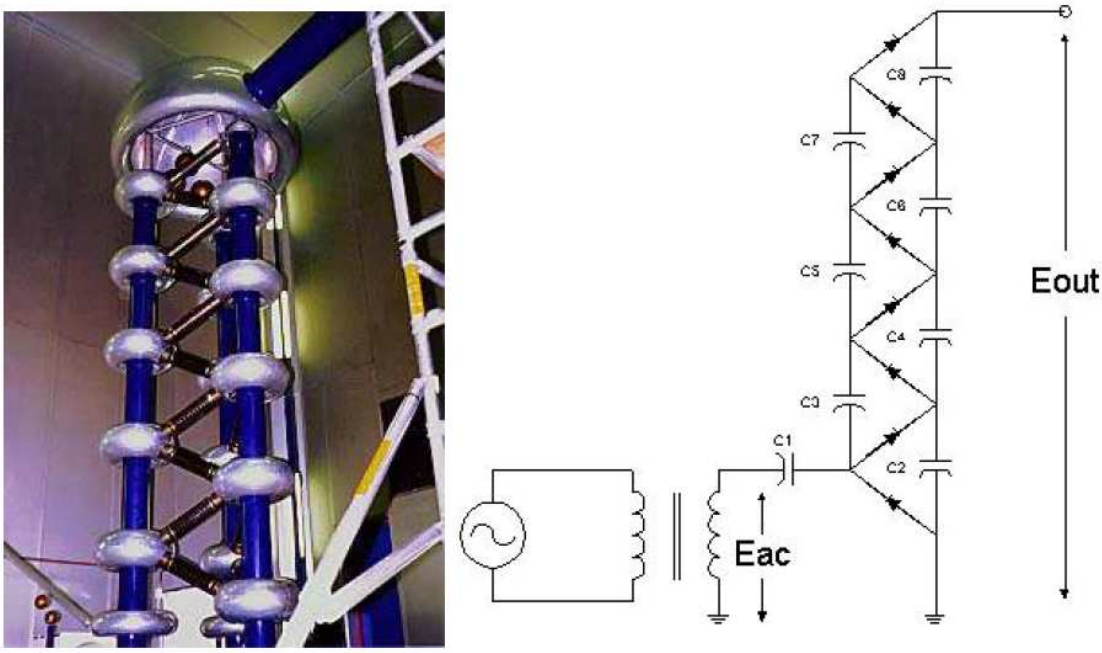
station cible. C'est un accélérateur triangulaire qui grâce à sa cavité radiofréquence et son refroidissement stochastique [21] va respectivement diminuer la dispersion en énergie, et corriger le faisceau par rétroaction, voir fig. 2.5.

- **l'accumulateur** est un anneau triangulaire se trouvant dans le même tunnel que le Debuncheur. Il récupère les faisceaux d'antiprotons provenant du Debuncheur, les stockes et les regroupe en paquets pour les envoyer dans l'injecteur principal. Il utilise également des cavités radio-fréquence et un refroidissement stochastique afin de minimiser la taille transverse des faisceaux et conserver leur énergie.

- **le recycleur** est un anneau de stockage qui se trouve dans le même tunnel que l'injecteur principal. Il permet d'optimiser le taux de formation des faisceaux d'antiprotons en recyclant les antiprotons restant en circulation dans le Tevatron en fin de période de collision. Ceux-ci sont alors décélérés dans le Tevatron jusqu'à 150 GeV, puis réinjectés dans l'injecteur principal qui les ralentit jusqu'à 8 GeV avant de les transférer dans le recycleur. Ce dernier, équipé de dispositifs de refroidissement des particules par faisceau d'électrons [22], refroidit les antiprotons et les prépare à une réinjection dans l'injecteur principal pour accélération.

2.1.3 Le Tevatron

Le Tevatron est un synchrotron d'une circonférence de 6.3 km permettant d'accélérer puis de collisionner les faisceaux de protons et d'antiprotons. Il peut, à l'aide des ses 772 dipôles et 180 quadrupôles supraconducteurs, accélérer les faisceaux en provenance de l'injecteur principal de 150 GeV à 980 GeV, pour produire des collisions avec une énergie de 1,96 TeV dans le centre de masse. Des sytèmes magnétiques de focalisation et d'écrasement des faisceaux permettent de produire les collisions



(a) Photographie du Cockcroft-Walton de Fermilab

(b) Schéma de principe d'un Cockcroft-Walton

FIG. 2.3 – Photo de l'accélérateur Cockcroft-Walton (a) : le dôme électriquement chargé contient des atomes d'hydrogène qui sont ionisés ; les cylindres verticaux sont des condensateurs. Schéma du Cockcroft-Walton (b) : une haute tension continue Eout est créée à partir d'une moyenne-tension alternative Eac par un ensemble de condensateurs c_i.

au centre des détecteurs DØ et CDF.
Chaque faisceau contient 36 paquets de particules, eux-mêmes répartis en trois super-paquets de 12 paquets chacun ; ces super-paquets sont espacés de 2.6 μs, alors que les paquets le sont de 396 ns. Les faisceaux de protons et d'antiprotons circulent en sens inverse dans la Tevatron, et collisionnent pendant quelques dizaines d'heures, après la luminosité instantanée devient trop faible (cf. 2.2.2). Cette période de prise donnée est un store.

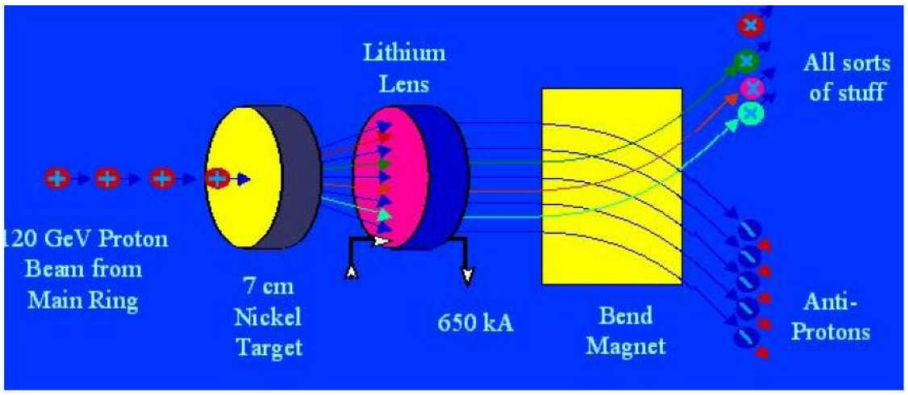


FIG. 2.4 – La station cible : la collision d'un faisceau de proton de 120 GeV sur une cible de nickel produit des antiprotons et d'autres particules, qui sont focalisés par une lentille avant de passer par un aimant qui sépare les antiprotons des autres particules.

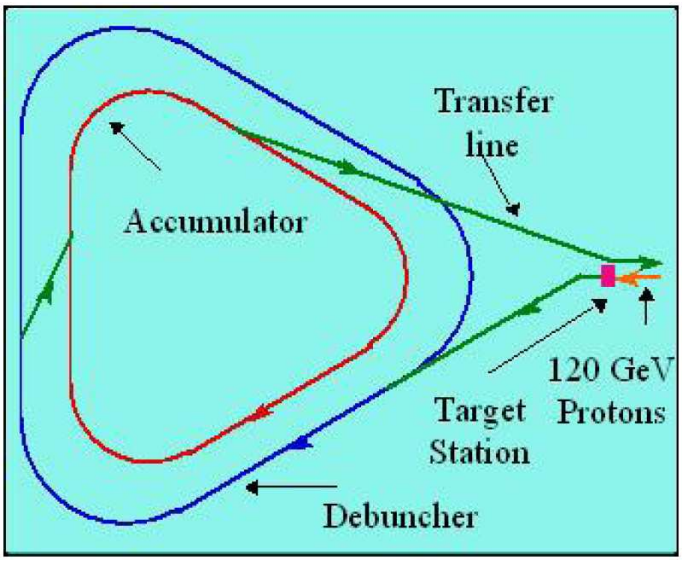


FIG. 2.5 – Le debuncher servant à la production des antiprotons. Le faisceau de proton est envoyé sur une cible de nikel, produisant ainsi des antiprotons. Ces derniers trés dispersé en énergie, son homogénéisé dans le Debuncher. Les antiprotons sont ensuite stockés dans l'accumulateur puis envoyés dans l'injecteur principal.

2.2 Le détecteur DØ

2.2.1 La zone des collisions

Le détecteur DØ est constitué de plusieurs sous-ensembles placés autour des faisceaux, voir section 2.2.3. Il a une longueur de 20 m et une hauteur de 10 m, voir fig. 2.6. On associe au détecteur DØ un repère orthonormé (x, y, z) ayant pour origine le centre du détecteur : l'axe z est orienté dans le sens des protons, l'axe y vers le haut et l'axe x vers l'extérieur du Tevatron. Le détecteur DØ couvre un angle solide maximal autour du point d'interaction et possède une bonne symétrie avant-arrière. On peut ainsi repérer les objets avec leurs coordonnées cylindriques (r, θ, ϕ), ayant pour origine le point d'interaction (ou vertex, cf. chapitre 3) : r est alors la distance à l'axe z dans le plan transverse (x, y), et θ et ϕ sont respectivement les angles polaire et azimutal. En général, l'angle polaire θ est exprimé en terme de pseudo-rapidité :

$$\eta \simeq -ln\ tan\frac{\theta}{2}. \tag{2.1}$$

On peut mesurer η à partir du centre du détecteur, ou bien à partir de la position z du vertex d'interaction primaire ; dans le premier cas on parle de "η détecteur", noté η_{det}, et dans le second on parle de "η physique", simplement noté η.

Lors d'une collision, l'impulsion du centre de masse des partons qui interagissent n'est pas nulle dans la direction z mais elle l'est approximativement dans le plan transverse. L'impulsion totale en z n'étant pas connue car l'acceptance est limitée en η, il est donc judicieux d'utiliser des variables transverses comme l'impulsion transverse $\vec{p_t}$ ou encore l'énergie transverse E_t définies par :

$$\vec{p_t} = \vec{p}\ sin\theta \quad , \quad E_t = Esin\theta. \tag{2.2}$$

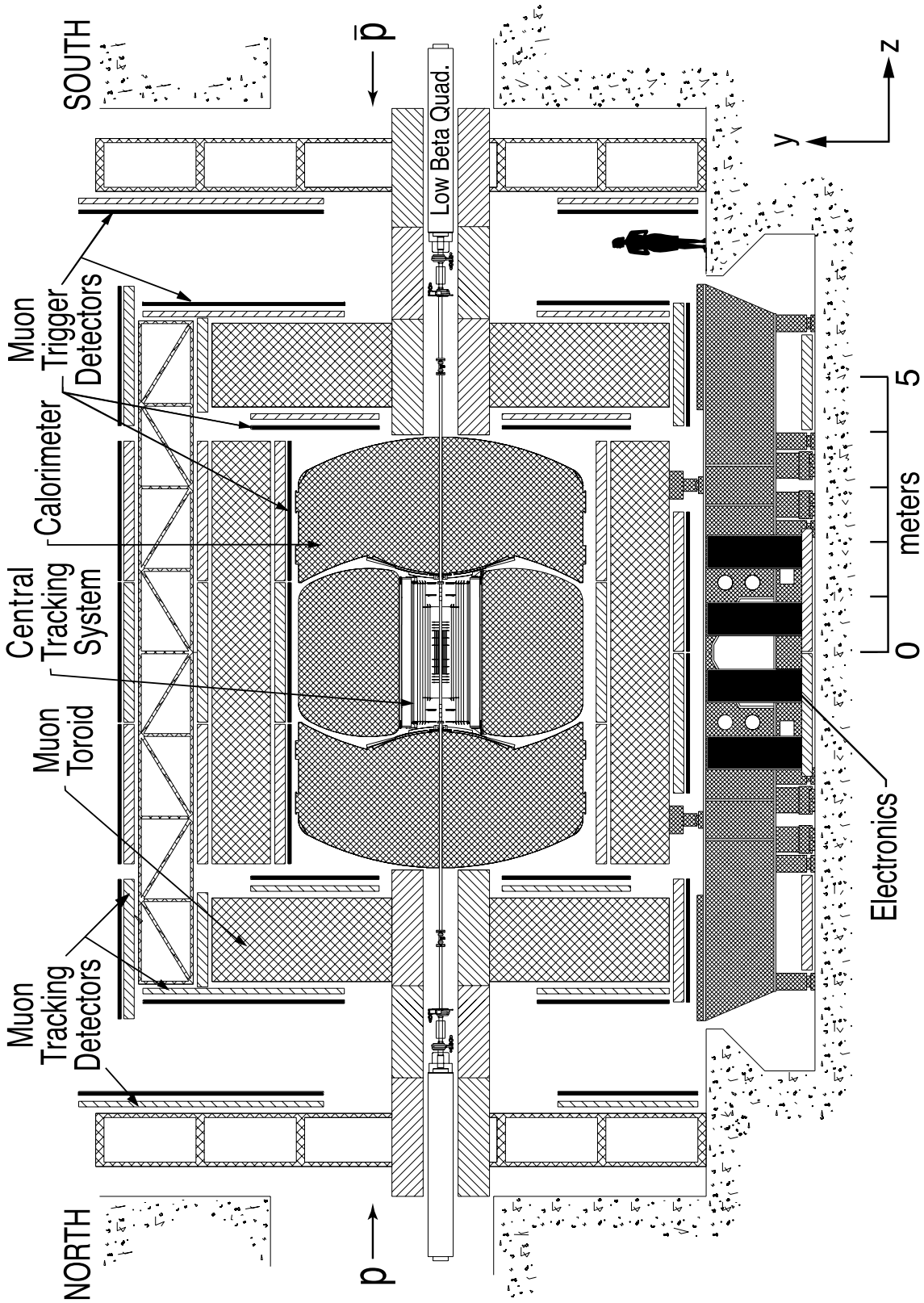


FIG. 2.6 – Vue en coupe longitudinale du detecteur DØ avec le trajectographe, le calorimètre et le système à muons.

2.2.2 La luminosité

2.2.2.1 Calcul de la luminosité

La luminosité dans les collisionneurs de particules à haute énergies est un facteur fondamental caractérisant la machine. La luminosité instantanée L est calculée à partir des caractéristiques des faisceaux par la relation :

$$L = \frac{f\ B\ N_p\ N_{\bar{p}}}{2\pi(\sigma_p^2 + \sigma_{\bar{p}}^2)}.F \tag{2.3}$$

où $f \simeq$43,7 kHz est la fréquence de rotation des paquets, $B = 36$ est le nombre de paquets par faisceau, N_p ($N_{\bar{p}}$) est le nombre de protons (d'antiprotons) par paquet, σ_p ($\sigma_{\bar{p}}$) est la dimension transverse du faisceau de protons (d'antiprotons) au point d'interaction, et F un facteur de forme qui dépend entre autres de la longueur des paquets. Les pics de luminosité intantanée typique au Run IIa (Runs IIb), cf. 2.2.2.3, sont de l'ordre de 40.10^{30} cm^{-2}s^{-1} (200.10^{30} cm^{-2}s^{-1}).

Expérimentalement, on est intéressé par la luminosité intégrée $\mathcal{L}$ qui est l'intégrale de la luminosité instantanée sur une période de temps donnée :

$$\mathcal{L} = \int_\tau L(t)\ dt. \tag{2.4}$$

En général, $\mathcal{L}$ est exprimée en inverses de sous-multiples du barn (b^{-1}) : un barn = 1b = 10^{-24}cm^2.
La luminosité intégrée calculée pour chaque grande période de prise de donnée au Tevatron est résumée dans la section 2.2.2.3.

Le nombre d'évènements observés d'un processus donné N^{proc} est relié à la section efficace du même processus σ^{proc} par :

$$N^{proc} = \epsilon\ .\ \mathcal{L}\ .\ \sigma^{proc}, \tag{2.5}$$

ϵ est un facteur correctif qui prend en compte les différentes efficacités et acceptances expérimentales.

2.2.2.2 Mesure de la luminosité

Les paramètres intervenant dans le calcul de la luminosité instantanée ne peuvent être connus à chaque croisement de faisceau, la luminosité intégrée est donc directement mesurée par des détecteurs de mesure de luminosité (Luminosity Monitor) [23], composés de deux ensembles de 24 scintillateurs, placés sur la face interne des bouchons du calorimètre, dans la région $2.7 \leq |\eta| \leq 4.4$, voir fig. 2.7. Chaque ensemble de scintillateurs comporte 24 portions de cercles identiques, possédant toutes un tube photomultiplicateur, voir fig. 2.8. Les scintillateurs mesurent la luminosité en détectant les particules issues des collisions $p\bar{p}$. En effet, le calcul de la luminosité au Tevatron est basé sur le processus de la collision $p\bar{p}$, de section efficace inélastique

d'environ 61 mb.

La mesure de la luminosité est importante pour la normalisation des données à analyser. Les évènements produits dans un store (cf. 2.1.3) sont enregistrés dans des unités logiques appelées runs. Chaque run (et chaque store) a un numéro d'identification unique. Cette segmentation permet un traitement plus simple des données, par exemple quand la qualité des données n'est pas correcte (problème dans l'un des sous-systèmes du détecteur, etc.) pendant un run, on rejette ce run et l'on garde le reste du store. De la même manière, chaque run est subdivisé en blocs de luminosité ayant chacun un identifiant unique, appelé LBN (Luminosity Block Number). Un bloc de luminosité est un laps de temps durant lequel toutes les quantités relatives à la luminosité sont mesurés. Pour correctement normaliser les données à analyser, les mauvais blocs de luminosité (pas de luminosité enregistrée, etc.) sont également rejetés.

2.2.2.3 Les périodes de prise de données

Au Tevatron, il y a plusieurs grandes périodes de prises de données appelées "Runs". Entre chaque Run (ou Shutdown), le détecteur subit des réparations ou des modifications afin d'améliorer ses performances pour le Run suivant. Les différents Runs au Tevatron sont les suivants :

- le Run I s'étend de 1992 à 1996,
- le Run IIa s'étend d'avril 2002 jusqu'à février 2006,
- le Run IIb1 a commencé en juin 2006 et s'est terminé en août 2007,
- le Run IIb2 a été collecté entre octobre 2007 et juin 2009,
- le Run IIb3 a débuté en septembre 2009.

Au cours du Run I, le Tevatron a délivré une luminosité intégrée de 140 pb^{-1} par expérience, avec une énergie au centre de masse de 1.8 TeV, permettant ainsi au quark top d'être découvert en 1995 [24] [25]. La grande période transitoire entre le Run I et le Run IIa a permit de faire d'importantes améliorations dans le complexe d'accélération de Fermilab, qui ont permis d'accroître l'énergie au centre de masse, qui est passé de 1.8 TeV pour le Run I à 1.96 TeV pour le Run IIa, suivit d'une forte augmentation de la luminosité. Le détecteur DØ a également été révisé pour croître les performances de ses différents sous-détecteurs, ainsi que son sytème de déclenchement et d'acquisition, (cf. 2.2.3 et 2.2.4). Les autres périodes de transitions, beaucoup plus courtes entre les Runs IIa et IIb1, les Runs IIb1 et IIb2, et les Runs IIb2 et IIb3 ont permis de faire des réparations, ou quelques nouvelles installations. Grâce à tous ces changements, on prévoit pour l'ensemble des Runs II une luminosité intégrée de plus de 10 fb^{-1} pour l'expérience DØ d'ici fin 2011, voir figure 2.9.

La recherche du Higgs dans cette thèse a été réalisée dans les données des Runs IIa, IIb1 et une partie du Run IIb2 enregistrées entre octobre 2007 et décembre 2008. Pour plus de clarté, dans la suite nous noterons "Run II" l'ensemble des Runs

IIa, IIb1, IIb2 et IIb3 ; nous noterons “Run IIa” le Run IIa ; et enfin nous noterons “Run IIb” l'ensemble des Runs IIb1, IIb2 et IIb3. Il sera par conséquent précisé pour chaque analyse de cette thèse, la composition des Runs étudiés.

2.2.3 Les sous-détecteurs

Les enjeux physiques du Run II sont de faire des mesures plus précises, et de mettre en évidence de signaux rares comme le boson de Higgs. L'architecture du détecteur DØ a donc été améliorée dans tous les domaines de détection. Les principaux éléments répartis du centre du détecteur vers l'extérieur, pour le Run II, sont décrits ci-dessous. Etant donné que j'ai effectué mes travaux de tâche de service sur le calorimètre, et que par ailleurs le calorimètre est l'une des composantes primordiales pour l'analyse de cette thèse, il sera présenté plus en détail.

2.2.3.1 Le détecteur de traces

Le détecteur de traces (ou trajectographe), permet de reconstruire la trace des particules chargées. Les informations sur la trace sont essentielles pour de nombreuses analyses à DØ et notamment l'identification des jets de quarks b dans la recherche du Higgs, (cf. 7.7.1). Le trajectographe est constitué de trois parties : la partie la plus centrale est un détecteur de vertex appelé SMT (Silicon Microstrip Tracker), il permet de mesurer la position du vertex d'interaction principal et de déterminer la trajectoire et l'impulsion des particules chargées ; ensuite un détecteur à fibres scintillantes appelé CFT (Central Fiber Tracker) entoure le SMT. Il sert a reconstruire la trajectoire des particules chargées pour mesurer leur impulsion ; enfin un aimant solénoïdal (Solenoidal Magnet) entoure le SMT et le CFT. Il produit un champ magnétique uniforme, qui va courber la trajectoire des particules chargées et permettre de mesurer leur charge et leur impulsion. Les matériaux ont été choisis afin de supporter les hautes doses de radiation du Run II. Le schéma du détecteur de traces de DØ est présenté à la figure 2.7.

La couche zéro [26] : elle a été installée durant le Shutdown entre le Run IIa et le Run IIb. Elle a pour but d'augmenter la protection du SMT et du CFT contre les radiations. Elle permet également d'améliorer la reconstruction des traces ainsi que la mesure du paramètre d'impact utile à l'identification des jets de quarks b.

Le SMT [27] : il s'étend sur une longueur de 64 cm, il a une partie centrale qui couvre $|\eta_{det}| < 1.5$ et une partie avant qui couvre la région $|\eta_{det}| < 3$.
La partie centrale se compose de 6 modules formés de tonneaux et de disques “F”, (cf. fig. 2.10). Chaque tonneau est constitué de support en Beryllium ayant quatre couches de détecteurs Silicium de rayons 2.7, 4.5, 6.6, et 9.5 cm, (cf. fig. 2.11). Les couches 1 et 3 sont composées d'une face recouverte de piste permettant la mesure en $r - \phi$. Les couches 2 et 4 sont formées d'une face permettant la mesure en $r - \phi$, et d'une autre sur laquelle les pistes formes entre elles un angle stéreo de 2 °. Cet agencement permet la reconstruction des points d'impact en 3 dimensions. Chaque tonneau est fermé par un disque “F” découpé en 12 sections et recouverts de deux

couches de détecteurs Silicium faisant un angle de 30° entre elles. Les disques "F" sont au nombre de douze et sont situés à $|z| = 12.5$, 25.3, 38.2, 43.1, 48.1 et 53.1 cm. Ils ont un rayon interne de 2.57 cm et un rayon externe de 9.96 cm
La partie avant comporte deux grands disques "H" situés à $|z| = 100.4$ et 121.0 cm, avec un rayon interne de 9.5 cm et rayon externe de 26 cm. Pour le Run IIb, les deux disques "H" les plus externes ont été enlevés et leur système de lecture est utilisé par la couche zéro.
La charge déposée sur chaque piste de silicium est recueillie par un canal individuel et traitée par un circuit intégré monté sur le détecteur. Le SMT compte 793 000 voix de lecture et a une d'épaisseur de 0.035 X_0[1].

Le CFT [28] : il couvre la région de rapidité $|\eta_{det}| < 2$, et est constitué de huit cylindres concentriques de rubans de fibres scintillantes, de rayon compris entre 20 et 52 cm, (voir fig. 2.7). Les deux cylindres centraux mesure 1.66 m de long et les six autres 2.52 m. Les cylindres ont chacun deux doublets de fibres, le doublet interne est aligné suivant l'axe du faisceau et le doublet externe fait un angle de $\pm 3°$ stéréo avec l'axe z. Cette agencement permet une reconstruction tridimensionnelle des trajectoires. Les fibres scintillantes sont au nombre de 76 800 avec un diamètre de 835 μm. Elles sont composées d'un mélange de polystyrène, de paraterphényl (1% en poid) et de 3-hydroxyflavone (1 500 ppm), permettant ainsi de réémettre l'énergie déposée par les particules sous forme de photons de longueur d'onde de 530 nm. Chaque fibre scintillante à une extrémité formée d'aluminium, permettant de réfléchir la lumière, pour minimiser les pertes, vers l'autre extrémité qui est couplée à une fibre transportant la lumière jusqu'à un Photo-Compteur de Lumière Visible (VLPC : Visible Light Photo Counter), qui assure la conversion des photons en signal de charge. Pour réduire le bruit thermique, un système de refroidissement cryogénique maitient la température du VLPC entre 6 et 13K. Le CFT compte 71 680 voix de lecture et a une d'épaisseur de 0.013 X_0.

L'aimant solénoïdal [31] : il entoure le SMT et le CFT, il a un diamètre externe de 1.42 m et une longueur de 2.73 m, (voir fig. 2.12). Il délivre un champ magnétique uniforme de 2 T parallèle à l'axe z, permettant de courber la trajectoire des particules chargées et ainsi de mesurer leur charge et leur impulsion. Il possède deux couches supraconductrices qui fonctionnent à une température proche de 10 k. L'aimant solénoïdal a un épaisseur de 0.9 X_0.

Performance des détecteur de traces L'efficacité de reconstruction est de plus de 95% dans la région centrale $|\eta| < 1.5$ pour des particules d'impulsion supérieure à 1.5 GeV. En deça, les particules tendent à spiraler avec un rayon de courbure n'impliquant plus qu'un nombre réduit de points dans les couches du CFT. La résolution sur l'impulsion transverse des traces de faible impulsion dans le trajectographe est de $\sigma_p/p_T^2 \approx 0.2\%$. La résolution sur la mesure du paramètre d'impact s'échelonne de 15 à 30 μm en $(r,\ \phi)$ en fonction de la pseudo-rapidité et pour différentes valeurs de l'impulsion transverse. La résolution sur la recontruction des vertex secondaires,

[1] X_0 est la longueur de radiation, c'est-à-dire la distance au bout de laquelle un électron a perdu une fraction $1 - \frac{1}{e}$, soit 63% de son énergie initiale.

basée sur des traces à haut paramètre d'impact, est de l'ordre de 40 μm dans le plan $(r,\ \phi)$, et 100 μm dans le plan $(r,\ z)$. Ces performances dépendent de la multiplicité de traces de l'évènement [29] [30].

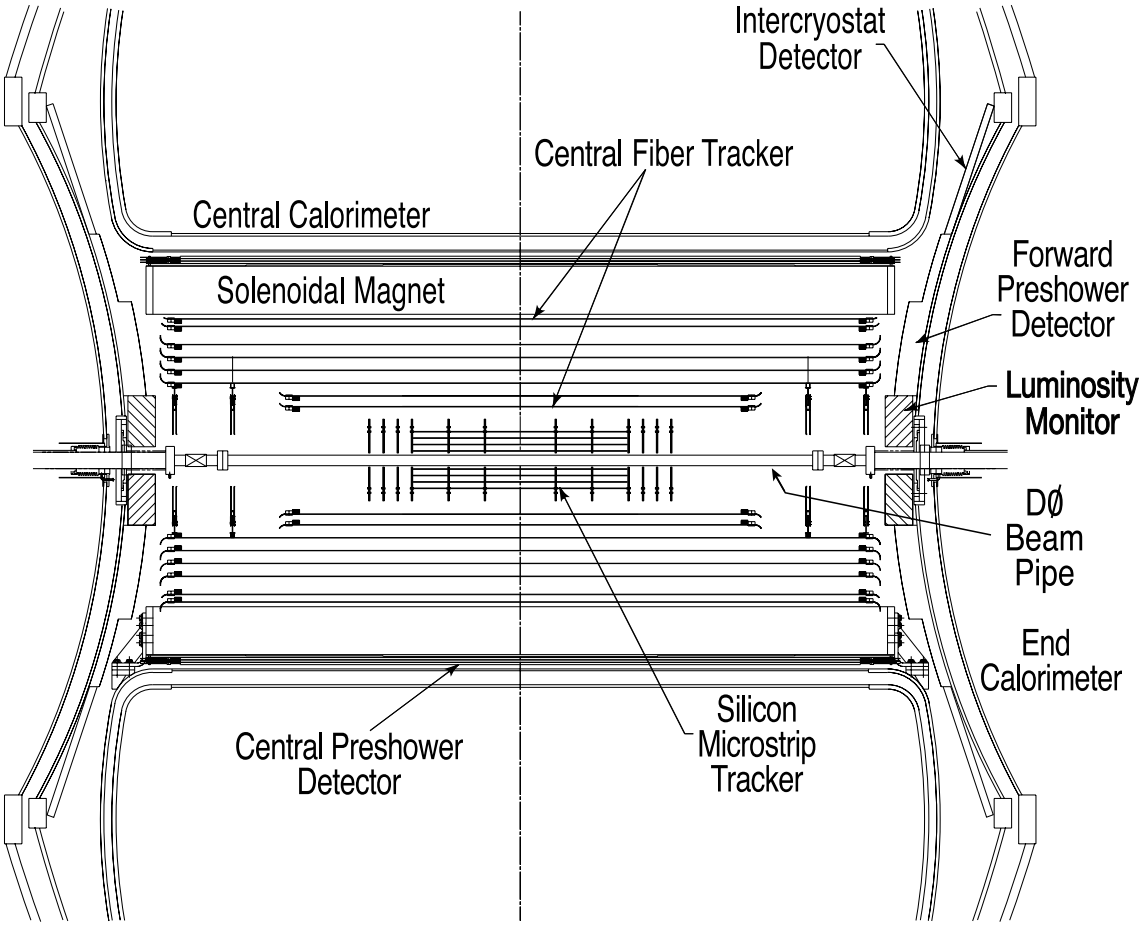


FIG. 2.7 – Schéma montrant les détecteurs de traces de DØ .

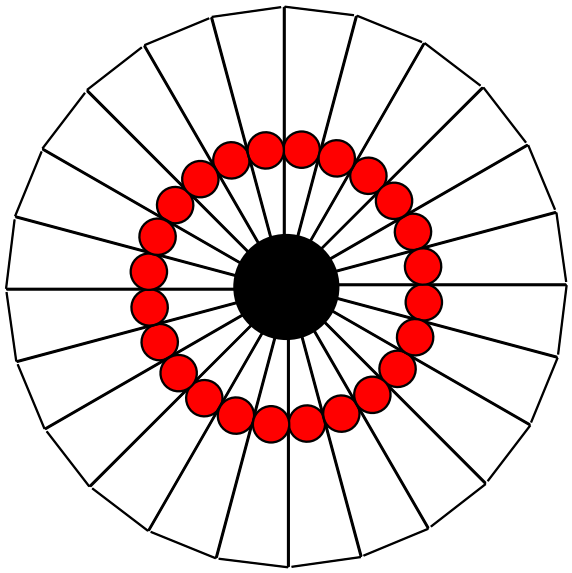


FIG. 2.8 – Vue des scintillateurs et des tubes photomultiplicateur (cercles en rouge) d'un compteur de luminosité.

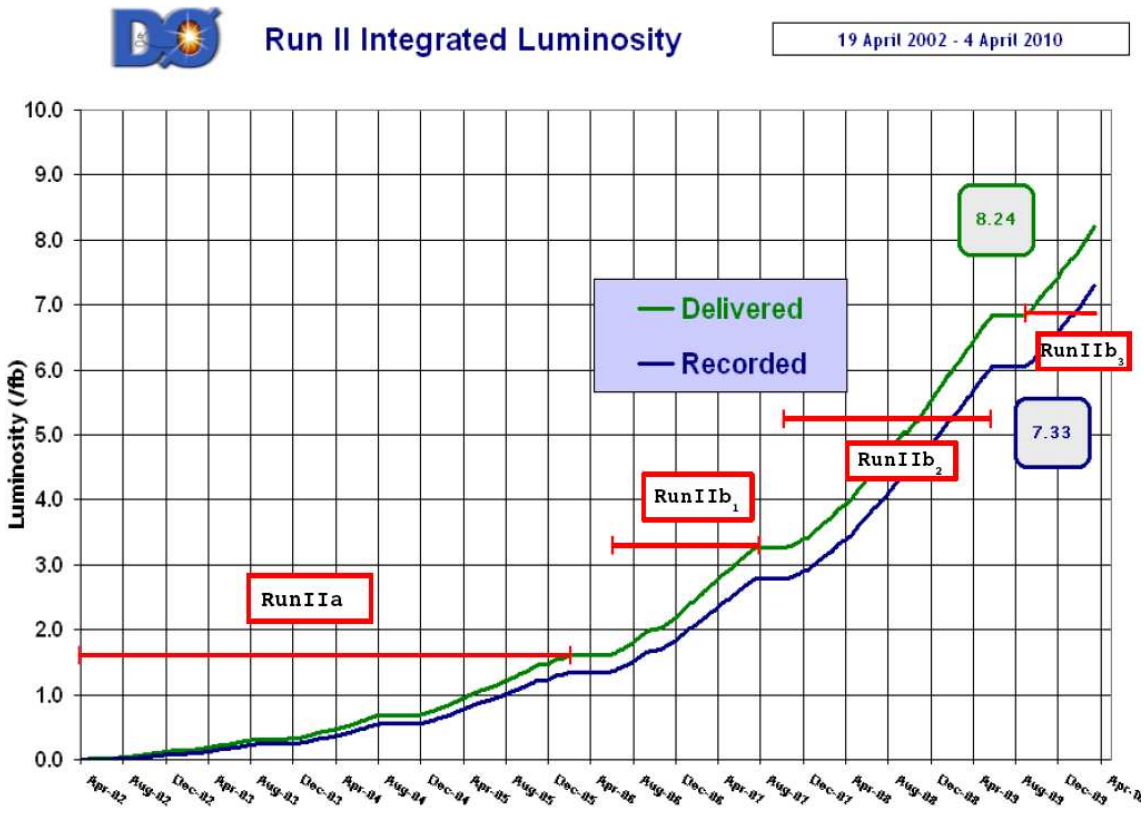


FIG. 2.9 – Evolution de la luminosité intégrée délivrée (courbe du haut en vert), et enregistrée (courbe du bas en bleu), par le détecteur DØ au cours du Run II. Les différentes périodes de prise de donnée au Run II qui sont le Run II_a, le Run IIb_1, le Run IIb_2 et le Run IIb_3, sont également présentées.

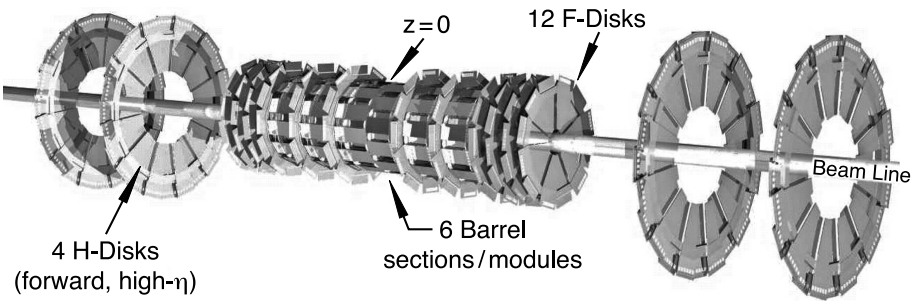


FIG. 2.10 – Schéma du détecteur de Vertex. Composé de six tonneaux, de douze disques F, et de quatre disques H.

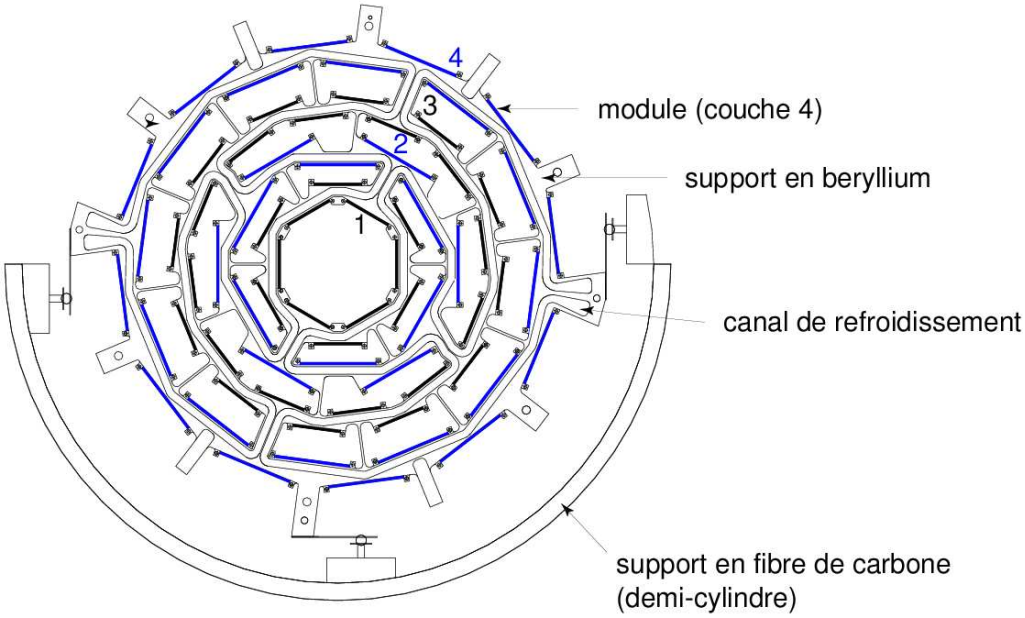


FIG. 2.11 – Coupe transverse d'un tonneau. On voit les supports en Beryllium et les quatre couches de détecteurs Silicium

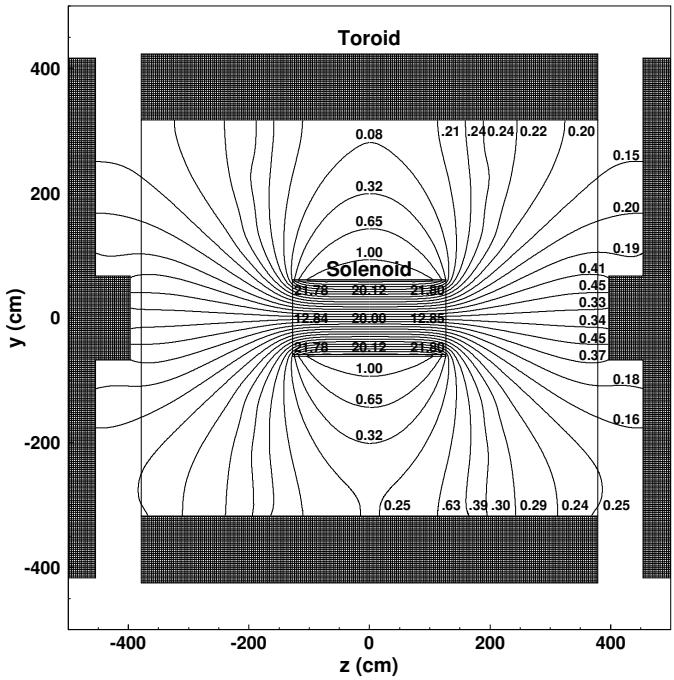


FIG. 2.12 – Schéma dans le plan (x, y) du champ magnétique central en KG, dans le détecteur DØ .

2.2.3.2 Les détecteurs de pieds de gerbes

Les détecteurs de pieds de gerbes [32] [33] se situent entre le trajectographe et le calorimètre, (voir figure 2.7). Ils servent à limiter la dégradation de la mesure de l'énergie par le calorimètre, à cause de l'aimant central. Ils permettent également d'identifier les particules en analysant les mécanismes de développement de gerbes des hadrons d'une part, et des électrons et photons d'autre part. Les détecteurs de pieds de gerbes sont composés de pistes de scintillateurs triangulaires, cette géométrie optimise la résolution sur la position. Les fibres sont composées de plastique de polystyrène, de paraterphényl (1%) et de 150 ppm au dyphényl stilbène. Les pistes sont traversées en leur centre par une fibre scintillante WLS (Wave Length Shifting) qui transporte la lumière vers l'électronique de lecture. Les détecteurs de pieds de gerbes ont une épaisseur de 1.3 X_0.
Dans la région centrale en rapidité $|\eta_{det}| < 1.3$, le détecteur (CPS : Central PreShower) est un cylindre long de 2.6 m couvrant l'aimant solénoïdal. Le CPS consiste en trois couches de pistes triangulaires, l'une fournissant la mesure longitudinale, les 2 autres formant un angle stéréo entre elles, permettant la reconstruction de points en trois dimensions.
Dans la région des bouchons, les détecteurs (FPS : Forward PreShower) couvrent la région $1.5 < |\eta_{det}| < 2.5$. Ils sont fixés sur la paroi du calorimètre bouchon. Ils sont composés de quatre couches, deux situées en amont d'une couche de radiateur en plomb, qui permettent la mesure de l'énergie d'ionisation de la particule, et deux autres situées derrières le radiateur, qui mesurent l'extension de la gerbe initiée dans le plomb.

Performance des détecteurs de pieds de gerbes
Les performances de ces détecteurs sont évaluées sur le pouvoir d'identification des électrons et photons et de séparation électron/pion. Pour une efficacité supérieure à 90% une réjection d'un facteur 2 à 3 est obtenue sur les pions. Ce facteur de rejection s'ajoute à celui obtenu en utilisant la forme de la gerbe reconstruite dans le calorimètre seul. Cette performance se traduit, à efficacité de sélection des électrons constante, par une amélioration du rejet des jets identifiés par erreur comme électrons : l'utilisation des pré-échantillonneurs dès le premier niveau de déclenchement en association fournit par exemple une réjection d'un facteur supérieur à 2 de jets identifiés faussement comme électrons comparé à l'usage du calorimètre seul [34] [35].

2.2.3.3 Le calorimètre

Le calorimètre [36] permet de mesurer l'énergie des objets électromagnétiques et des hadrons, ainsi que l'énergie transverse manquante. Il permet également l'identification des électrons, photons et jets, et il participe à la détection des muons et des neutrinos.
Quand les particules traversent le calorimètre, ils interagissent avec le milieu absorbant en produisant une "gerbe" de particules, et déposent par conséquent leur énergie dans ce milieu. Les gerbes résultantes dépendent de la nature de la particule incidente. Il existe deux types de gerbe :
– les gerbes électromagnétique sont créées par les électrons et les photons. Les électrons perdent de l'énergie par rayonnement continu de freinage (ou Bremsstrahlung) en émettant des photons, qui produisent à leur tour des paires électrons positrons qui vont également émettre des photons. Ce processus s'arrête quand l'énergie des particules créées devient trop faible.
– les gerbes hadroniques sont créées par les hadrons qui interagissent avec le milieu absorbant, créant d'autres hadrons comme les pions chargés ou les protons, et qui se désintègrent par la suite.
Les gerbes hadroniques et électromagnétiques provenant de réactions différentes, ont donc des caractéristiques et notamment des formes différentes. La H-Matrix, cf. 3.3, est par exemple une variable qui tient compte des informations sur la forme de la gerbe pour distinguer les objets électromagnétiques des objets hadroniques. Les grandes différences entre ces deux types de gerbe sont les suivantes : la gerbe hadronique a une extension longitudinale et transversale plus importante que la gerbe électromagnétique ; et la gerbe électromagnétique commence plus tôt dans le calorimètre. Le structure du calorimètre de DØ en un calorimètre électromagnétique suivi d'un calorimètre hadronique est donc justifiée.

Le calorimètre de DØ est un calorimètre à échantillonnage, en effet chaque module composé de deux cellules, est une alternance entre le milieu absorbeur (sa composition dépend de la position de la cellule dans le calorimètre) dans lequel se développent les gerbes, et le milieu actif, dans lequel se trouve une électrode de lecture en cuivre, ionisable par les particules et composée d'argon liquide, voir fig. 2.13.

Le choix de ce dernier est motivé entre autres par la simplicité de l'étalonnage, et pour sa résistance aux radiations. Ces raisons contrebalancent l'inconvénient d'avoir le détecteur enfermé dans des cryostats complexes, maintenus à une température proche de 90 K. Les tours sont définies par l'empilement en profondeur des cellules pour un couple (η_{det}, ϕ) donné, elles servent au déclenchement et à la reconstruction. Le calorimètre est composé de trois parties se trouvant chacune est dans un cryostat : la région centrale (CC pour Central Calorimeter), les régions avant constituées de deux bouchons identiques Nord et Sud (EC pour End Calorimeter), et la région inter-cryostat (ICR pour Inter Cryostat Region), voir fig. 2.14.

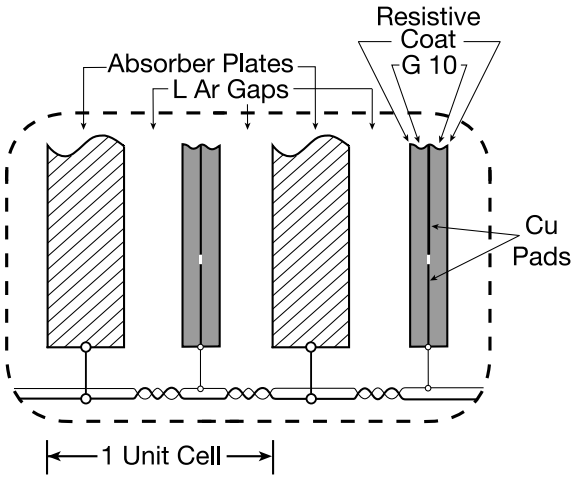


FIG. 2.13 – Schéma d'un module du calorimètre, il est composé de deux cellules. La cellule est l'élément de base du calorimètre.

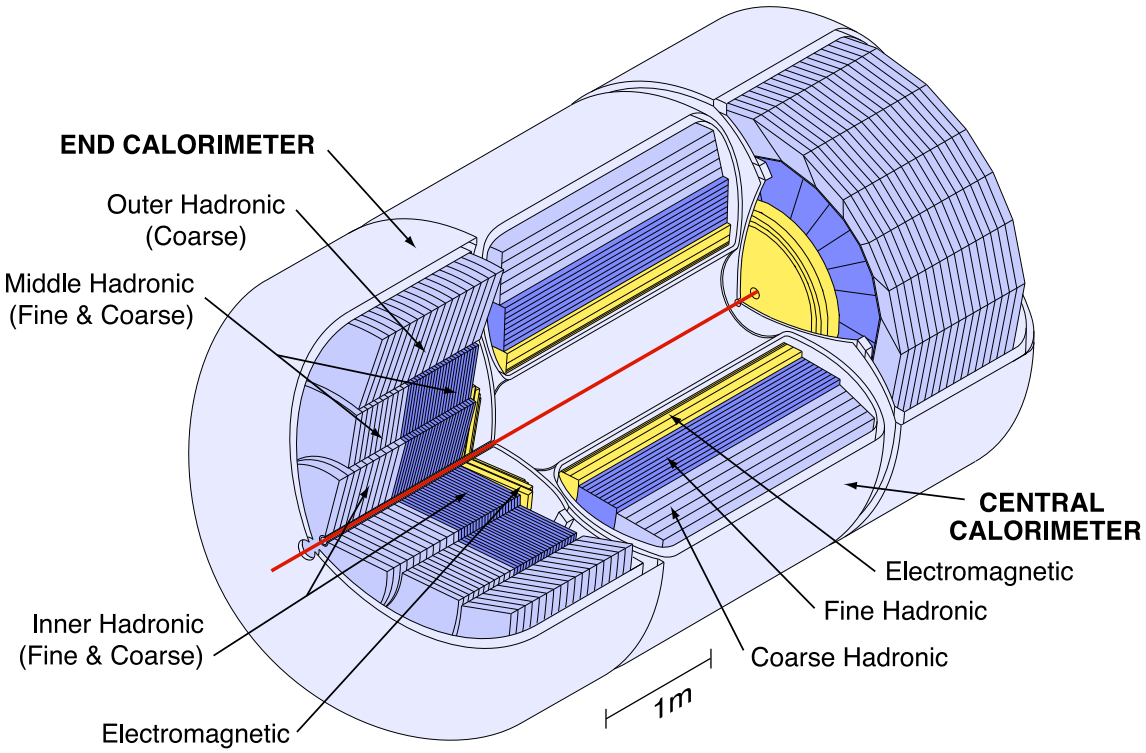


FIG. 2.14 – Les trois calorimètres (central et bouchons)de DØ dans leurs cryostats.

Le calorimètre central (CC) : il couvre la région $|\eta_{det}| < 1.1$, de forme cylindrique, il a une longueur de 2.26 m et un rayon interne (externe) de 0.75 (2.22) m. Le CC comporte trois couches : premièrement on trouve la couche électromagnétique (calorimètre EM), la plus proche du point d'interaction, le milieu absorbeur des cellules est composé de plaques d'uranium appauvri de 3 mm d'épaisseur. Le calorimètre EM est encore subdivisé en quatre sous-couches d'épaisseur 1.4, 2.0, 6.8 et 9.8 X_0, de la plus proche à la plus lointaine du point d'interaction ; deuxièmement on trouve la couche hadronique fine (calorimètre FH), le milieu absorbeur est composé de plaques constituées d'un alliage uranium-niobium (2%) de 6 mm d'épaisseur. Le calorimètre FH comporte trois sous-couches d'épaisseur 1.3, 1.0 et 0.9 X_0 ; troisièmement on trouve la couche hadronique grossière (calorimètre CH), le milieu absorbeur est composé de plaque de cuivre de 46.5 mm d'épaisseur. Le calorimètre CH ne possède qu'une couche d'épaisseur 3.2 X_0.
Les cellules des différentes couches sont segmentées en $\Delta\eta_{det} \times \Delta\phi = 0.1 \times 0.1$, afin de correspondre à la taille des gerbes électromagnétiques (1-2 cm) et hadroniques ($\sim$10 cm). Cette segmentation permet par exemple de déterminer la nature et la forme des jets partoniques, typiquement contenus dans un cône de rayon $\Delta R = \sqrt{\Delta\eta_{det}^2 + \Delta\phi^2} \simeq 0.5$. La troisième couche du calorimètre électromagnétique est plus finement segmentée $\Delta\eta \times \Delta\phi = 0.05 \times 0.05$, car c'est là où les gerbes électromagnétiques déposaient le maximum de leur énergie durant le Run I. Pour le Run II, l'ajout de l'aimant solénoïdale d'une épaisseur de près de 0.9 X_0 a eu pour conséquence que les gerbes se développent plus tôt, déposant donc plus d'énergie dans la deuxième couche du calorimètre. Comme mentionné en 2.2.3.2, l'ajout des FPS et CPS au Run II sont également motivés par ces inconvénients.

Les calorimètres bouchons (EC) : ils couvrent la région $1.5 < |\eta_{det}| < 4.0$. ils sont également constitués de trois couches (EM, FH et CH) :
– la couche EM à quatre sous-couches d'épaisseur de 1.6, 2.6, 7.9 et 9.3 X_0. L'absorbeur de la première couche est en fer, et celui des trois autres est en uranium appauvri.
– la partie hadronique est divisée en trois sections : une section interne qui compte quatre sous-couches FH avec des absorbeurs en uranium appauvri de longueur 1.1 λ_A (λ_A est la longueur d'absorption), et une sous-couche CH avec des absorbeurs en inox de longueur 4.1 λ_A ; une section médiane avec quatre sous-couches FH munies d'absorbeurs de longueur 0.9 λ_A, et une sous-couche CH d'absorbeur de longueur 4.4 λ_A ; et enfin une section externe formée de trois sous-couches CH avec des absorbeurs en inox de longueur 6.0 λ_A.
La segmentation de toutes les couches sauf celle de la troisième sous-couche EM est de $\Delta\eta_{det} \times \Delta\phi = 0.1 \times 0.1$. La granularité de la troisième sous-couche EM est plus fine et vaut $\Delta\eta_{det} \times \Delta\phi = 0.05 \times 0.05$.

La région inter-cryostat (ICR) : elle correspond à la zone non couverte par le calorimètre électromagnétique et elle est très peu instrumentée, $1.1 < |\eta_{det}| < 1.5$. Afin d'augmenter l'herméticité du calorimètre dans cette région, des détecteurs inter-

cryostat (ICD : InterCryostat Detector) et Massless Gap (MG) ont été installés [37] [38]. Ils permettent ainsi d'améliorer la mesure de l'énergie des jets, et celle de l'énergie transverse manquante. L'épaisseur des matériaux dans l'ICR est de 1.1 X_0.

- les détecteurs ICD couvrent toute la région $1.1 < |\eta_{det}| < 1.5$, (voir fig. 2.15). Leurs plaques de scintillateurs et l'électronique ont entièrement été refaites pour le Run II. Ils sont fixés sur la face externe des cryostats des bouchons autour des FPS. Sur chaque bouchon, l'ICD se compose de 16 tuiles, chacune couvre une région en $\Delta\eta_{det} \times \Delta\phi = 0.3 \times 0.4$. Chaque tuile est divisée en 12 scintillateurs qui couvrent chacun une zone $\Delta\eta_{det} \times \Delta\phi = 0.1 \times 0.1$. Les scintillateurs sont reliés à des tubes photomultiplicateurs par des guides d'onde.

- les détecteurs Massless Gaps sont de simples cellules sans plaques absorbeurs (leurs parois servant d'absorbeur). Elles sont plongées dans l'argon liquide, sur la face interne des cryostats des calorimètres CC et EC.

Performance du calorimètre

Les performances du calorimètre sont traitées dans la section 7.9.

Electronique du calorimètre

L'électronique du calorimètre est découpée en 12 châssis ou crates (voir fig. 2.16). Chaque châssis a 12 cartes ADC (Analog to Digital Converter) qui digitalisent le signal de dérive des électrons d'ionisation. Chaque carte ADC comporte 8 cartes BLS (Base Line Subtraction) qui soustraient la queue du signal enregistré lors d'une précédente collision (dit baseline), du signal mesuré. Les BLS se composent de quatre tours d'électronique pouvant accueillir 12 voies de lectures ou canaux. Au total on compte 55 296 voies de lectures pour le calorimètre, mais seuls 47 032 sont utilisées. Chaque canal est donc relié à un circuit électronique, permettant de lire le signal, et dont les éléments principaux sont un préamplificateur, un modulateur de forme, un BLS et un ADC, (voir fig. 2.17).
A la sortie de la cellule du calorimètre, le signal est donc amplifié. La vitesse de dérive des électrons est proche de 450 ns, mais les temps de charge et de décharge du préamplificateur sont respectivement de 450 ns et 15 μs, le préamplificateur n'a donc pas le temps de complètement se décharger pour la prochaine collision, et la queue du signal produit ainsi un baseline pour le signal suivant. A la sortie du préamplificateur, le signal passe par le BLS qui va lui soustraire le baseline. Le signal, après soustraction de la ligne de base, du piedestal, est ensuite modulé suivant une Gaussienne, de sorte que son maximum se trouve au environ de 320 ns, voir fig. 2.18. Le signal modulé est lu toutes les 132 ns, puis il est digitalisé par l'ADC et stocké pour être analysé.

Pour correctement convertir le signal de sortie de l'électronique de lecture en une énergie en GeV, le calorimètre subit plusieurs calibrations listées ci-dessous.

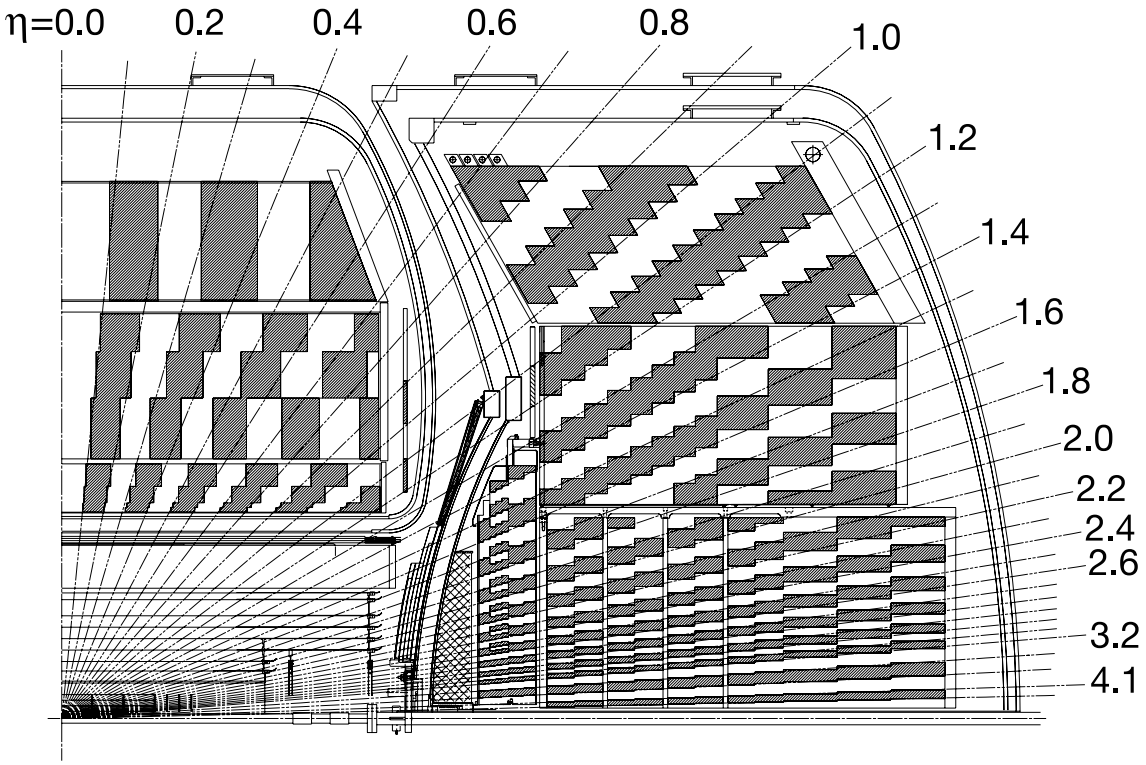


FIG. 2.15 – Schéma en coupe longitudinale du calorimètre, les détecteurs ICD se trouvent dans le zone inter-cryostat et couvrent la région $1.1 < |\eta_{det}| < 1.5$.

– Correction de la lecture de l'électronique calorimètre : l'électronique du calorimètre ne lit pas toujours le signal électrique au maximum de son amplitude. Il faut donc corriger ces valeurs mal lues. J'ai déterminé des facteurs de correction avec la méthode des trois échantillon, lors d'un de mes travaux de tâche des service, comme détaillé dans le chapitre 4.

– La calibration électronique : les signaux enregistrés sont proportionnels à l'énergie déposée. Des signaux électriques (dit DAC) aux caractéristiques connues sont injectés dans les canaux de lecture du calorimètre, puis le signal est digitalisé avec les ADC (on parle de coups ADC). La linéarité de la réponse électronique est étudiée en comparant la réponse électronique en fonction de l'amplitude du signal injecté. Par la suite, le signal ADC est corrigé des effets de non-linéarité avec une fonction qui dépend du nombre de coups ADC lus [61]. Enfin, des corrections en gains sont calculées en comparant le signal ADC avec le signal DAC.

– La calibration d'échelle : Pour convertir le signal ADC en une valeur d'énergie en GeV, le signal ADC est multiplié par une constante (1 ACD $\sim$ 0.3 MeV). Cette énergie qui correspond à l'énergie déposée dans l'argon liquide, est reliée à l'énergie

totale de la cellule par un poids qui dépend de la couche du calorimètre. Enfin, l'énergie totale de la particule électromagnétique est obtenue en sommant l'énergie de toutes les cellules de la gerbe.

– Correction des effets géométriques : cette correction obtenue sur des évènements simulés dépend de l'énergie de l'objet électromagnétique, de sa position par rapport au centre du détecteur, et de son angle d'incidence. La simulation permet de déterminer l'énergie perdue dans le calorimètre comme la différence entre l'énergie simulée et l'énergie reconstruite.

– Calibration en énergie du bosons Z : la constante de calibration de l'échelle d'énergie au pic de masse du Z a été déterminée par un ajustement de l'énergie des paires des électrons sur le pic de la distribution de la masse du boson Z [62].

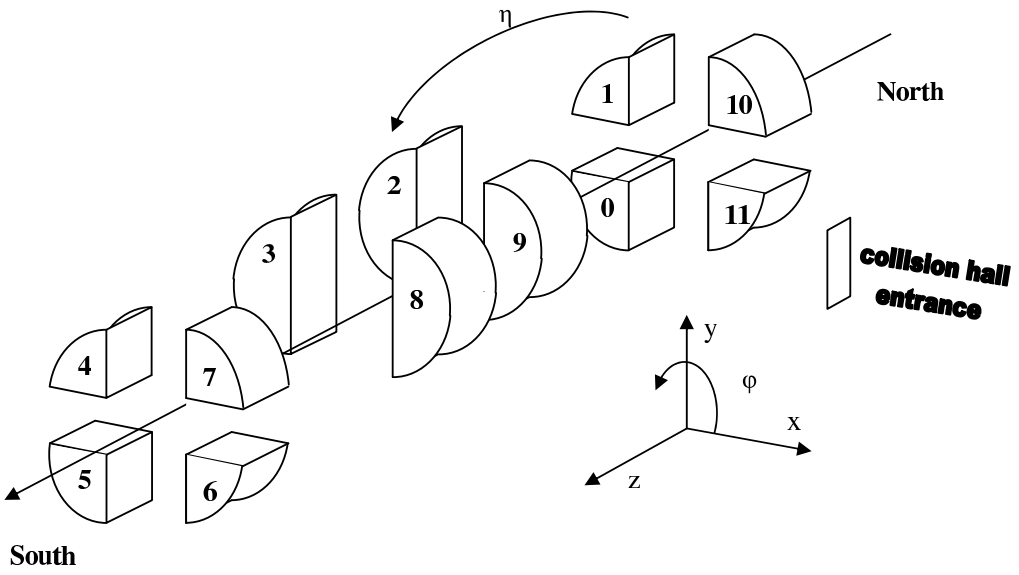


FIG. 2.16 – Vue des 12 châssis du calorimètre.

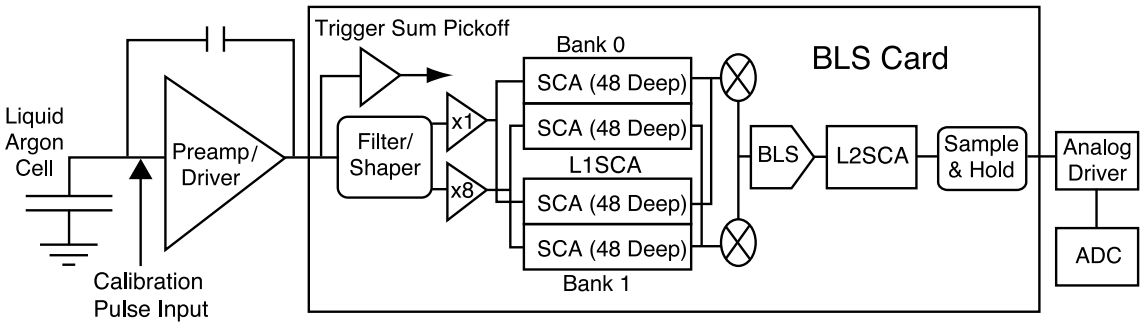


FIG. 2.17 – Electronique du calorimètre. Les éléments principaux sont un préamplificateur (Preamp/Driver), un modulateur de forme (Filter/Shaper), une carte BLS (BLS Card) et une carte ADC (ADC)

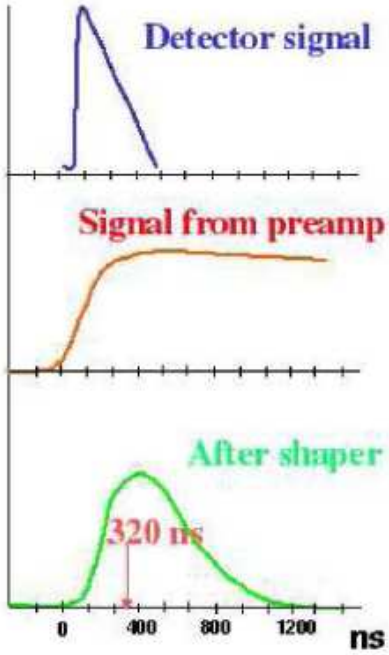


FIG. 2.18 – La forme du signal dans les différentes parties de l'électronique du calorimètre : en haut, le signal détecté ; au milieu, le signal amplifié ; en bas, le signal modulé.

Bruit de fond du calorimètre

L'énergie détectée dans le calorimètre ne provient pas que des particules engendrées dans les collisions $p\bar{p}$. Il est donc important de mesurer la contribution du bruit pour garantir une bonne performance du calorimètre. Voici trois principaux bruits du calorimètre :

– *Ring of fire* : à cause du manque d'espace dans la région inter-cryostat, l'électrode alimentant les couches électromagnétiques du calorimètre à une forme circulaire. Quand cette électrode capte un bruit, par exemple provenant d'un chantier externe, un dépôt d'énergie fictif est perçu pour toutes les cellules d'une même couche et du même η_{det}, cf. figure 2.19. En 2003, le système d'isolation électrique du détecteur a été remplacé, et la fréquence de ce bruit a nettement diminué.

– *Coherent noise* : correspond à un décalage cohérent des piédestaux de toutes les cellules d'une ou de plusieurs cartes ADC. Ce bruit est mal connu, mais les évènements qui y sont affectés présente une grande occupation d'une carte ADC et une faible largeur de la distribution du nombre de coups d'ADC. Ce bruit génère par exemple de faux jets, qui ne sont pas confirmés au niveau 1 du système de déclenchement, cf. 3.4.

– *Spanish fan* : ce bruit apparu en juin 2006, touche les cellules du calorimètre se trouvant le long du cryostat, dans la région $0.7 < |\eta_{det}| < 0.8$, en formant deux

demi-roues, cf. figure 2.20. Ce bruit se caractérise par une petite fraction électromagnétique des jets et une grande fraction de leur énergie déposée dans la couche CH.

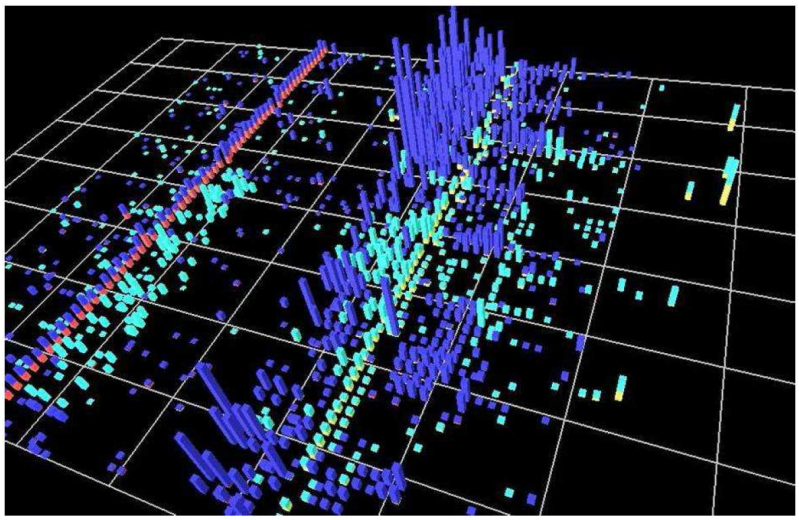


FIG. 2.19 – Evènement touché par le bruit *ring of fire*, dans le plan (η_{det}, ϕ).

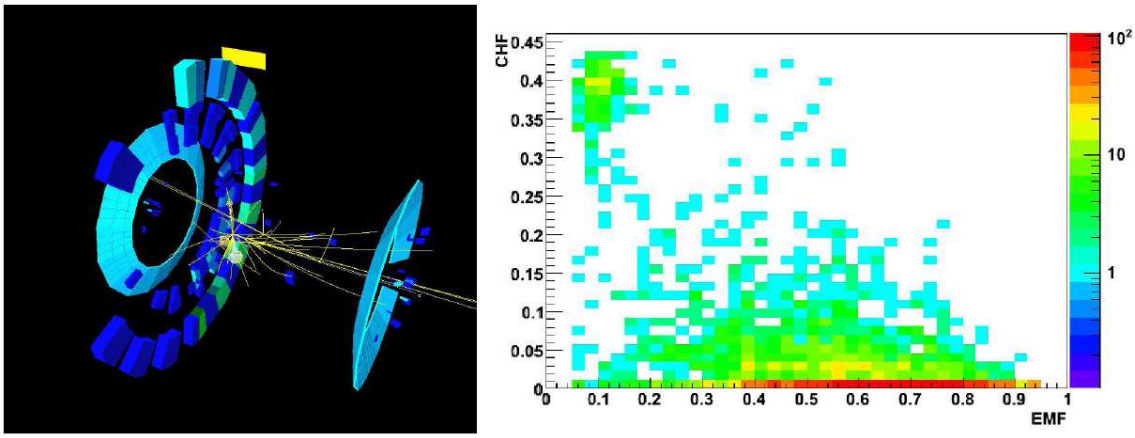


FIG. 2.20 – A gauche, représentation des dépôts d'énergie pour un évènement affecté par le bruit *spanishfan*. A droite, fraction d'énergie dans la couche CH (CHF) en fonction de la fraction électromagnétique (EMF) pour des jets se trouvant dans la région $0.7 < |\eta_{det}| < 0.8$. On y voit la présence du bruit à basse EMF et grande CHF.

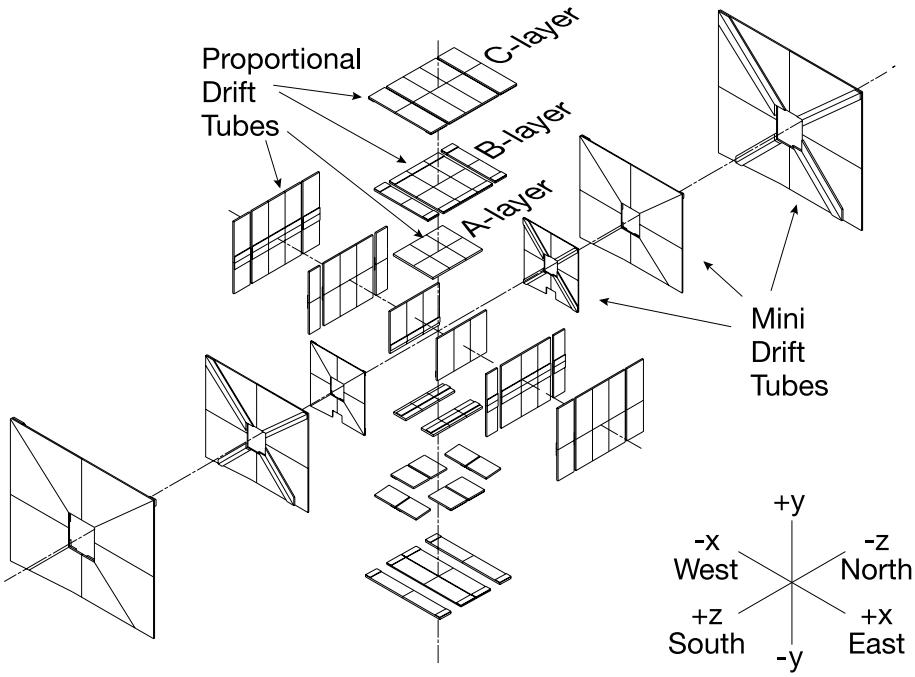


FIG. 2.21 – Eclaté des chambres proportionnelles à dérive du détecteur de muons.

2.2.3.4 Les détecteurs de muons

Le système de détection de muons [39] [40] couvre la région jusqu'à $|\eta_{det}| < 2$, il permet d'identifier les muons, et de déterminer leur trajectoire et impulsion. Il a trois couche A, B et C du centre vers l'extrémité du détecteur, et des aimants toroidaux placés entre les couche A et B permettent de mesurer l'impulsion du muon à l'aide de la courbure de la trajectoire[2], voir fig. 2.6.
Le système de détection de muons se compose d'une partie centrale (WAMUS : Wide Angle MUon System) couvrant la région $|\eta_{det}| < 1$, et d'une partie avant et arrière (FAMUS : Forward Angle MUon System) couvrant la région $1 < |\eta_{det}| < 2$. Le WAMUS est composé de chambres proportionnelles à dérive (PDT : Proportional Drift Tubes). Les FAMUS ont été installés pour le Run II. Ils sont composés de minichambres à dérive (MDT : Mini Drift Tubes), voir fig. 2.21. Les WAMUS et FAMUS ont respectivement une et trois couches de scintillateurs, qui sont utilisées par le système de déclenchement. Par ailleurs, pour réduire le bruit de fond cosmique, des fibres scintillantes entourant le WAMUS, mesure la différence de temps entre le moment où un muon traverse un PDT et celui où il frappe le détecteur, ce qui permet de vérifier qu'un muon provient bien d'un croisement de faisceau, voir fig. 2.22.

2.2.4 Le système de déclenchement

Au Tevatron, étant donné le grand nombre de collisions et faible section efficace des évènements recherchés, l'intégration d'un système de déclenchement (ou trigger)

[2]Ecart entre la direction du muon à l'entrée et à la sortie de l'aimant.

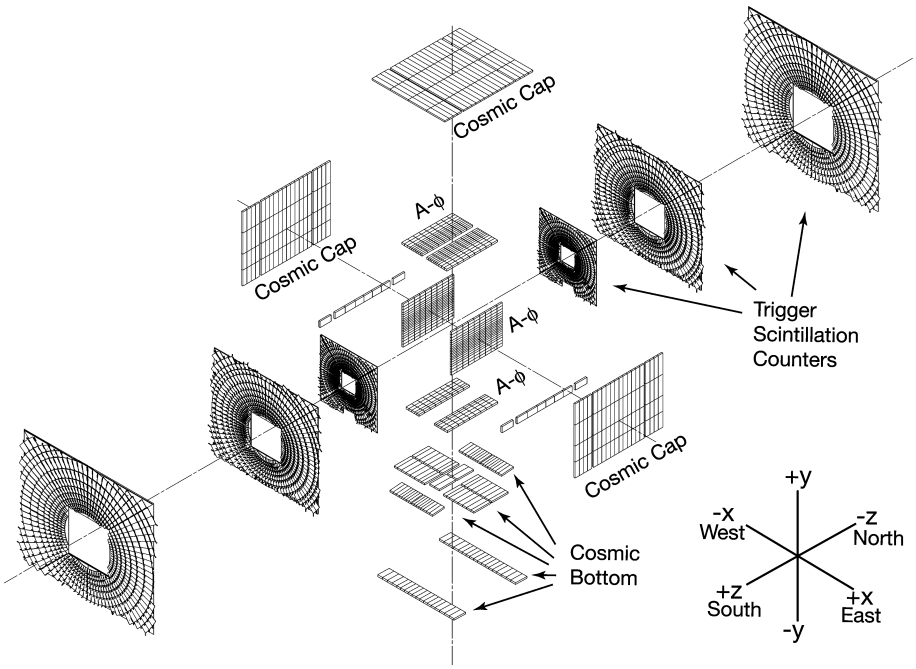


FIG. 2.22 – Eclaté des couches de scintillateurs et de détecteurs anti-cosmiques.

et d'acquisition spécifique aux topologies du signal recherché est indispensable. Le système de déclenchement est composé des niveaux suivants, (voir fig. 2.23) :

- le niveau 0 sélectionne les évènements dans lesquels une collision $p\bar{p}$ inélastique a eu lieu. Ces évènements sont appelés évènements de biais minimum. Lorsque c'est le cas, un signal est mesuré par les détecteurs de luminosité, et le niveau 0 est validé. La fréquence de sortie du niveau 0 est de 1.7 MHz.
- le niveau 1 utilise les informations du CFT, des détecteurs de pied de gerbes, du calorimètre et du spectromètre à muon. A ce stade la prise de décision de l'électronique rapide est de 132 ns et la fréquence en sortie est réduite à 2 kHz.

- le niveau 2 utilise les informations fournies par les mêmes détecteurs qu'au niveau 1 plus celles du SMT, on peut ainsi avoir des informations sur la présence d'objets physiques (électrons, traces, etc.) dans les évènements. Le niveau 2 a un temps de décision de 100 μs et les évènements ont une fréquence de 1kHz.
- le niveau 3 reconstruit partiellement les évènements avec des algorithmes simples. La fréquence de sortie des évènements est réduit à 50 Hz.

Les niveaux 1 et 2 du système de déclenchement ont été modifiés pour le RunIIb, afin de faire face à l'augmentation de la luminosité [41] [42]. Les déclenchements du calorimètre, du CFT et des détecteurs de pied de gerbes ont été modifiés au niveau 1 ; le déclenchement des trajectographe a été amélioré au niveau 2.

Les niveaux 1, 2 et 3 correspondent à un ensemble de liste de termes (un terme

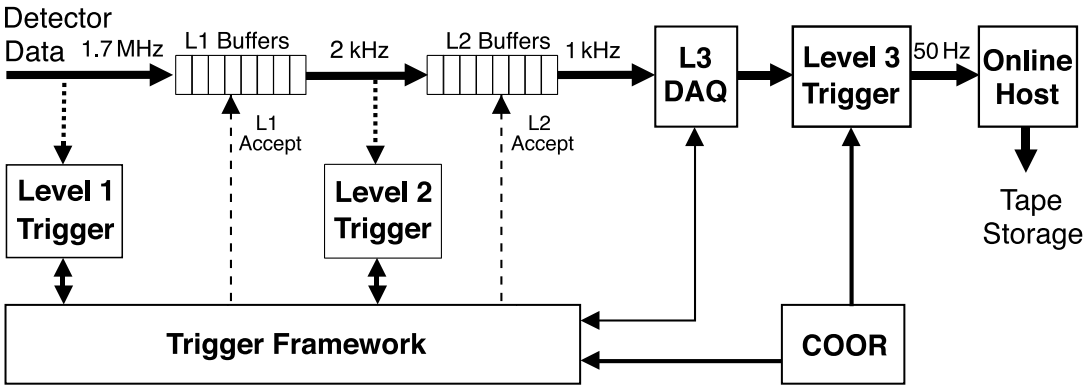


FIG. 2.23 – Architecture du système de déclenchement de DØ . Les contraintes sur les taux d'évènements et la durée de prise de décision sont indiquées à chaque niveau.

correspond à un ensemble de critères sur l'évènement), qui seront ensuite combinés, et appelés trigger. Chaque trigger correspond donc à une topologie d'évènement différente. Certains triggers ont une acceptance trop élevée, pour résoudre ce problème, des seuils de saturation (dit prescale) sont posés sur ces triggers. Par exemple, un trigger avec un prescale de 10 signifie que seul 1 évènement sur 10 qui satisfont à ses critères sera retenu.
Etant donné que le système de déclenchement n'est pas décrit par la chaîne de simulation de DØ , cf. 2.2.5, les analyses doivent utiliser des triggers ayant une efficacité proche de 100%, ou bien déterminer leur efficacité afin de repondérer les évènements simulés.

2.2.5 Reconstruction des évènements et format des données

La reconstruction des données

A la sortie du niveau L3, on récupère des données brutes (en anglais raw data) qui doivent être mises dans un format approprié pour être analysées. Les données brutes passent par le programme de reconstruction d0reco (DØ RECOnstruction program), qui va traduire leur signaux en objets physique (électrons, jets, etc.). La reconstruction des évènements se fait hors-ligne, après le stockage des données brutes sur bandes, et nécessite quatre grandes étapes :

- la première étape consiste à décoder l'information brute, l'association des canaux électroniques aux différents sous-détecteurs et l'application à chaque sous-détecteur de constantes de calibration. Puis les signaux sont regroupés pour former des amas (calorimètre ou détecteurs de pied de gerbe) ou des coups (trajectographes), auxquels sont associée la position dans l'espace de l'élément de détecteur correspondant ;
- la deuxième étape reconstruit les traces avec les coups reconstruits précédemment ;
- la troisième étape reconstruit les vertex à partir des traces. L'algorithme re-

construit d'abord les vertex primaires, puis les vertex secondaires, (voir section 3.2) ;
- la dernière étape identifie les particules et objets physiques, à partir des informations des trois étapes précédentes, ainsi que de tous les sous-détecteurs.

La chaîne de simulation

DØ possède une chaîne de simulation, indispensable dans les analyses, car elle permet de prévoir les bruits de fond et les signaux attendus. Elle est composée de plusieurs étapes :

- la génération du processus physique avec des logiciels spécialisés comme PYTHIA ou ALPGEN (voir section 7.3.1) ;
- l'interaction des particules avec les différents détecteurs à l'aide du logiciel d0gstar (pour DØ GEANT Simulation of the Total Apparatus Response) utilisant le programme de simulation GEANT-3 [43] ;
- et la reconstruction de la chaîne d'électronique du détecteur (d0sim) [44].

Notons que le système de déclenchement n'est pas pris en compte dans la simulation. Par la suite, les évènements simulés passent par la même chaîne de reconstruction que les données réelles, d0reco.

Format des données

d0reco produit des fichiers dit ThumbNails en format orienté objet (C++), voir fig. 2.24. A partir des fichiers thumbnails, on produit des fichiers CAF-trees qui peuvent être analysés avec l'outil de visualisation et d'analyse ROOT [45]. Pour limiter la taille des fichiers CAF-trees et faciliter les analyses, les données sont réduites. On appelle cette opération "skimming" ou écrèmage, et l'échantillon réduit est le "skim". Les skims sont réalisés par la collaboration suivant les intérêts de chaque groupe d'analyse, puis mis à leur disposition. Les fichiers CAF-trees contiennent moins d'information mais ils sont plus compacts et facile à utiliser avec CAFe (pour Common Analysis Framework Environment) [46]. CAFe est un cadre d'analyse commun orienté objet, composé de divers ensembles informatiques appelés "packages", et mise à la disposition des groupes de la collaboration. Il permet d'une part de limiter l'existence de formats multiples et la duplication des efforts au sein de la collaboration, et d'autre part il offre les outils nécessaires aux diverses analyses, comme la possibilité de transformer les fichiers CAF-trees en fichiers n-tuples qui contiennent juste les informations utiles à l'analyse, ou encore l'utilisation de logiciels pour appliquer des corrections, ou calculer des limites. C'est dans CAFe que les différentes analyses de cette thèse ont été effectuées.

2.3 Conclusion

Ce chapitre à présenté le complexe d'accélération de Fermilab, la production des faisceaux de protons et d'antiprotons ainsi que l'anneau du Tevatron. La zone de collision ainsi que les différents sous systèmes du détecteur DØ et notamment le

trajectographe et le calorimètre qui sont d'une importance capitale pour la recherche du boson de Higgs de cette thèse ont été vus. Enfin, le système de déclenchement et d'acquisition ainsi que la reconstruction des évènements on été abordés.

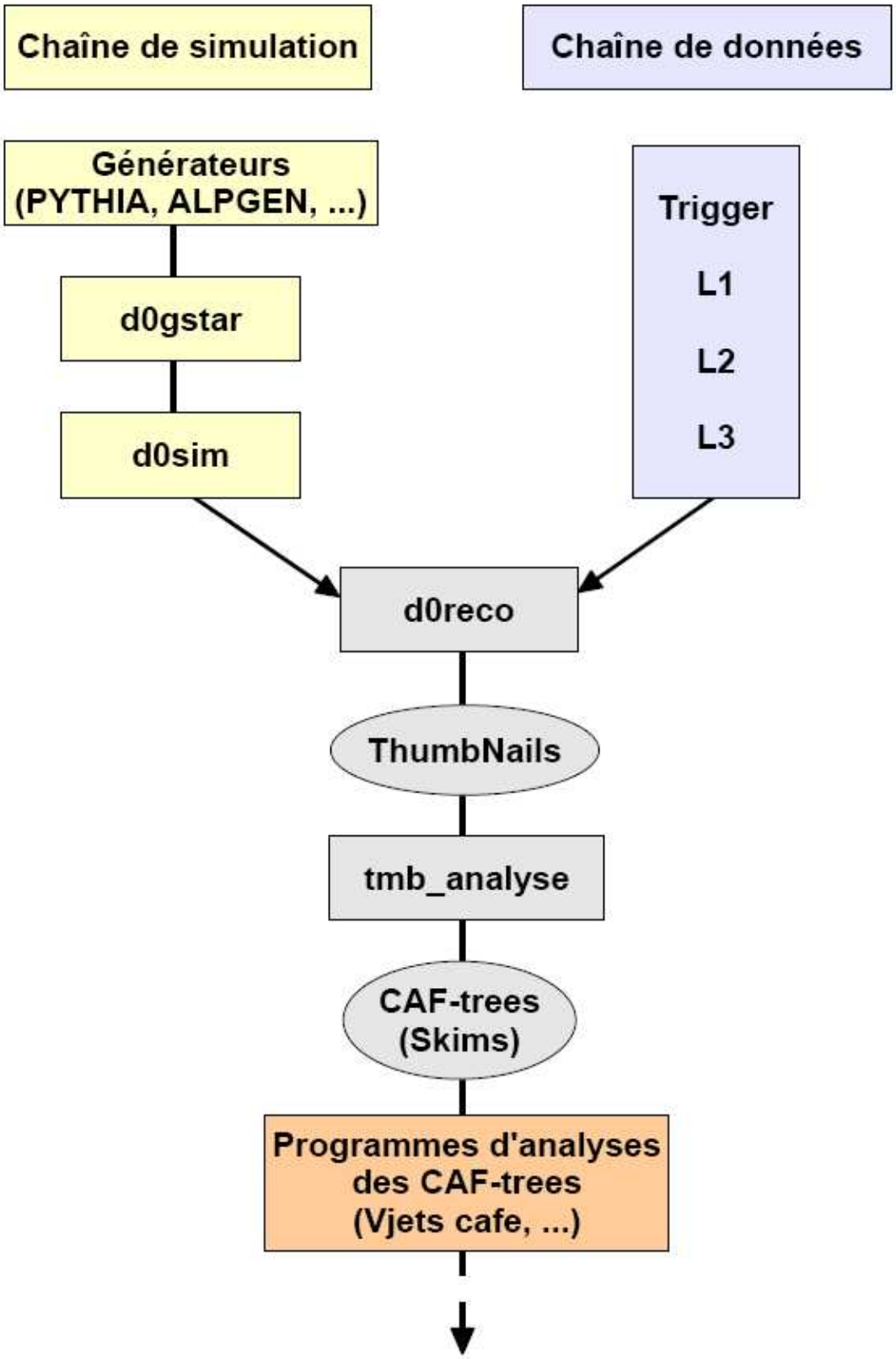


FIG. 2.24 – Chaîne de donnée et de simulation de DØ .

Chapitre 3

Objets physiques utilisés

Contents

3.1 Les traces **57**
3.2 Les vertex primaire et secondaire **57**
3.3 Les objets électromagnétiques **58**
3.4 Les jets hadroniques **59**
3.5 L'énergie transverse manquante **64**
3.6 Conclusion **64**

Ce chapitre explique comment les objets physiques utilisés dans cette analyse comme les traces, les vertex, les électrons et les photons, ou encore les jets hadroniques et l'énergie transverse manquante, sont reconstruits et identifiés à partir des données du détecteur. Une grande partie de ce travail de thèse ayant servi à identifier les électrons dans le calorimètre, ceux-ci sont présentés plus en détail dans les chapitres 5 et 6.

3.1 Les traces

Les traces sont reconstruites à l'aide des informations fournies par le trajectographe interne. En traversant le champ magnétique, les particules chargées laissent une trace le long de leur trajectoire quand ils interagissent avec le SMT et le CFT. Généralement, les évènements laissent 10^4 à 10^6 impacts, dans le trajectographe interne. Des algorithmes de reconstruction de traces utilisent ces impacts pour retrouver les traces [47] [48].

3.2 Les vertex primaire et secondaire

Les vertex primaires (PV : Primary Vertice) sont les points d'interaction des collisions $p\bar{p}$. Les vertex secondaires (SV : Secondary Vertice) sont déplacés par rapport au PV, ils proviennent de la désintégration de particules qui ont une durée de vie longue comme le quark b. Les vertex sont construits à l'aide d'algorithme

utilisant les traces d'impulsion transverse supérieure à 0.5 GeV ayant au moins deux impacts dans le SMT [49].

3.3 Les objets électromagnétiques

Les objets électromagnétiques sont les électrons et les photons. Quand ils traversent le calorimètre, ils forment une gerbe électromagnétique (cf. 2.2.3.3). Ils sont reconstruits à partir des tours du calorimètre à l'aide de l'algorithme de cône simple (Simple Cone Algorithme) [50] [51] [52] qui fonctionne de la manière suivante :

1. on recherche la tour calorimétrique dont l'impulsion transverse (au moins supérieure à 0.5 GeV) est la plus élevée. La direction en (η_{det}, ϕ) de cette tour est prise comme direction de référence ;
2. les tours se trouvant dans un rayon ΔR défini dans l'espace par :
$$\Delta R = \sqrt{\Delta\eta_{det}^2 + \Delta\phi^2} = 0.4 \tag{3.1}$$
autour de la direction de référence lui sont associées pour former un amas de tours. Ce dernier prend une nouvelle valeur d'impulsion transverse, de pseudo-rapidité η_{det} et d'angle ϕ, et donc une nouvelle direction ;
3. les étapes 1 et 2 sont itérées sur les autres tours de calorimètre d'impulsion transverse supérieure à 0.5 GeV, et les amas de tours formés seront des candidats électromagnétiques si leur impulsion transverse totale excède 1.5 GeV.

Outre les objets électromagnétiques, les hadrons déposent aussi de l'énergie dans le calorimètre et forment des gerbes. Ainsi, pour rejeter les candidats qui ne ressemblent pas aux objets électromagnétiques, on applique des coupures d'identification sur certaines variables [53]. Quelques coupures fréquemment utilisées sont expliquées ci-dessous :

EMfrac : la fraction d'énergie déposée dans la partie électromagnétique du calorimètre, voir fig. 3.1. Elle est définie comme le rapport entre l'énergie déposée dans les couches électromagnétiques du calorimètre et l'énergie totale déposée :

$$EMfrac = \frac{E_{EM}}{E_{tot}}, \tag{3.2}$$

généralement, ce rapport est supérieure à 90%.

Iso : l'objet électromagnétique doit être isolé, autrement dit l'activité calorimétrique autour du candidat doit être inférieure à 0.1, voir fig. 3.2. L'isolation est définie par :

$$Iso = \frac{E_{tot}(0.4) - E_{EM}(0.2)}{E_{EM}(0.2)}, \tag{3.3}$$

$E_{tot}(0.4)$ (resp. $E_{EM}(0.2)$) est l'énergie déposée dans le calorimètre (resp. les couches électromagnétiques) dans un cône de rayon ΔR = 0.4 (resp. 0.2) autour du candidat.

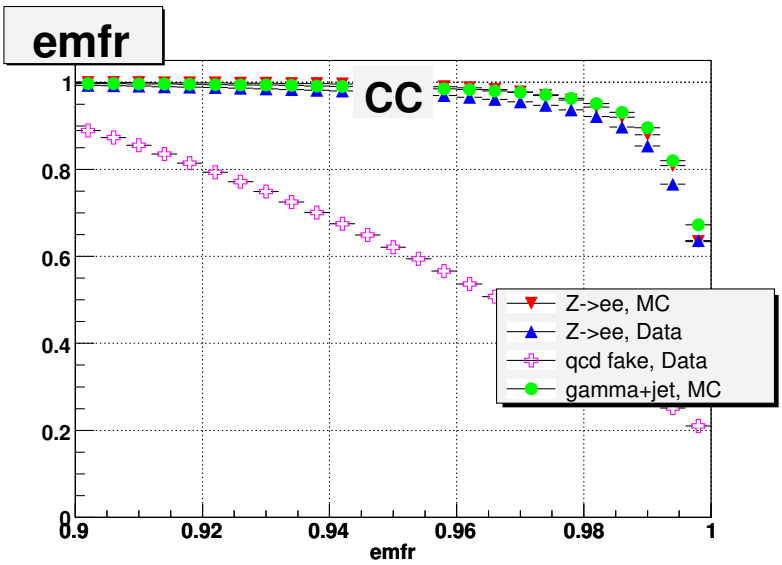


FIG. 3.1 – Distribution de la variable EMfrac dans le calorimètre central.

H-Matrix : les objets électromagnétiques doivent avoir une bonne forme de gerbe. Effectivement, les gerbes électromagnétiques ont une forme différente de celle des gerbes hadroniques, (cf. 2.2.3.3), ainsi on peut sélectionner les objets électromagnétiques à l'aide des informations sur la forme de leur gerbe. Pour cela, on sélectionne les gerbes dont la H-Matrix est inférieure à une certaine valeur, en générale 35 (20) dans le CC (EC) du calorimètre, voir fig. 3.3. La H-Matrix est un estimateur de χ^2, calculé à partir de l'inverse de la matrice de covariance construite avec 7 (HMx7 : H-Matrix7) ou 8 (HMx8 : H-Matrix8) variables qui décrivent la forme de la gerbe [53].

3.4 Les jets hadroniques

Lors des collisions hadroniques, à cause de l'interaction forte qu'ils subissent, les quarks et les gluons s'hadronisent (voir section 1.2) et forment un amas (ou jet) dans le calorimètre, voir figure 3.4. La reconstruction, l'identification et la détermination de la mesure de l'énergie des jets, nécessitent plusieurs étapes qui sont décrites ci-dessous.

La reconstruction : trois algorithmes sont utilisés pour la reconstruction des jets hadroniques : d'abord l'algorithme de cône simple, décrit en 3.3, identifie les amas du calorimètre ; ensuite l'algorithme de cône Run II rassemble les tours touchées par les amas proches en cônes stables (ou proto-jets) de rayon R = 0.5 (JCCB) ou R = 0.7 (JCCA), le choix s'est porté sur les JCCB dans les analyses de cette thèse ; et enfin une procédure de recouvrement entre jets (Splitting And Merging) différencie

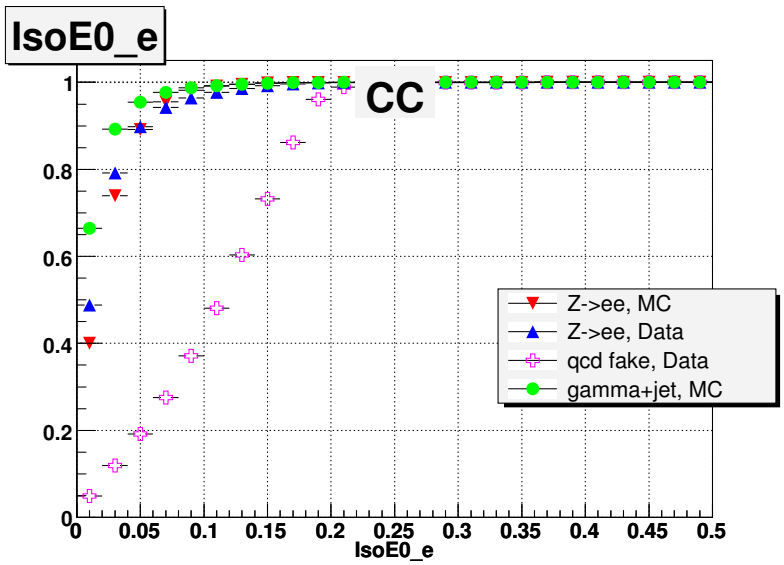


FIG. 3.2 – Distribution de la variable Iso dans le calorimètre central.

ou confond deux-à-deux les proto-jets, pour former les jets finaux qui doivent avoir une énergie transverse supérieure à 6 GeV.

L'identification : les jets reconstruits précédemment ne proviennent pas forcément d'un parton, et pour minimiser le taux de mauvaise identification, on applique des coupures d'identification [54]. Les candidats qui passent alors ces coupures sont qualifiés de "bons" jets. En exemple on peut mentionner les coupures suivantes :

- la fraction électromagnétique doit être $0.05 < \text{EMfrac} < 0.95$. EMfrac est bornée inférieurement pour rejeter les candidats qui ont plus de chance d'être du bruit de fond, et supérieurement afin de rejeter les candidats électromagnétiques (voir section 3.3).
- pour limiter les bruits électroniques importants dans les couches hadroniques grossières du calorimètre, on demande que la fraction d'énergie déposée dans celles-ci soit inférieure à 0.4.
- 90% de l'énergie du jet doit être contenue dans au moins deux tours, pour éviter les tours bruyantes.
- le jet doit être confirmé au niveau 1 du système de déclenchement, c'est-à-dire que le rapport entre l'impulsion transverse du jet et l'énergie des tours de déclenchement dans un cône $\Delta R < 0.5$ autour du centre du jet doit être supérieure à 0.5. A la sortie du calorimètre les signaux électriques sont envoyés à deux électroniques : une électronique de trigger L1 et une électronique de lecture de précision. Cette dernière comporte un grand gain qui fait que des bruits électroniques peuvent dans certaines conditions, simuler un dépôt calorimétrique semblable à celui d'objets physiques. La confirmation au niveau 1 réduit donc ce taux de faux jets générés par l'électronique.

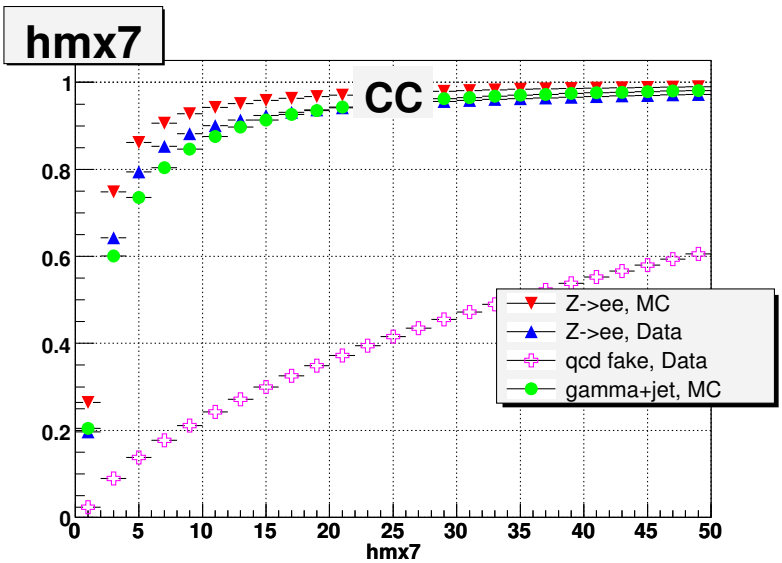


FIG. 3.3 – Distribution de la variable H-Matrix dans le calorimètre central.

La correction en énergie : la mesure de l'énergie des jets nécessite un calorimètre performant, mais également une bonne compréhension des effets qui perturbent leur mesure. La calibration en énergie des jets (JES : Jet Energie Scale) [55] [56] permet de rectifier les effets physiques et instrumentaux et donc de ramener l'énergie du jet mesurée dans le calorimètre à sa "vrai" valeur correspondant au niveau jet de particule. La JES est définie par :

$$E_{jet}^{part} = \frac{E_{jet}^{det} - O}{R \times S} \tag{3.4}$$

E_{jet}^{part} est l'énergie du jet au niveau particule et E_{jet}^{det} est l'énergie mesurée ; O est la correction de seuil (offset), c'est-à-dire l'énergie dans le cône du jet mais non associée à l'interaction dure, (par exemple les bruits électroniques, la radiation résiduelle de l'uranium du calorimètre, l'empilement avec les croisements de faisceaux précédents, ou les interactions multiples $p\bar{p}$ autres que les collisions dures), cette correction dépend du rayon du cône et du η_{det} du jet ; R est la réponse corrigée du calorimètre, elle tient compte par exemple de la perte d'énergie dans certaine région du calorimètre, elle dépend du η_{det} et de l'énergie du jet ; S est la correction de gerbe, ce facteur tient compte de la fraction d'énergie déposée hors du cône du jet par des particules du jet de particules, et également de l'énergie déposée dans le cône par des particules qui n'appartiennent pas au jet de particules.
Les jets simulés subissent des corrections supplémentaires comme décrit dans la section 7.3.2.

La correction des jets semi-muoniques : les jets qui présentent une (ou plusieurs) désintégration semi-muonique nécessitent une correction supplémentaire appelée correction semi-muonique [57]. En effet, la désintégration semi-muonique du

jet peu par exemple créer un neutrino et un muon parmis d'autres particules. Le neutrino n'interagit pas avec le détecteur et le muon interagit très peu avec le calorimètre, ce qui crée de l'énergie transverse manquante, (cf. 3.5). La correction semi-muonique permet donc de corriger ces effets. Cette correction est particulièrement importante pour les jets de quark b, qui dans plus de 10% des cas ont une désintégration semi-muonique du type $b \rightarrow \mu\nu X$.

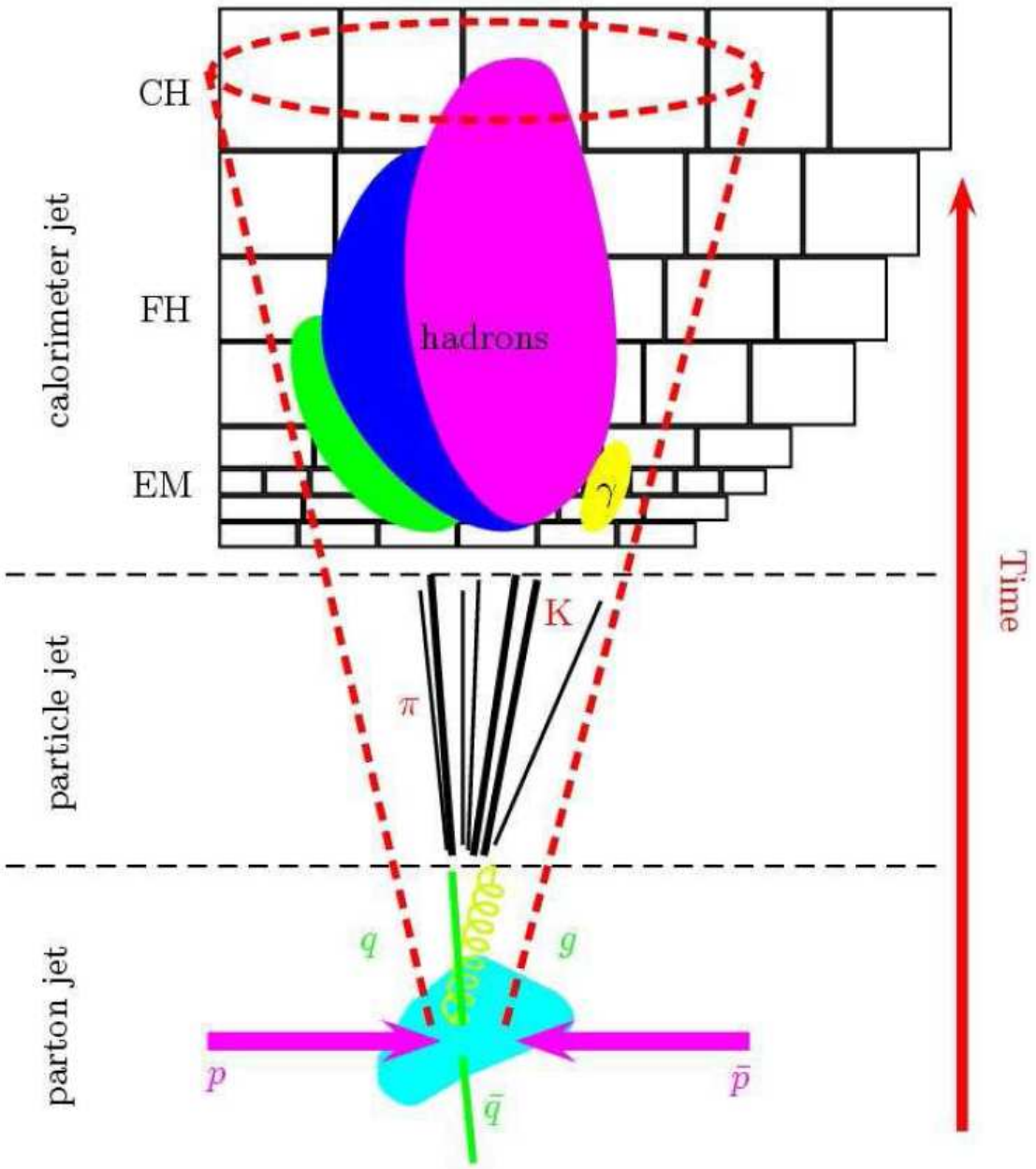


FIG. 3.4 – Schéma d'un jet hadronique en trois étapes : les partons issus de la collision dure émettant des gluons évoluent en parton jets (parton jet) ; le jet de particules (particle jet) correspond à l'hadronisation des partons et à la désintégration des hadrons crées en particules plus stables (pions, kaons, etc.) ; le jet dans le calorimètre (calorimeter jet) est la phase où ces particules interagissent avec les différentes parties (EM, FH et CH) du calorimètre pour former des gerbes.

3.5 L'énergie transverse manquante

L'énergie transverse manquante [58] est l'énergie qui échappe à la détection, elle peut être la signature d'un neutrino, car ce dernier n'interagit pas avec le détecteur. Sa détermination repose sur le principe de la conservation de l'impulsion transverse lors des interactions. Ce principe implique que la somme vectorielle des impulsions transverses p_T de toutes les particules produites dans l'interaction doit être nulle. Si la somme des p_T mesurées n'est pas nulle, c'est qu'une ou plusieurs particules ont emportées de l'énergie sans être détectées. Leur énergie transverse est appelée énergie transverse manquante, et elle est notée $\not{E}_T$. L'énergie transverse manquante est donc définie par l'opposé de la somme des énergies transverses déposées dans les cellules du calorimètre (les couches hadroniques grossières du calorimètre, trop bruyantes, ne sont pas comptabilisées) :

$$\vec{\not{E}_T} = -\sum^{cellules} \vec{p_T}. \tag{3.5}$$

L'énergie transverse manquante est une variable sensible car elle dépend des corrections apportées par la calibration des objets électromagnétiques et des jets d'une part, et des muons qui interagissent faiblement avec le calorimètre d'autre part. On doit donc lui appliquer les corrections de :

- la calibration en énergie des objets électromagnétiques (resp. jets), qui tient compte du changement de l'énergie des objets électromagnétiques (resp. jets) ;

- l'impulsion transverse des muons. Etant bien mesurée par le trajectographe et le spectromètre à muon, l'impulsion transverse des muons est soustraite de l'énergie transverse manquante

3.6 Conclusion

Ce chapitre a été consacré à la présentation des objets physiques utilisés dans ce manuscrit. D'abord les traces ainsi que les vertex primaire et secondaire ont été introduits. Ensuite les objets électromagnétiques, en particulier les électrons et les jets, très important dans l'analyse de cette thèse ont été détaillés. Enfin l'énergie transverse manquante a été définie.

Chapitre 4

Correction de l'énergie mesurée dans le calorimètre

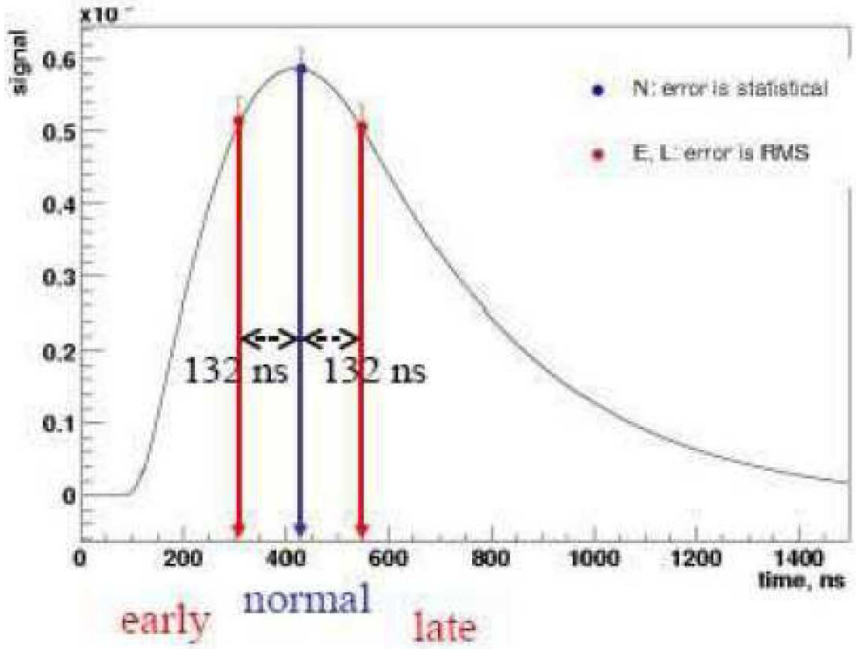
Contents

4.1 Méthode des trois échantillons 65
4.2 Echantillons de données 67
4.3 Sélection des évènements 67
4.3.1 Détermination du piédestal 67
4.3.2 Sélection du signal 67
4.4 Détermination des facteurs de correction 71
4.5 Conclusion 71

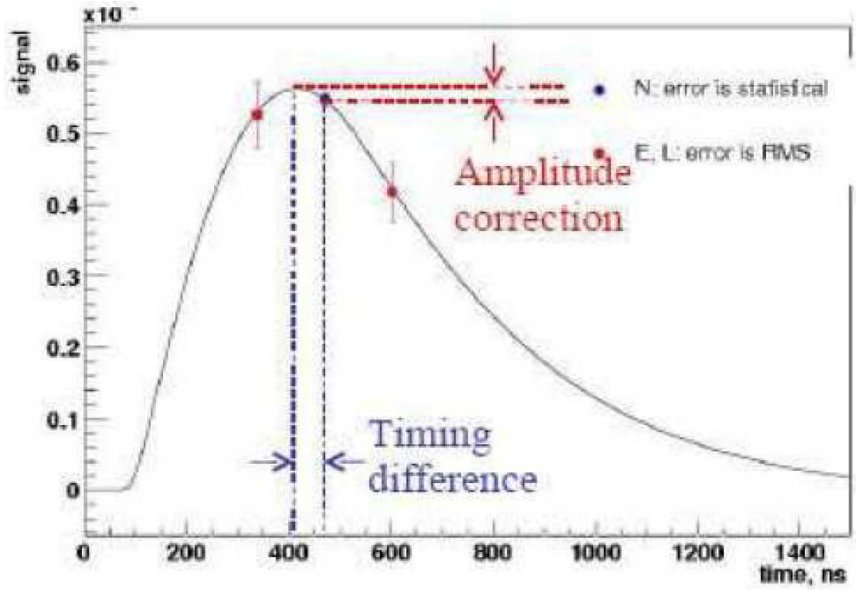
4.1 Méthode des trois échantillons

Pour diverses raisons (la calibration électronique, la longueur des câbles, le remplacement de matériaux, etc.), il peut y avoir un décalage de la lecture de l'électronique du calorimètre (cf. 2.2.3.3). Ceci a pour conséquence que le signal n'est pas lu au maximum de son amplitude, comme illustré sur la figure 4.1. Afin de corriger ces valeurs mal lues, et donc de déterminer la valeur du signal au maximum de son amplitude, on utilise la méthode dite des trois échantillons (ou Triple Sample Method) [59]. Cette technique se compose des trois étapes suivantes :

1. d'abord, l'amplitude du signal mesurée par les canaux du calorimètre, est lue trois fois, au temps t (N : Normal), t-dt (E : Early), et t+dt (L : Late), avec dt = 132 ns, cf. 4.1 ;
2. ensuite, pour chaque canal, on ajuste une fonction simulée [60] (attendue pour la forme du signal), ayant comme paramètre libre la normalisation et le temps, aux trois valeurs précédemment mesurées du signal ;
3. enfin, le facteur de correction à appliquer au signal pour corriger sa valeur mal lue, est calculé pour chaque canal.



(a) Bonne lecture.



(b) Mauvaise lecture.

FIG. 4.1 – Fig. 4.1(a), représentation d'un signal bien lu : de gauche à droite, le signal lu au temps t-dt (early), t (normal) et t+dt (late), avec dt = 132 ns ; Au temps t, le signal est correctement lu au maximum de son amplitude. Fig. 4.1(b), représentation d'un signal mal lu : au temps t, le signal n'est pas lu au maximum de son amplitude, cette mesure doit donc être corrigée.

4.2 Echantillons de données

Les données étudiées [63] ont été collectées lors de runs spéciaux enregistrés en février et décembre 2006, en février 2007 et en mai 2008, avec une luminosité instantanée faible comprise entre $27.10^{30}\text{cm}^{-2}\text{s}^{-1}$ et $100.10^{30}\text{cm}^{-2}\text{s}^{-1}$. L'ensemble des triggers triple timing a été utilisé, il requiert au moins un signal dans n'importe qu'elle tour du calorimètre. Un peu plus de 600 000 évènements ont été recueillis pour cette analyse.
Contrairement aux autres analyses de cette thèse, les données ont directement été récupérées par nos soins, dans la base de donnée de DØ appelée SAM [64], sous forme de fichiers thumbnails (cf. 2.2.5), puis nous les avons transformées en fichiers CAF-trees pour les analyser dans CAFe.

4.3 Sélection des évènements

4.3.1 Détermination du piédestal

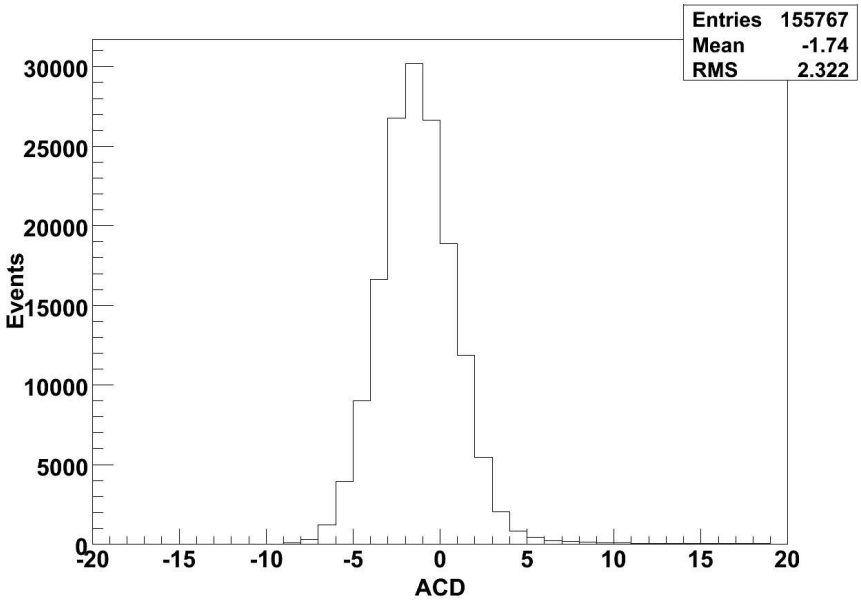
Pour connaître précisément le signal physique enregistré par l'électronique du calorimètre, cf. 2.2.3.3, le piédestal doit être déterminé. Le piedestal d'une voie électronique correspond à la charge mesurée par l'ADC en l'absence de signal. C'est le "zéro" électronique du canal en question. Il est important de connaître précisément sa position qui détermine la quantité à soustraire pour mesurer le signal physique. Pour cette analyse, les piédestaux sont compris entre ± 30 coups ADC. Donc pour sélectionner le piédestal, on pose les coupures minimales (ADCmin) et maximales (ADCmax) suivantes sur les coups ADC pour chaque signal :

- ADCmin0 = -30, ADCmin1 = -30, ADCmin2 = -30, pour respectivement le signal N, E et L.
- ADCmax0 = 30 , ADCmax1 = 30, ADCmax2 = 30, pour respectivement le signal N, E et L.

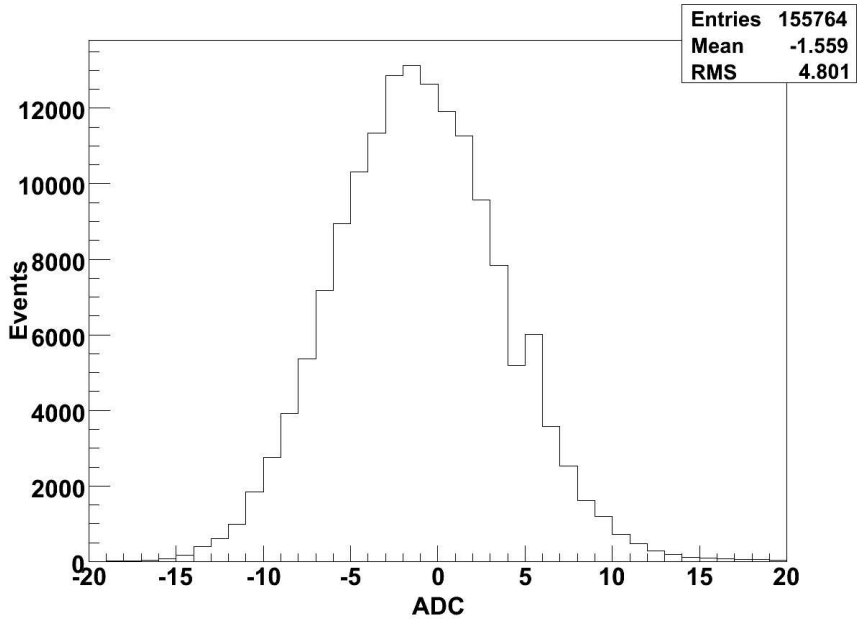
On obtient deux types de distributions de piédestal, voir fig. 4.2 : les distributions étroites correspondent aux canaux électromagnétiques ; et les distributions plus larges correspondent aux canaux hadroniques. Cette différence est due aux capacités des canaux hadroniques qui sont plus importantes.

4.3.2 Sélection du signal

La vraie valeur des signaux, correspondant à la différence entre la valeur du signal mesurée par l'ADC et la valeur moyenne des piédestaux, a été calculée pour chaque canal. Par la suite, les trois signaux de chaque canal ont été ajustés avec une courbe de référence simulée qui tient compte des caractéristiques de chaque canal (capacité, type de préamplificateur, longueur des câbles etc.), cf fig. 4.1. L'ajustement trouve le meilleur temps permettant d'avoir une lecture optimale de la valeur maximale de l'amplitude du signal. Pour chaque canal, l'ajustement a été réalisé sur la distribution des rapports des signaux E/N et L/N. L'avantage d'utiliser la rapport des signaux est que celui-ci est pratiquement indépendant du piédestal, ou des différentes coupures.

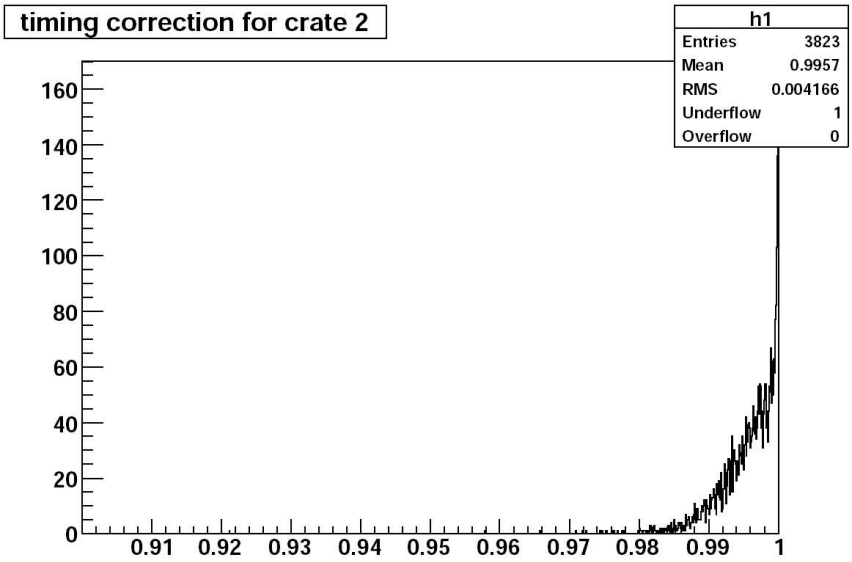


(a) Canal électromagnétique.

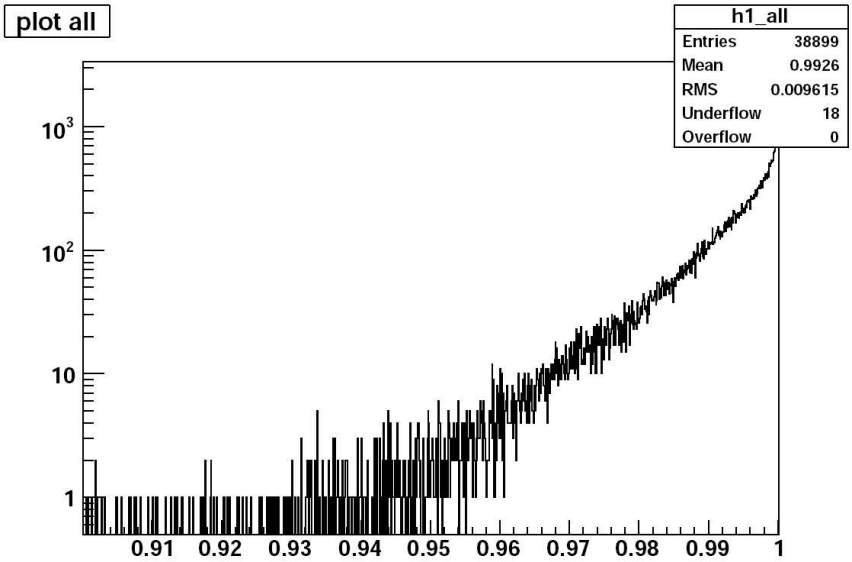


(b) Canal hadronique.

FIG. 4.2 – Distribution typique des piédestaux pour un canal électromagnétique 4.2(a) et hadronique 4.2(b).



(a) Facteurs de correction pour le châssis numéro 2.

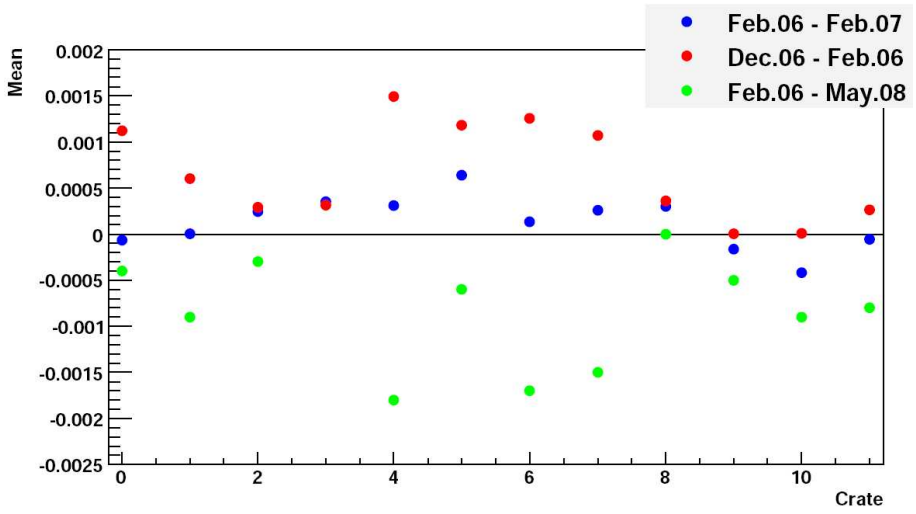


(b) Facteurs de correction pour tous les châssis.

FIG. 4.3 – Les facteurs de correction pour les données analysées de février 2006, pour le châssis numéro 2 fig. 4.3(a) en échelle linéaire, et pour tous les 12 châssis fig. 4.3(b) en échelle logarithmique.

11 0.69% 0.66% 0.56% 0.61%	9 0.67% 0.42% 0.42% 0.36%	8 0.28% 0.36% 0.31% 0.32%	6 1.31% 1.29% 1.06% 1.14%	crate Feb. 2006 Dec.2006 Feb. 2007 May 2008
0 0.42% 0.51% 0.31% 0.38%	2 0.43% 0.50% 0.46% 0.40%	3 0.35% 0.38% 0.34% 0.38%	5 0.87% 0.86% 0.70% 0.81%	
1 0.82% 0.82% 0.76% 0.73%			4 1.81% 2.01% 1.65% 1.63%	
10 0.70% 0.71% 0.60% 0.61%	9 0.41% 0.42% 0.42% 0.36%	8 0.32% 0.36% 0.31% 0.32%	7 1.20% 1.22% 1.06% 1.05%	

(a) Vue d'ensemble des facteurs de correction.



(b) Différence de la moyenne des facteurs de correction.

FIG. 4.4 – 4.4(a) vue d'ensemble des facteurs de correction : la première ligne (rouge) indique le numéro du châssis ; la deuxième (bleue), troisième (noire), quatrième (orange), et cinquième (verte) lignes indiquent les facteurs de correction calculés pour les données collectées respectivement en février et décembre 2006, février 2007 et mai 2008. 4.4(b) différence de la moyenne des facteurs de correction : entre février 2006 et février 2007 (bleue) ; entre décembre et février 2006 (rouge) ; et entre février 2006 et mai 2008 (verte).

4.4 Détermination des facteurs de correction

L'ajustement permet de corriger le temps de lecture, et ainsi de déterminer les facteurs de correction pour chaque canal. La figure 4.3 montre les facteurs de correction obtenus pour le châssis numéro 2 et pour tous les châssis. La figure 4.4 montre l'ensemble des facteurs de correction calculés pour chaque châssis et pour chaque période. En général, plus de 90% des canaux ont un facteur de correction supérieur à 0.996. On trouve également sur cette figure, la différence de la moyenne des facteurs de correction calculés pour différentes combinaisons des quatre échantillons analysés (cf. 4.2), cette différence est proche de 0. Cela montre que les facteurs de correction, en plus d'être faibles, sont stables dans le temps.

4.5 Conclusion

Une méthode servant à corriger la mesure de l'énergie par le calorimètre a été détaillée. Cette étude a montré que la plupart des facteurs de correction sont petits, et que leurs variations dans le temps sont négligeables. Ces facteurs de correction n'ont donc pas été utilisés car ils sont pratiquement identiques aux facteurs de correction établis précédemment [65].

Chapitre 5

Certification de l'identification des électrons dans le calorimètre

Contents

5.1 Méthode "Tag and Probe" **72**
5.2 Echantillons de données **73**
5.3 Sélection des évènements **74**
5.4 Les évènements simulés **75**
5.5 Le bruit de fond multijet **75**
5.6 Identification des électrons **76**
5.6.1 Critères de sélection 76
5.6.2 Efficacités d'identification 77
5.6.3 Correction des efficacités 77
5.7 Conclusion **78**

L'augmentation de la luminosité au Run II a nécessité de nombreux changements au niveau du détecteur, qui ont été mentionnés au chapitre 2.2.3. Ces changements concernent également les critères de sélection et d'identification des électrons. Ce chapitre décrit les moyens mis en œuvre pour certifier l'identification des électrons, dans la partie centrale (CC) et les parties avant (EC) du calorimètre, pour le Run IIb. Ma contribution à la certification de l'identification des électrons dans la région inter-cryostat (ICR) du calorimètre, qui est le complément de cette analyse, est présentée séparément dans le chapitre 6.

5.1 Méthode "Tag and Probe"

Les critères de sélection des électrons qui ont été définis pour le Run IIa ne sont pas appropriés pour le Run IIb. En effet, à cause de l'augmentation de la luminosité, on observe une baisse des efficacités pouvant atteindre 10% dans les données du Run IIb, voir figure 5.1. L'enjeu de cette analyse est donc de définir de nouvelles coupures pour le Run IIb, avec des variables peu sensibles à la luminosité d'une part, et qui offrent un bon rapport signal sur bruit de fond d'autre part [66]. En général,

Période	Skim
Run IIb1 (juin 2006 - août 2007)	CSG_CAF_2EMhighpt_PASS2_p21.10.00
Run IIb2 (octobre 2007 - février 2008)	CSG_CAF_2EMhighpt_PASS4_p21.10.00_p20.12.00 CSG_CAF_2EMhighpt_PASS4_p21.10.00_p20.12.01 CSG_CAF_2EMhighpt_PASS4_p21.10.00_p20.12.02

TAB. 5.1 – Skims utilisés pour l'analyse du Run IIb.

pour calculer les efficacités d'identification des électrons, on utilise la méthode "Tag and Probe" [67]. D'abord, on sélectionne des évènements possédant au moins deux électrons issus de la désintégration d'un boson Z, et dont la masse invariante après reconstruction se trouve dans la fenêtre de masse du boson Z, c'est-à-dire 80 GeV $< M_{ee} <$ 100 GeV ; ensuite, l'un des deux électrons, le "Tag", qui doit passer des coupures très serrées (Tight), sert à étiqueter le deuxième électron afin de réduire le bruit de fond ; l'autre électron, le "Probe", doit passer des coupures appropriées, dépendant de l'efficacité à déterminer ; enfin, l'efficacité ϵ est calculée par :

$$\epsilon = \frac{N_{pass}}{N_{pass} + N_{fail}}, \tag{5.1}$$

N_{pass} désigne le nombre des électrons probes qui ont passé les coupures appropriées de l'efficacité à déterminer, et N_{fail} correspond à ceux qui ont échoué.

5.2 Echantillons de données

Données : l'étude porte sur des échantillons de données du Run IIb (pour cette analyse, le Run IIb se compose du Run IIb1 et une partie du Run IIb2, cf. 2.2.2.3), enregistrés entre juin 2006 et février 2008, voir tableau 5.1. Les données ont été reconstruites avec le programme de reconstruction de DØ (d0reco), (cf. 2.2.5), adapté à la haute luminosité.
Cette analyse a été réalisée en utilisant les skims 2EMhightpt du Run IIb, (cf. 2.2.5). Ce dernier est formé d'évènements avec au moins deux objets électromagnétiques d'impulsion transverse supérieure à 12 GeV.

Bruit de fond multijet : le bruit de fond multijet (ou QCD) (cf. 7.4) est sélectionné à partir d'un autre échantillon de données, le skim QCD. Ce dernier comprend les évènements ayant au moins un jet d'impulsion transverse supérieure à 8 GeV ; ils doivent avoir déclenché au moins un des triggers simples ou doubles des jets, (les triggers simples et doubles des jets requièrent respectivement un et deux jets dans les évènements). Les échantillons de données utilisés pour sélectionner le bruit de fond multijet sont regroupés dans le tableau 5.2.

L'ensemble des échantillons de données utilisés pour cette analyse sont sous la forme de fichiers CAF-trees, et ils ont été analysés dans CAFe, (cf. 2.2.5).

Période	Skim
Run IIb2	CSG_CAF_QCD_PASS4_p21.10.00_p20.12.00
(octobre 2007 - février 2008)	CSG_CAF_QCD_PASS4_p21.10.00_p20.12.01
	CSG_CAF_QCD_PASS4_p21.10.00_p20.12.02

TAB. 5.2 – Skims utilisés pour les sélection du bruit de fond multijet, pour l'analyse du Run IIb.

5.3 Sélection des évènements

On sélectionne les évènements ayant au moins un électron candidat Tag et un électron candidat Probe. L'électron candidat Tag doit passer les coupures d'identification très sévères suivantes :

- avoir une impulsion transverse $p_T^{EM} > 25$ GeV ;
- avoir un type ID = 10 ou ± 11 ; le type ID permet de sélectionner des candidats issus de l'algorithme de cône simple décrit dans la section 3.3. Il vaut 10 si le candidat est un photon, et +11 (-11) si c'est un positron (électron) ;
- son isolation doit être Iso < 0.07 ;
- la fraction d'énergie qu'il a déposée dans le calorimètre doit être, EMfrac $>$ 0.97 ;
- le candidat électron Tag doit avoir une bonne forme de gerbe dans la partie centrale de calorimètre, H-Matrix7 < 20, (cf. section 3.3) ;
- IsoHC4 < 2 GeV, c'est l'isolation des traces des objets électromagnétiques. Elle est définie par la somme de l'impulsion transverse des traces p_T^{traces}, (avec $p_T^{traces} > 0.5$ GeV), dans un cône de rayon $0.05 < R < 0.4$ autour de la direction de l'objet électromagnétique, voir fig. 5.2 ;
- LHood8 > 0.8, likelihood ou maximum de vraisemblance [68]. C'est un estimateur construit à partir de 8 variables, comme par exemple l'impulsion transverse totale des traces dans un cône de rayon 0.4 autour de la trace du candidat, le rapport entre l'énergie transverse du candidat et l'impulsion transverse des traces, ou encore les H-Matrix à 7 (CC) et 8 (EC) variables. Le Likelihood est utilisé pour la sélection des électrons seulement, car il utilise les propriétés de la trace associée au candidat (les photons ne laissent pas de traces dans le trajectographe). En général on demande Likelihood > 0.20 pour une identification lâche, et Likelihood > 0.8 pour une sélection plus sévère, voir fig. 5.3 ;
- NN-elec7 > 0.7 [69], est la variable de sortie d'un réseau de neurones (cf. 7.7.1) qui permet de sélectionner les électrons dans la partie centrale du calorimètre. Il est construit à partir des 7 variables suivantes : la fraction d'énergie déposée dans la première couche du calorimètre électromagnétique (EM1) ; le nombre de cellules touchées dans la couche EM1 dans un cône de rayon 0.2 ; l'isolation

des traces (IsoHC4) ; le nombre de cellules de la couche EM1 dans un cône de rayon compris entre 0.2 et 0.4 autour de la direction du candidat ; le nombre de traces dans un cône de rayon 0.05 ; le nombre d'amas dans le CPS dans un cône rayon 0.1 ; et une combinaison du nombre de coups et de l'énergie mesurée dans le CPS. La variable de sortie de réseau de neurones NN-elec7 est un nombre compris entre 1 (si c'est du signal) et 0 (si c'est du bruit de fond), voir fig. 5.4.
La variable NN-elec3 à la différence du NN-elec7 est utilisée pour identifier les électrons dans les parties avant du calorimètre. Elle est construite à partir des trois variables qui sont le nombre de cellules touchées dans la couche EM1 dans un cône de rayon 0.2, l'IsoHC4 et la H-Matrix8.

L'électron candidat Probe doit avoir les spécificités suivantes :

- $p_T^{EM} > 10$ GeV ;
- ID $= 10$ ou ± 11 ;
- Iso < 0.15 ;
- EMfrac > 0.90 ;
- IsoHC4 < 6 GeV.

Les électrons candidats Tag et Probe doivent se trouver dans le calorimètre central (CC) défini par $|\eta_{det}| < 1.1$ ou bien dans la région avant du calorimètre délimitée par $1.5 < |\eta_{det}| < 3.2$. Notons que l'efficacité de cette pré-sélection est pratiquement de 100% . Enfin, pour réduire le bruit de fond multijet, la distribution de la masse invariante des ces deux électrons candidats doit être consistante avec celle de la masse du boson Z, c'est-à-dire quelle doit être comprise entre 80 GeV et 100 GeV.

5.4 Les évènements simulés

Les évènements du signal $Z \to e^+e^-$ ont été simulés dans l'intervalle de masse compris entre 60 et 130 GeV avec le générateur d'évènement PYTHIA (cf. 7.3.1). L'échantillon d'évènements simulés utilisé est référencé dans la base de données par CSG_CAF_MCv4-55052_p21.11.00. Par la suite, ces évènements simulés sont passés par la même chaîne de reconstruction d0reco que les données (cf. 2.2.5). Enfin, pour éviter tous biais, les évènements simulés ont été sélectionnés avec les mêmes coupures d'identification que les données, (cf. 5.3).

5.5 Le bruit de fond multijet

Les évènements constituants le bruit de fond multijet doivent avoir les caractéristiques suivantes :

- avoir au moins un candidat électron Probe ;

- le candidat Probe doit être distant de tous les "bons" jets (cf. 3.4), de sorte que $\Delta R(\eta_{det}, \phi) > 0.7$;
- avoir au moins un bon jet d'impulsion transverse $p_T > 15$ GeV (après correction de l'énergie des jets, cf. 3.4), qui soit distant du candidat électron Probe de $\Delta\phi > 2.8$;
- avoir une énergie transverse manquante $\not{E}_T < 10$ GeV ;
- la reconstruction de la masse invariante du candidat électron Probe avec celle de n'importe qu'elle autre candidat électromagnétique (s'il existe) doit être inférieure (supérieure) à 60 (130) GeV.

5.6 Identification des électrons

5.6.1 Critères de sélection

Les variables choisies pour construire les différents ensembles de définition des électrons ont l'avantage de peu dépendre de la luminosité, et elles ont un bon rapport signal sur bruit de fond. Il existe six qualités (ou définitions) d'électrons, de la plus lâche à la plus serrée elles sont nommées : VLoose, Loose, MLoose1, MLoose2, Medium et Tight. Chaque qualité requiert des critères différents qui doivent tous être satisfaits. Ces critères sont résumés dans le tableau 5.3 (5.4) pour la partie centrale (les bouchons) du calorimètre. La majorité des variables listées ont été définies dans les sections 3.3 et 5.3. Les variables TrkM [70] et Hit-Road [71] sont utilisées en association dans la partie centrale du calorimètre, elle sont définies ci-dessous :

- la TrkM est la probabilité spatiale que la trace soit associée à un candidat électromagnétique, voir fig. 5.5, elle est définie à partir de :

$$\chi^2 = \left(\frac{\Delta\eta_{det}}{\sigma_{\eta_{det}}}\right)^2 + \left(\frac{\Delta\phi}{\sigma_\phi}\right)^2, \tag{5.2}$$

  $\Delta\eta_{det}$ et $\Delta\phi$ sont respectivement les différences de position en η_{det} et en ϕ entre la trace et l'objet électromagnétique, $\sigma_{\eta_{det}}$ et σ_ϕ sont les écarts types respectifs ;

- la variable Hit-Road est un likelihood qui utilise les coups laissés le long de la trajectoire des particules dans le trajectographe, et l'énergie transverse de la particule, pour mieux identifier les électrons, les photons et le bruit de fond. En effet, les photons ne laissent pas de trace dans le trajectographe, et la trace laissée par les électrons est différente de celle laissée par les bruits de fond. Enfin, la variable Hit-Road, comprise entre 0 et 1, sera d'autant plus élevée que l'objet ressemblera à un électron, voir fig. 5.6 ;
- la variable SigPhi correspond à la largeur en cm de l'objet électromagnétique dans la troisième couche du calorimètre (EM3), dans le plan (r, ϕ), voir fig. 5.7.

En général, les définitions sont construites de façon de plus en plus stricte en resserrant les coupures déjà existantes, ou en ajoutant de nouvelles coupures sur d'autres

Variables	VLoose	Loose	MLoose1	MLoose2	Medium	Tight
Iso <	0.10	0.10	0.07	0.07	0.07	0.07
EMfr >	0.95	0.95	0.97	0.97	0.97	0.97
IsoHC4 <	3.5	3.0	2.5	2.5	2.5	2.5
H-Matrix7 <	35	35	25	25	25	25
NN-elec7 >	—	0.2	0.2	0.6	0.6	0.6
LHood8 >	—	—	—	—	0.2	0.8
TrkM >	—	0.001	0.001	0.001	0.001	0.001
ou Hit-Road >	—	0.4	0.4	0.5	0.5	0.5

TAB. 5.3 – Définitions des électrons dans la région centrale du calorimètre.

Variables	VLoose	Loose	MLoose1	MLoose2	Medium	Tight
Iso <	0.10	0.10	0.10	0.07	0.05	0.07
EMfr >	0.95	0.95	0.95	0.95	0.97	0.97
IsoHC4 <	3.5	2.0	2.0	2.0	1.0	2.0
H-Matrix8 <	35	20	15	15	10	15
NN-elec3 >	—	0.4	0.4	0.4	0.4	0.4
TrkM >	—	—	—	—	—	0.0
SigPhi <	—	—	20.0	20.0	20.0	20.0

TAB. 5.4 – Définitions des électrons dans les bouchons du calorimètre.

variables. Il n'y a pas de variation trop brutale entre les efficacités des définitions adjacentes, elle est généralement comprise entre 4 et 7%. Seules les définitions Medium et Tight dans les bouchons du calorimètre ont une variation plus importante, car l'efficacité du signal baisse plus fortement, entre 15 et 30%, à cause de la faible efficacité de reconstruction des traces dans cette région. En général les analyses privilégient une l'efficacité aussi élevée que possible, par conséquent, seule la coupure Tight, la plus sévère, demande une trace associée au candidat électromagnétique (TrkM > 0.0) dans les bouchons du calorimètre.

5.6.2 Efficacités d'identification

La figure 5.8 montre des exemples de distributions des efficacités d'identification des électrons en fonction de leur impulsions transverses p_T, et des angles η_{det} et ϕ pour différentes définitions et dans les différentes parties du calorimètre.

5.6.3 Correction des efficacités

En général, la simulation ne reproduit pas parfaitement la réalité, et donc la distribution des évènements simulés est légèrement différente de celle des données. Pour que la simulation décrive au mieux la réalité, on corrige la simulation avec

des facteurs de correction (noté SF pour Scale Factor). Le SF est le rapport entre l'efficacité des données et celle de la simulation dans chaque bin i tel que :

$$SF_i = \epsilon_i^{Data}/\epsilon_i^{MC}. \tag{5.3}$$

SF_i est calculé pour chaque définition en fonction de p_T^{EM}, η_{det}, et ϕ_{det}, dans les CC et EC du calorimètre. Il tient compte des incertitudes statistiques et systématiques liées aux données et à la simulation. La figure 5.9 montre deux exemples de distributions des facteurs de correction obtenus en fonction de l'impulsion transverse p_T, pour les deux définitions VLoose et Tight dans les CC et EC du calorimètre.

5.7 Conclusion

L'augmentation de la luminosité entre le RunIIa et le RunIIb a nécessité de nombreux changements, notamment les critères de sélection et d'identification des électrons dans le calorimètre ont dû être revus, l'enjeu étant de garder un rapport signal sur bruit de fond aussi élevé que possible. Dans ce chapitre, les différentes étapes de la certification des électrons ont été présentées : les coupures d'identification des électrons pour le Run IIb ont été définies pour la partie centrale et les bouchons du calorimètre ; des facteurs de correction ont par la suite été calculés pour chaque définition, et pour chaque région donnée du calorimètre. Ces coupures certifiées ont servi à l'analyse de cette thèse.

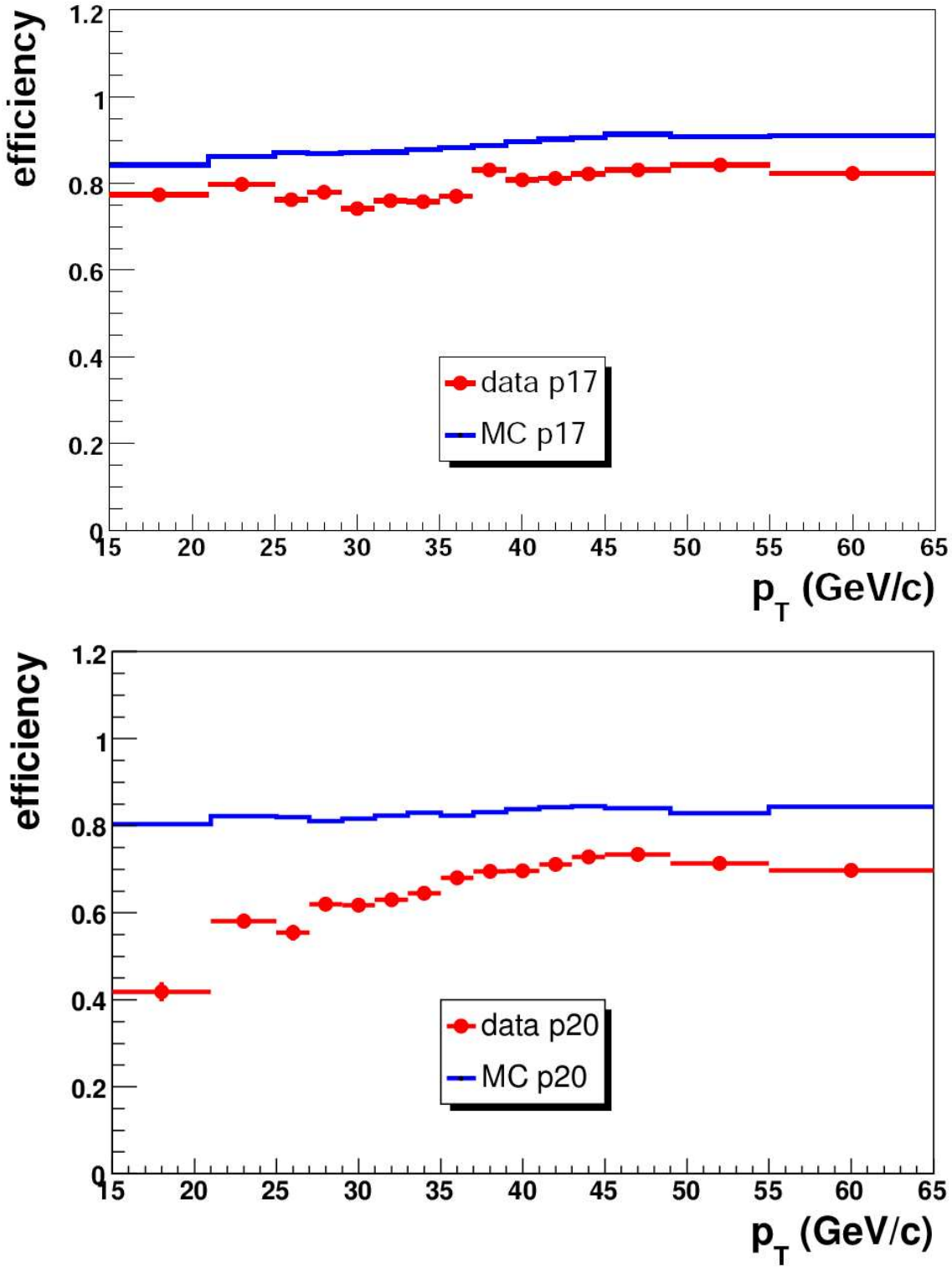


FIG. 5.1 – Distribution des efficacités d'identification des électrons (efficiency) en fonction de l'impulsion transverse (p_T). En haut, la coupure Tight-trk (moyennement serrée) définie pour les données du Run IIa (p17). En bas, la même coupure Tight-trk sur les données du Run IIb (p20). Les points (rouges) représentent les données et le trait (bleu) la simulation.

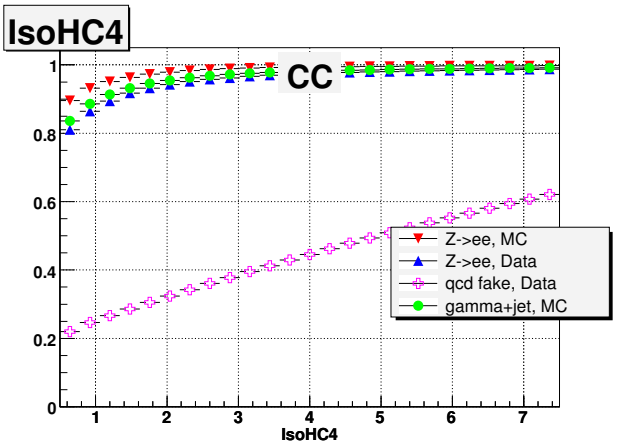


FIG. 5.2 – Distribution de la variable IsoHC4 dans le calorimètre central.

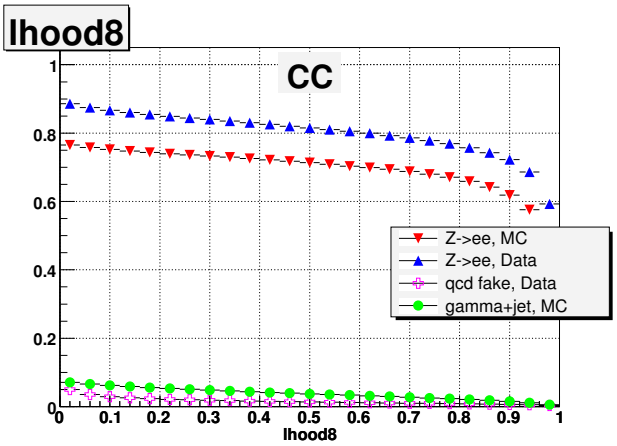


FIG. 5.3 – Distribution de la variable LHood8 dans le calorimètre central.

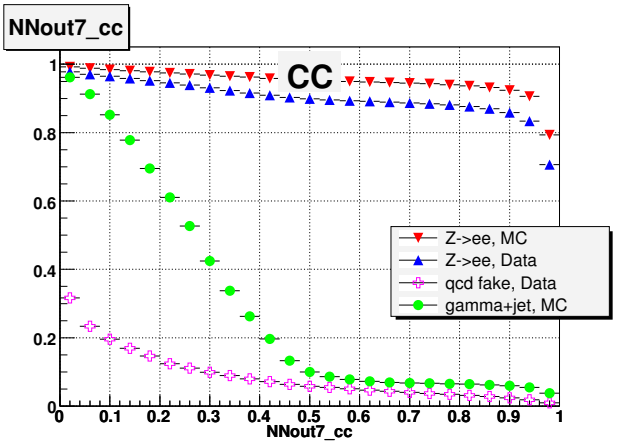


FIG. 5.4 – Distribution de la variable NN-elec7 dans le calorimètre central.

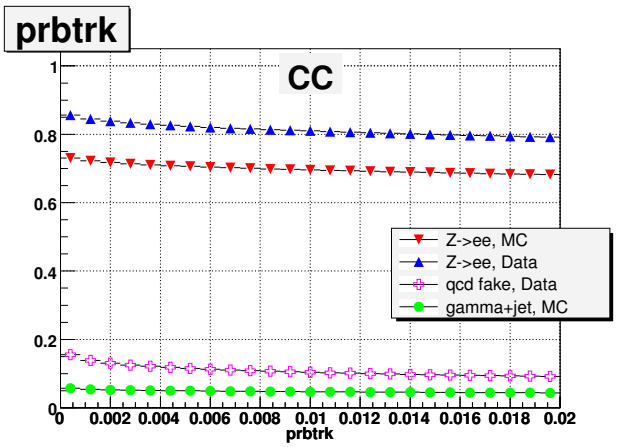


FIG. 5.5 – Distribution de la variable TrkM dans le calorimètre central.

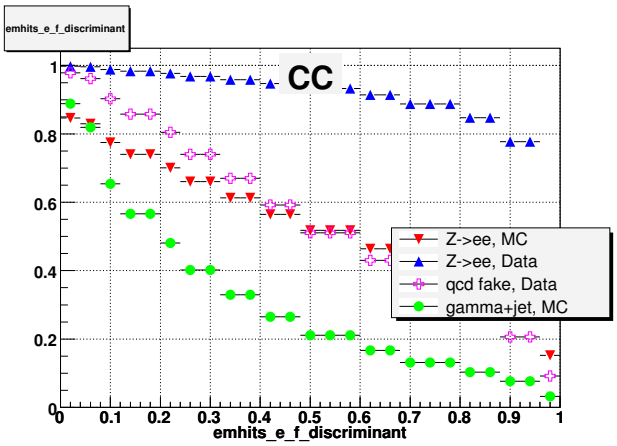


FIG. 5.6 – Distribution de la variable Hit-Road dans le calorimètre central.

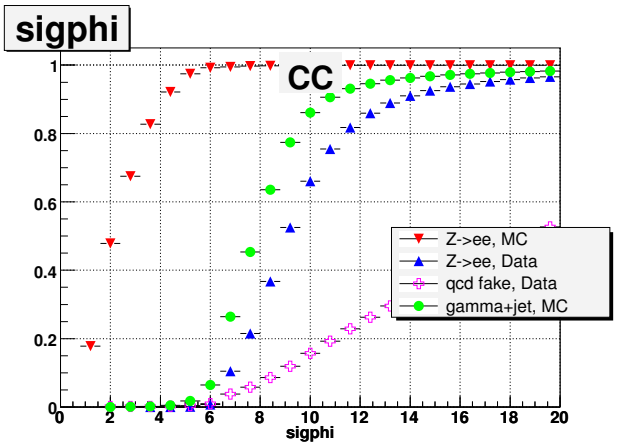
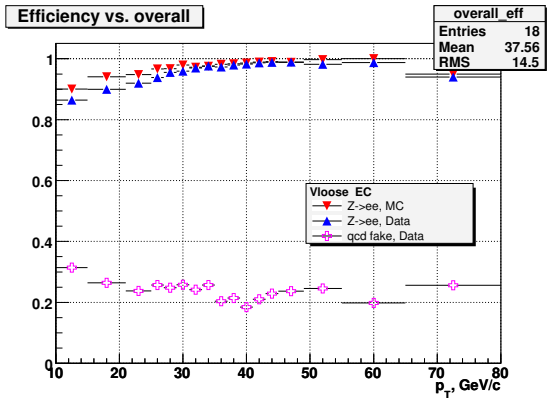
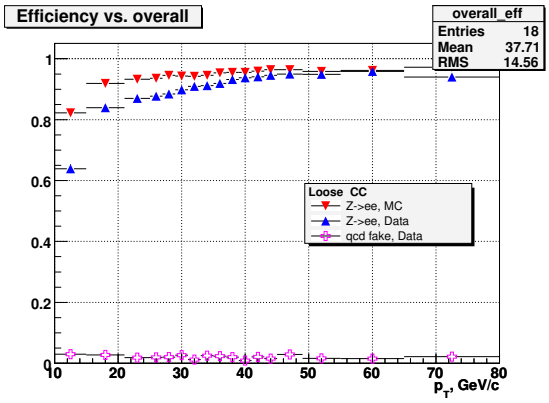


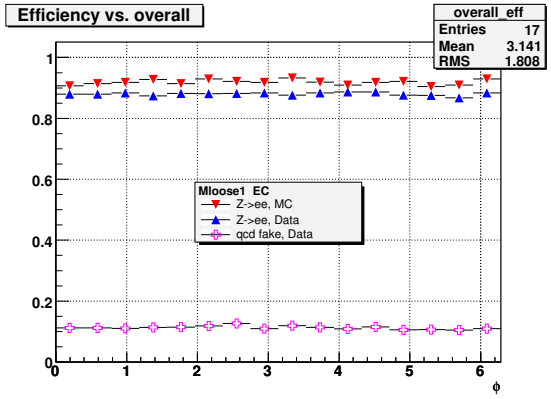
FIG. 5.7 – Distribution de la variable SigPhi dans le calorimètre central.



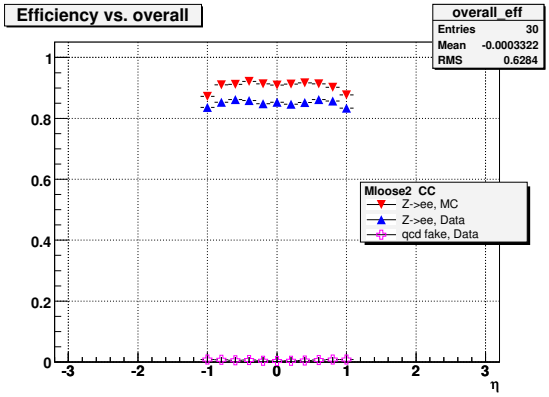
(a) p_T VLoose EC



(b) p_T Loose CC

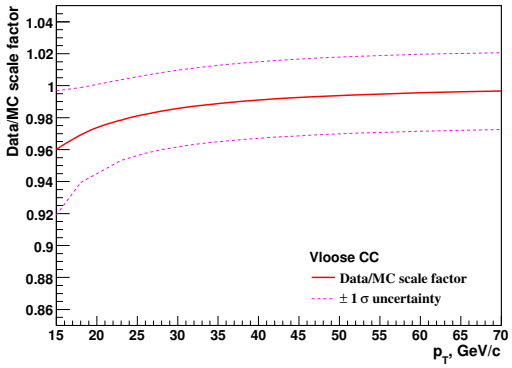


(c) ϕ MLoose1 EC

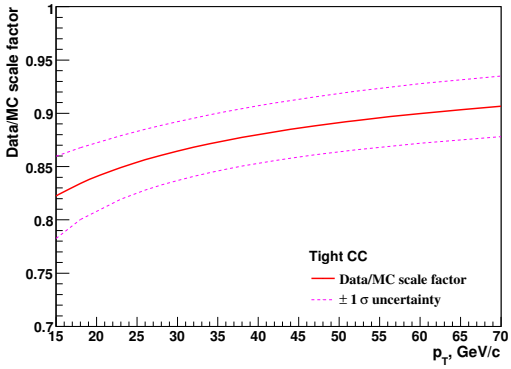


(d) η_{det} MLoose2 CC

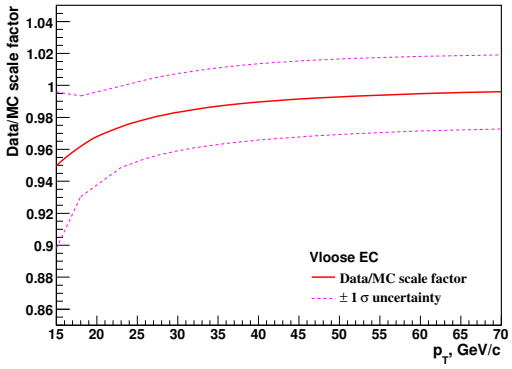
FIG. 5.8 – Efficacité d'identification des électrons en fonction de : 5.8(a) (5.8(b)) l'impulsion transverse pour la définition VLoose (Loose) dans le EC (CC) ; et de l'angle ϕ 5.8(c) (η_{det} 5.8(d)) pour la définition MLoose1 (MLoose2) dans le EC (CC).



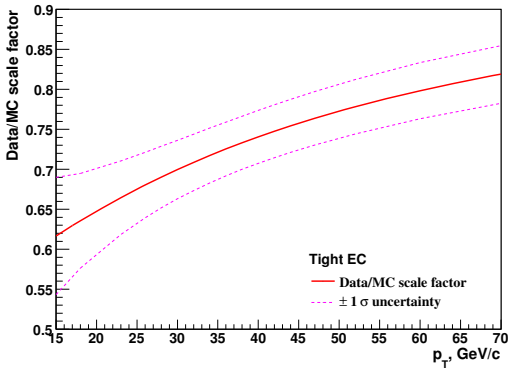
(a) SF VLoose CC



(b) SF Tight CC



(c) SF VLoose EC



(d) SF Tight EC

FIG. 5.9 – Rapport des efficacités Data/MC (trait plein), et les incertitudes (pointillés), en fonction du p_T^{EM}, pour les définitions 5.9(a) VLoose et 5.9(b) Tight dans le CC, et pour les définitions 5.9(c) VLoose et 5.9(d) Tight dans le EC.

Chapitre 6

Certification de l'identification des électrons dans l'ICR

Contents

6.1 Introduction **84**
6.2 Les objets taus **85**
6.3 Echantillon de données **86**
6.4 Sélection des évènements **86**
6.5 Les évènements simulés **87**
6.6 Efficacités **87**
6.6.1 Efficacité du trajectographe 88
6.6.2 Efficacité de reconstruction des objets taus 88
6.6.3 Efficacité du réseau de neurones tau 90
6.6.4 Efficacité totale 92
6.7 Bruit de fond multijet **92**
6.8 Facteurs de correction **92**
6.9 Conclusion **93**

Ce chapitre, consacré à la certification de l'identification des électrons dans la région inter-cryostat (ICR), est un complément du précédent chapitre. L'ajout des électrons se trouvant dans la région inter-cryostat permet en moyenne d'augmenter l'efficacité du signal de près de 17%. Nous verrons dans les prochaines pages, les différentes étapes de la certification de l'identification des électrons dans cette région du détecteur.

6.1 Introduction

Le détecteur DØ a un bon calorimètre doté d'une excellente herméticité dans sa partie centrale (CC) et ses bouchons (EC). Cependant, la région inter-cryostat (ICR) du détecteur définie par $1.1 \leq |\eta_{det}| \leq 1.5$, à l'exception des bords (c'est-à-dire la région $1.1 \leq |\eta_{det}| \leq 1.2$ et $1.4 \leq |\eta_{det}| \leq 1.5$), n'est pas couverte par le calorimètre électromagnétique et est très peu instrumentée. Ces lacunes entraînent

une dégradation de la mesure de l'énergie dans l'ICR, cf. 2.2.3.3. Le détecteur inter-cryostat (ICD) permet de combler l'absence des couches électromagnétiques et donc de limiter cette dégradation. Le but de cette analyse est de déterminer l'efficacité de reconstruction des électrons dans l'ICR, ainsi que les facteurs de corrections. L'objectif étant d'inclure les électrons de l'ICR dans les analyses, pour augmenter la statistique. Ces résultats seront utilisés par la suite pour la recherche du Higgs de cette thèse.

6.2 Les objets taus

Comme mentionné précédemment, la région inter-cryostat n'est pratiquement pas couverte par le calorimètre électromagnétique, mais elle l'est entièrement par le trajectographe (cf. 2.2.3.1). Lors de leur passage dans l'ICR, les électrons ne déposent donc quasiment pas d'énergie dans le calorimètre électromagnétique, ce qui complique leur identification par les méthodes habituelles, basées sur l'énergie laissée dans le calorimètre électromagnétique. En revanche, les électrons laissent une trace dans le trajectographe, et déposent une partie de leur énergie dans l'ICD. Ces informations peuvent donc être utilisées pour la reconstruction des électrons dans cette région du détecteur. Les objets taus présentent de nombreuses ressemblances avec les objets électromagnétiques dans l'ICR, à savoir qu'ils laissent une (ou plusieurs) trace(s) dans le trajectographe et peuvent (ou pratiquement pas) déposer de l'énergie dans l'ICD. L'idée est donc d'adapter les coupures de sélection des objets taus, déjà existantes, aux objets électromagnétiques. Il convient dans un premier temps de définir les objets tau.

Les objets taus [72] sont les leptons candidats taus (τ), ils ont une durée de vie très courte de l'ordre de 290.10^{-15} s, et se désintègrent donc très vite dans le détecteur en laissant une ou plusieurs traces dans le trajectographe, et éventuellement un dépôt d'énergie dans le calorimètre. Dans l'ICR, on peut trouver trois types de taus, cf. fig. 6.1 :

- le tau de type 1 se désintègre en un $\pi^{\pm}$. Il ne laisse pas d'énergie dans le calorimètre électromagnétique, mais laisse une trace dans le trajectographe. Ce type de taus est trouvé dans le centre de l'ICR, là où les couches électromagnétiques sont inexistantes ;
- le tau de type 2 se désintègre en un $\pi^{\pm}$ avec un ou plusieurs π^0. Il laisse un dépôt d'énergie dans le calorimètre électromagnétique, et une trace dans le trajectographe. Ce type de taus est trouvé au bord de l'ICR, là où les couches électromagnétiques sont limitées ;
- le tau de type 3 se désintègre en plusieurs $\pi^{\pm}$ avec un ou plusieurs π^0. Il peut ou pas laisser un dépôt d'énergie dans le calorimètre électromagnétique, selon qu'il se trouve respectivement au bord ou au centre de l'ICR, et il laisse plusieurs traces dans le trajectographe.

Les candidats taus sont reconstruits à partir de deux algorithmes [73], le premier identifie les amas du calorimètre à l'aide de l'algorithme de cône simple, (cf. 3.3), et le deuxième utilise les traces de hautes impulsions transverses pour former les

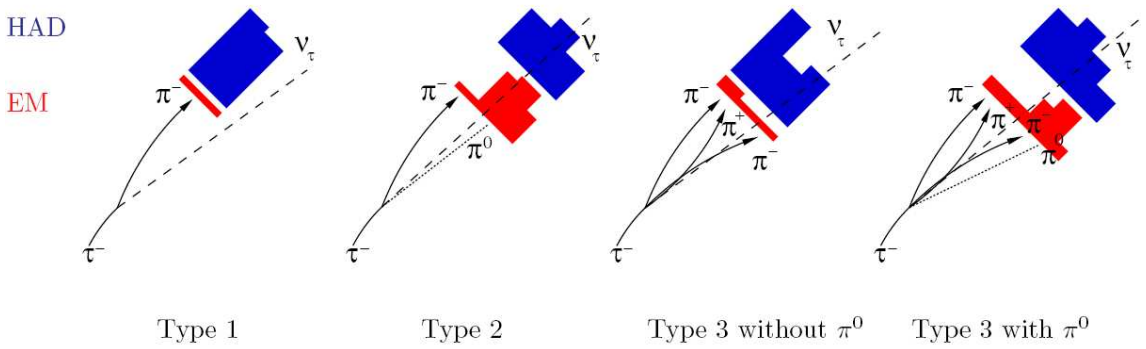


FIG. 6.1 – Caractéristiques des trois types de tau : le tau de type 1 se désintègre en $\tau \rightarrow \pi^{\pm}\nu_{\tau}$, il a donc une trace associée et il ne dépose pas d'énergie dans les couches électromagnétiques du calorimètre (EM en rouge) ; le tau de type 2 se désintègre en $\tau \rightarrow \pi^{\pm}$ n$\pi^0\nu_{\tau}$ où n indique le nombre de π^0, ce type de tau possède donc une trace et dépose de l'énergie dans les couches EM ; le tau de type 3 se désintègre en $\tau \rightarrow \pi^{\pm}\pi^{\mp}\pi^{\pm}\nu_{\tau}$ nπ^0 avec n$\geq$ 0, il possède donc au moins deux traces et laisse éventuellement un dépôt d'énergie dans les couches EM.

candidats taus. Puis les taus sont sélectionnés à partir des candidats taus, à l'aide de coupures de sélection appropriées.

6.3 Echantillon de données

L'étude porte sur les données du Run IIa enregistrées entre avril 2002 et février 2006, et reconstruites avec le programme de reconstruction de DØ (d0reco). Les données ont été sélectionnées à partir du skim EMinclusive, et elles sont présentées sous la forme de fichiers CAF-trees (cf. 2.2.5). Le skim EMinclusive utilisé est référencé par "CSG_CAF_EMinclusive_PASS3_p18.14.00" dans la base de donnée. La luminosité intégrée associée, une fois que le trigger JT_125TT (ce trigger requiert au moins un jet d'énergie transverse supérieure à 125 GeV) a été demandé, est de 1.074 ± 0.066 fb^{-1}. L'analyse a été effectuée dans le cadre de travail de CAFe.

6.4 Sélection des évènements

La méthode "Tag and Probe" a été utilisée pour déterminer les efficacités, voir section 5.1. On sélectionne donc les évènements ayant au moins un électron candidat Tag et un électron candidat Probe. Le candidat Tag est sélectionné dans les régions CC et EC du calorimètre, avec les coupures très serrées (Tight) suivantes :

- Iso > 0.07 ;
- EMfr > 0.97 ;
- H-Matrix7 < 25 dans le CC, et H-Matrix8 < 15 dans les EC ;

- IsoHC4 < 2.5 dans le CC, et IsoHC4 < 2.0 dans les EC ;
- TrkM > 0.0 ou Hit-Road > 0.5, dans le CC ;
- TrkM > 0.0 dans les EC ;
- NN-elec7 > 0.6 dans le CC, et NN-elec3 > 0.4 dans les EC ;
- SigPhi < 20 cm dans les EC ;
- LHood8 > 0.8 ;
- E/p < 4 dans le CC. E et p sont respectivement l'énergie déposée dans le calorimètre, et l'impulsion transverse de la trace du candidat.

Le candidat Probe est sélectionné dans la région ICR, il doit passer les coupures d'identification adaptées à chaque efficacité à déterminer. L'efficacité de reconstruction des électrons dans l'ICR est le produit de trois efficacités différentes. Ces dernières ainsi que les coupures d'identification requises pour chaque candidat tau, sont présentées dans la section 6.6.

6.5 Les évènements simulés

Des évènements Z+jets légers, Z+jets b et Z+jets c ont été simulés dans l'intervalle de masse compris entre 15 GeV et 250 GeV, avec le générateur d'évènement ALPGEN interfacé à PYTHIA, (cf. 7.3.1), puis ils sont passés par la même chaîne de reconstruction (d0reco) que les données, (cf 2.2.5). Comme pour les données, les efficacités ont été déterminées avec la méthode "Tag and Probe". L'électron candidat Tag doit passer les mêmes coupures de sélection Tight que les données listées en 6.4, et l'électron candidat Probe doit comme les données, passer les mêmes coupures d'identification adaptées à chaque efficacité à déterminer, et qui sont mentionnées en 6.6.

La simulation ne reproduisant pas parfaitement la réalité, des corrections lui sont appliquées afin qu'elle décrive au mieux les données (cf. 7.3.2) : l'efficacité d'identification de l'électron candidat Tag a été corrigée comme conseillé en [74] ; l'impulsion transverse du boson Z a été corrigée, et le vertex primaire du boson Z, ainsi que le profil de la luminosité instantanée ont été repondérés. Par ailleurs les échantillons Z+jets légers, Z+jets b et Z+jets c ont subi un traitement de "filtrage des saveurs lourdes" (cf. 7.3.1). Enfin, la section efficace de production des différents processus a été corrigée. Le tableau 7.2 montre les échantillons de simulation utilisés.

6.6 Efficacités

Le calcul de l'efficacité de reconstruction des électrons dans l'ICR nécessite la détermination des trois efficacités décrites ci-dessous. Les efficacités ont été calculées avec la méthode Tag and Probe. Le candidat électron Tag doit passer des coupures Tight (cf. 6.4), et le candidat électron Probe doit passer des coupures adéquates

aux différentes efficacités à calculer, et qui sont expliquées dans les prochaines sous-sections.

6.6.1 Efficacité du trajectographe

L'efficacité du trajectographe est la probabilité qu'un candidat tau se trouvant dans l'ICR, ait produit une trace dans le trajectographe. Les évènements entrant dans le dénominateur doivent avoir un électron de qualité Tight et un candidat électron Probe, de sorte que la distribution de la masse invariante de ces deux électrons soit compatible avec celle de la masse du boson Z, c'est-à-dire comprise entre 70 GeV et 100 GeV. L'électron Probe doit passer les coupures de sélections suivantes :

- avoir une énergie transverse $E_T > 10$ GeV ;
- être dans la région ICR $1.1 \leq |\eta_{det}| \leq 1.5$;
- NN-L3Tau [75] est la variable de sortie d'un réseau de neurones qui permet de réduire le bruit de fond. Elle est construite à partir de cinq variables comme l'énergie transverse de la particule, ou l'énergie déposée dans les couches électromagnétique du calorimètre. La coupure faite sur cette variable dépend de la région η_{det}, et elle a été optimisée de sorte que l'efficacité de sélection du signal soit supérieure ou égale à 95% ;
- LHood-T3 > 0.03, c'est un likelihood défini pour les objets taus. Il permet de réduire le bruit de fond en gardant environ 95% de l'efficacité du signal. Il est calculé à partir de la fraction d'énergie déposée dans l'ICD et dans les couches électromagnétiques du calorimètre en fonction de η_{det} ;
- IsoHC4 < 3.0 GeV, cf. 5.3, permet également de réduire le bruit de fond.

Les évènements entrant dans le numérateur du calcul de l'efficacité du trajectographe doivent avoir un électron candidat Tag et un électron candidat Probe, respectant tous les critères cités ci-dessus. Les évènements doivent en plus être associés à au moins une trace passant les coupures suivantes :

- L'impulsion transverse de la trace doit être $p_T^{trace} > 20$ GeV ;
- l'extrapolation de la trace dans le calorimètre doit se trouver dans la région ICR $1.1 \leq |\eta_{det}^{trace}| \leq 1.5$;
- la trace doit avoir produit au moins un coup dans le SMT.

L'efficacité du trajectographe en fonction de la luminosité, de l'impulsion transverse, et des angles η_{det} et ϕ, est représentée à la figure 6.2. En moyenne, l'efficacité du trajectographe est de 0.735 $\pm$0.003 (0.785 $\pm$ 0.003) pour les données (la simulation).

6.6.2 Efficacité de reconstruction des objets taus

L'efficacité de reconstruction des objets taus est la probabilité qu'un candidat tau se trouvant dans l'ICR, et ayant produit une trace dans le trajectographe, passe les

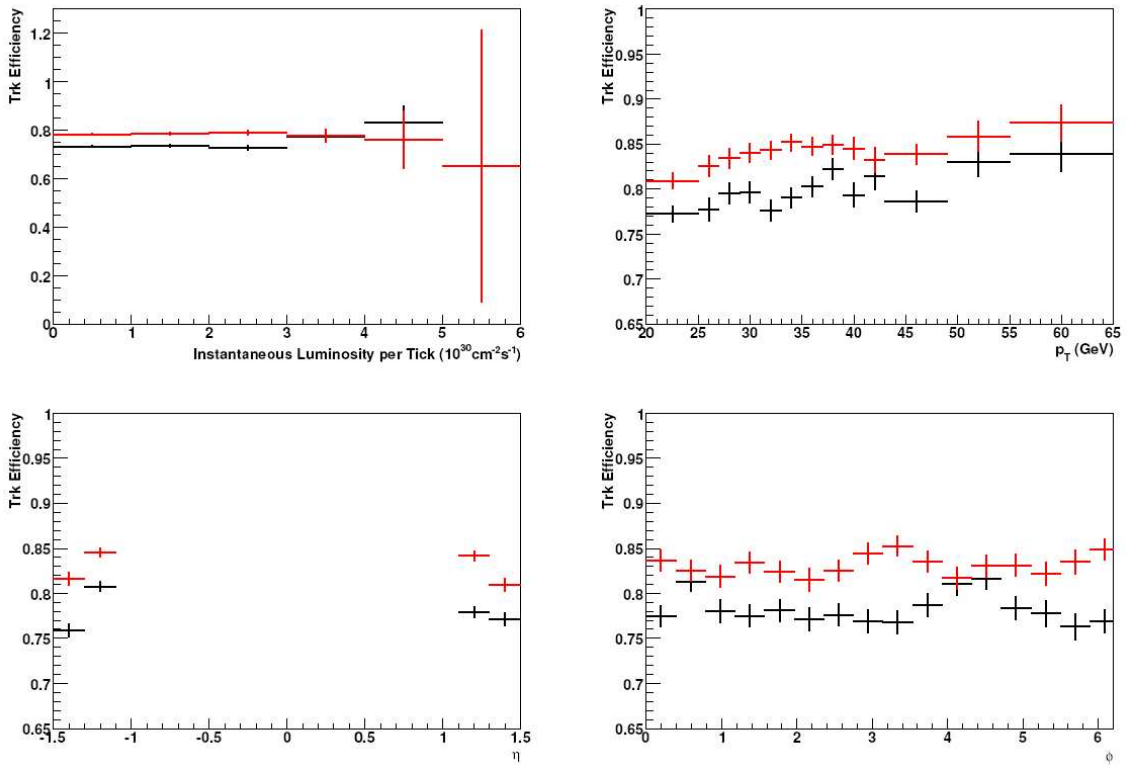


FIG. 6.2 – Efficacité du trajectographe (Trk Efficiency) en fonction de la luminosité instantanée (en haut à gauche), de l'impulsion transverse (en haut à droite), et des angles η_{det} (en bas à gauche) et ϕ (en bas à droite). Les croix noires représentent les données et les croix rouges les évènements simulés.

critères de sélection des objets taus. Les évènements entrant dans le dénominateur, pour le calcul de cette efficacité doivent avoir un candidat électron Tag de qualité Tight (cf. 6.4), et un candidat électron Probe associé à une trace. Cette dernière doit passer les mêmes coupures que celles de la trace listées ci-dessus, et en plus elle doit respecter les autres critères suivants :

- la trace doit être isolée, IsoHC4 < 3.0 GeV ;
- la distance entre la trace et l'électron Tag doit être $\Delta\phi(e^{Tag}, \text{trace}) > 170°$;
- la reconstruction de la masse invariante de la trace et de l'électron Tag doit être comprise entre 70 et 110 GeV ;
- la trace et l'électron Tag doivent être de charges opposées.

Les évènements entrant dans le numérateur, doivent avoir un candidat électron Probe associé à la trace, et qui respecte les conditions suivantes :

- être dans la région ICR $1.1 \leq |\eta_{det}| \leq 1.5$;

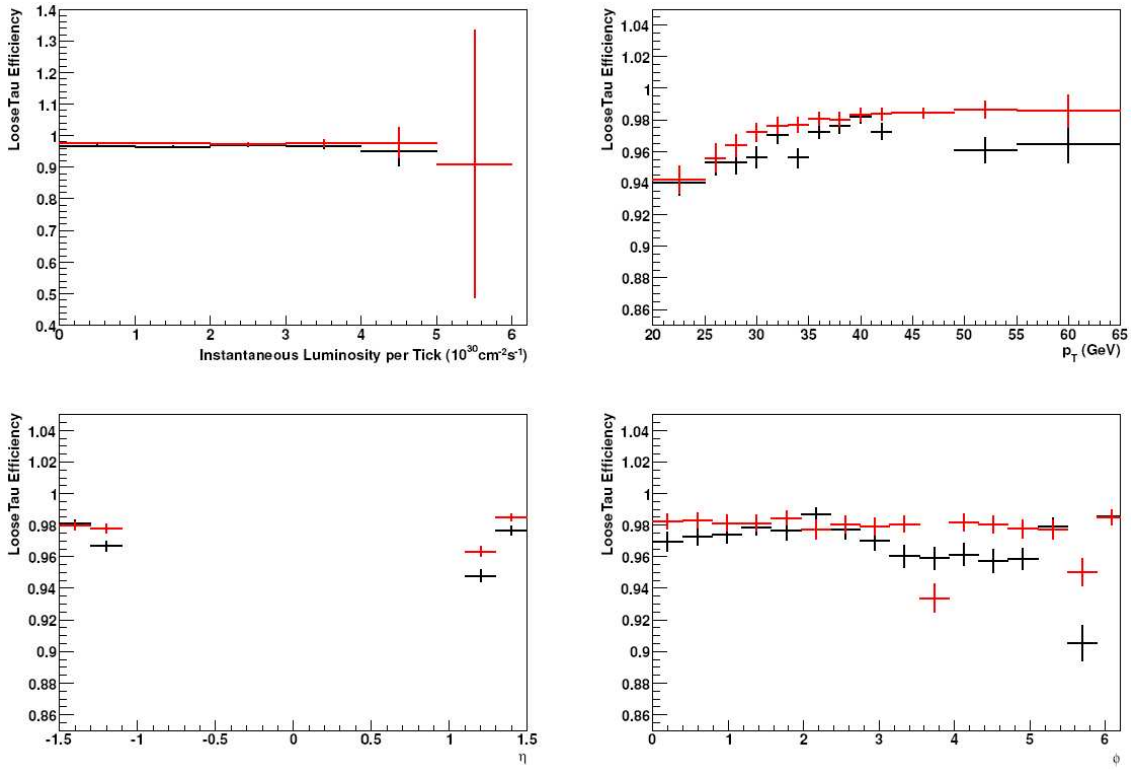


FIG. 6.3 – Efficacité de reconstruction des objets taus en fonction de la luminosité instantanée (en haut à gauche), de l'impulsion transverse (en haut à droite), et des angles η_{det} (en bas à gauche) et ϕ (en bas à droite). Les croix noires représentent les données et les croix rouges les évènements simulés.

- avoir une énergie transverse $E_T > 10$ GeV ;
- les candidats électron Probe reconstruits comme des taus de type 1, doivent avoir déposé une fraction d'énergie supérieure à 0.1 dans l'ICD ;
- les candidats électron Probe reconstruits comme des taus de type 2, doivent avoir un likelihood LHood-T3 > 0.15.

L'efficacité de reconstruction des objets taus en fonction de la luminosité, de l'impulsion transverse, et des angles η_{det} et ϕ, est représentée à la figure 6.3. En moyenne, l'efficacité de reconstruction des objets taus est de 0.967 $\pm$0.002 (0.974 $\pm$ 0.002) pour les données (la simulation).

6.6.3 Efficacité du réseau de neurones tau

L'efficacité du réseau de neurones tau, ou NN-τ, est la probabilité qu'un candidat tau se trouvant dans l'ICR, et ayant produit une trace dans le trajectographe, passe les critères de sélection des objets taus et également une des trois coupures (Loose, Medium et Tight) du NN-τ. Les évènements entrant dans le dénominateur doivent avoir un candidat électron Tag de qualité Tight (cf. 6.4), et un candidat

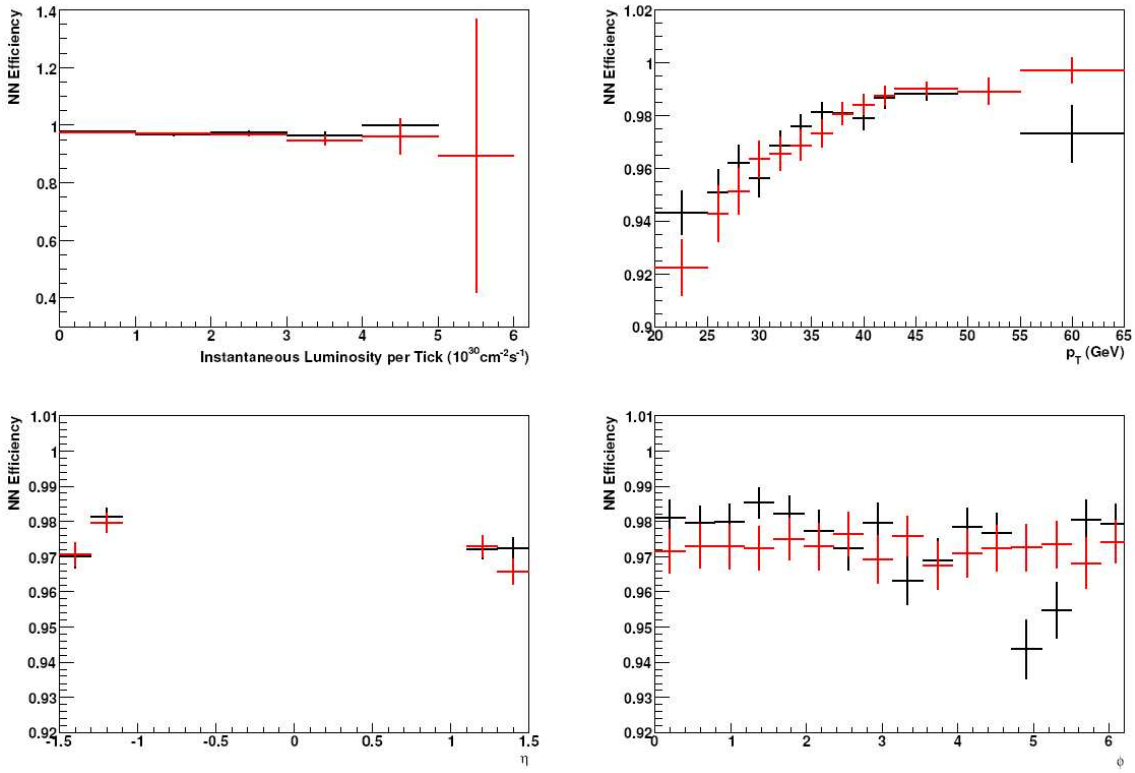


FIG. 6.4 – Efficacité du réseau de neurones tau pour la coupure intermédiaire (NN-τ = 0.7), en fonction de la luminosité (en haut à gauche), de l'impulsion transverse (en haut à droite), et des angles η_{det} (en bas à gauche), et ϕ (en bas à droite). Les croix noires représentent les données et les croix rouges les évènements simulés.

Probe ayant passé les mêmes coupures listées en 6.6.2. Les évènements entrant dans le numérateur, doivent avoir un candidat électron Probe ayant passé une des coupures du NN-τ [76]. Il y a trois qualités de coupures NN-τ : la coupure lâche ou Loose pour NN-$\tau = 0.2$; la coupure intermédiaire ou Medium pour NN-$\tau = 0.7$; et la coupure serrée ou Tight pour NN-$\tau = 0.9$. Le NN-τ est calculé différemment pour chaque type de tau, il utilise entre autres l'énergie déposée dans les deux premières couches électromagnétiques du calorimètre, l'énergie du candidat tau dans un cône de rayon R < 0.5, ou encore l'impulsion transverse des traces p_T^{trace} et des candidats taus p_T^{τ}.

L'efficacité de reconstruction du réseau de neurones tau NN-τ pour la coupure Medium, en fonction de la luminosité, de l'impulsion transverse, et des angles η_{det} et ϕ, est représentée à la figure 6.4. En moyenne, l'efficacité du réseau de neurones NN-τ pour les sévérités Loose, Medium et Tight sont respectivement de 0.9974 $\pm$0.0004 (0.9960 $\pm$0.0007), 0.974 $\pm$0.001 (0.972 $\pm$0.002) et 0.916 $\pm$0.002 (0.915 $\pm$0.003) pour les données (la simulation).

6.6.4 Efficacité totale

L'efficacité totale d'identification des électrons dans la région inter-cryostat correspond au produit des trois efficacités calculées précédemment : l'efficacité du trajectographe, l'efficacité de reconstruction des objets taus, et l'efficacité du réseau de neurones tau.

6.7 Bruit de fond multijet

Les échantillons de données étudiés pour déterminer les efficacités contiennent peu d'évènements multijets. Ces derniers proviennent essentiellement de la contamination des processus multijets, γ+jet, et W+jet, où un (ou plusieurs) jet(s) est mal identifié comme un électron. Le choix d'un électron Tag de qualité Tight et les coupures appliquées à l'électron Probe pour calculer des différentes efficacités, permettent de considérablement réduire le bruit de fond multijet. Les coupures les plus significatives sur la réduction du bruit de fond lors de la sélection du candidat électron Probe sont $\Delta\phi(\mathrm{e}^{Tag}, \mathrm{e}^{Probe}) > 170°$, IsoHC4 < 3.0 GeV et $\mathrm{p}_T^{trace} > 20$ GeV. Il est difficile d'estimer le bruit de fond multijet à partir de la distribution de la masse invariante des deux électrons, en faisant un ajustement sur le spectre de la masse invariante du boson Z avec une fonction exponentielle, à cause de la longue queue dans la distribution de nos données qui s'étend aux basses masses. Cependant, quand l'une des trois coupures citées précédemment est relâchée, on est capable de réaliser cet ajustement. La coupure sur le p_T^{trace} a été choisie, et est donc supprimée pour cette opération. Ensuite la coupure p_T^{trace} a de nouveau été appliquée, puis nous avons comparé le nombre d'évènements qui ont été rejetés par l'ajout de cette coupure, avec celui des évènements de bruit de fond précédemment estimés avec l'ajustement. La différence entre ces deux nombres correspond au bruit de fond multijet pour cette analyse. Il représente environ $1.0 \pm 0.4\%$ des évènements, autrement dit, l'effet de la soustraction du bruit de fond multijet dans le calcul de nos efficacités est négligeable.

6.8 Facteurs de correction

Comme la simulation ne décrit pas parfaitement les données, les évènements simulés ont été corrigés avec des facteurs de correction que nous avons calculés, pour les trois efficacités déterminées précédemment. Ces facteurs de correction correspondent au rapport, pour chaque bin, entre le nombre des données et celui des évènements simulés, voir section 5.6.3. Puis nous avons transcrit ces informations dans des fichiers spéciaux appelés "spc files". Enfin nous les avons fait certifier et mis à la disposition des analystes, dans le package informatique "emid_eff". La fig. 6.5 montre la distribution de la masse invariante de l'électron dans le calorimètre et de l'électron dans la région inter-cryostat, réalisée avec les spc files que nous avons produits. L'accord entre les données et la simulation est correct.

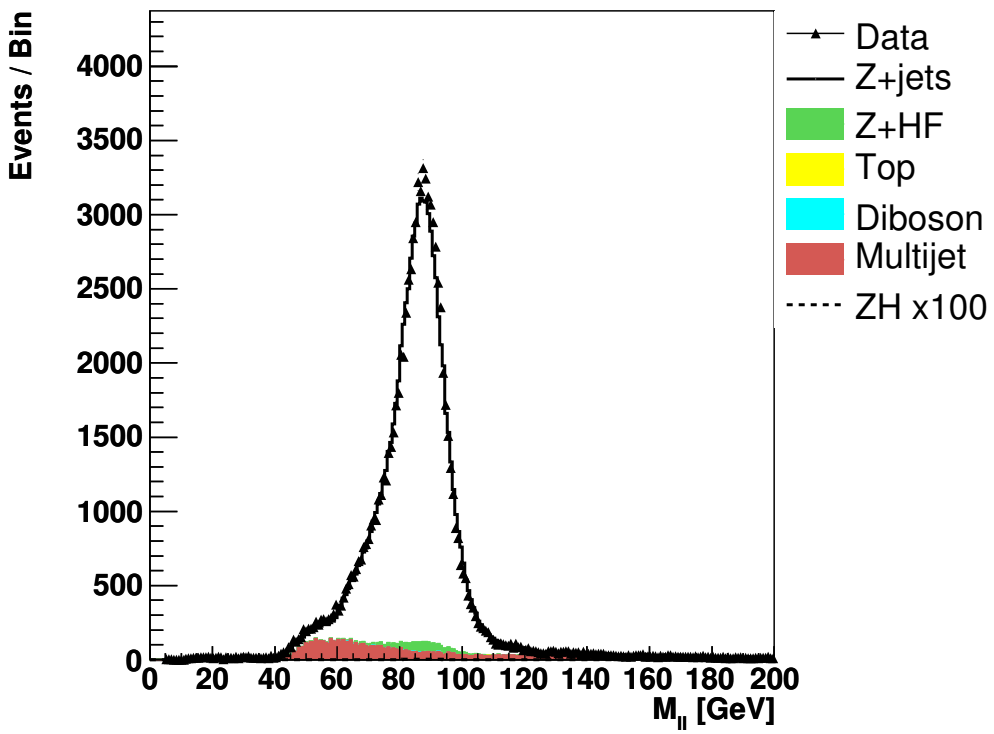


FIG. 6.5 – Distribution de la masse invariante de l'électron dans le calorimètre et de l'électron dans la région inter-cryostat.

6.9 Conclusion

Dans ce chapitre les différentes étapes nécessaires à la certification de l'identification des électrons dans la région inter-cryostat, ainsi que les facteurs de corrections ont été calculés pour les données du Run IIa. L'ajout des électrons se trouvant dans la région inter-cryostat permet en moyenne d'augmenter de plus de 17% l'efficacité du signal. Ces résultats ont été utilisés pour la recherche du Higgs de cette thèse présentée dans le prochain chapitre.

Chapitre 7

Recherche du boson de Higgs dans le canal $ZH \rightarrow e^+e^-b\bar{b}$

Contents

7.1 Principe de l'analyse 95
7.2 Les données 98
7.2.1 Echantillons de données 98
7.2.2 Qualité des données et luminosité 98
7.2.3 Pré-sélection des évènements 99
7.3 Les évènements simulés 102
7.3.1 Génération des évènements simulés 106
7.3.2 Corrections des évènements simulés 106
7.4 Détermination du bruit de fond multijet 109
7.4.1 Les évènements multijets dans l'analyse ee 110
7.4.2 Les évènements multijets dans l'analyse ee^{icr} 111
7.5 Normalisation des bruits de fond 111
7.6 Distributions de contrôle avant étiquetage des jets de b 117
7.7 Etiquetage des jets de b 122
7.7.1 Principe de fonctionnement du b-tagging 122
7.7.2 Le b-tagging dans l'analyse ZH 123
7.7.3 Correction de l'étiquetabilité des jets 125
7.8 Distributions de contrôle après étiquetage des jets de b 125
7.9 Ajustement cinématique 129
7.9.1 Principe 129
7.9.2 Résultats 130
7.10 Analyse discriminante 131
7.10.1 Construction des arbres 131
7.10.2 Variables utilisées 132
7.10.3 Variable discriminante finale 138
7.11 Conclusion 138

Ce chapitre décrit la recherche du boson de Higgs du SM dans le canal $ZH \rightarrow e^+e^-b\bar{b}$. Les résultats de cette analyse ont été présentés à la conférence Moriond en 2009 [77]. A ce jour, ils sont en cours d'approbation par la collaboration DØ pour être publiés au journal Physical Review Letters.
Le signal recherché pour cette analyse est la production d'un boson de Higgs associé à un boson Z, où le boson de Higgs se désintègre en deux jets de quark b, et le boson Z en deux électrons. Cette analyse a été réalisée conjointement avec l'analyse $ZH \rightarrow \mu^+\mu^-b\bar{b}$ qui présente la même topologie que notre signal. Les analyses ont été menées dans le cadre de travail de CAFe (cf. 2.2.5).

7.1 Principe de l'analyse

La recherche du Higgs de cette thèse s'est effectuée dans les données des Runs IIa et IIb du détecteur DØ collectées entre avril 2002 et décembre 2008. La topologie finale recherchée est deux leptons chargés issus d'un boson Z, et deux jets de quark b issus du boson de Higgs. Dans le cas où les leptons chargés sont des électrons, ces derniers sont sélectionnés dans le calorimètre (cette analyse est notée ee). Pour augmenter l'acceptance du signal dans le canal $ZH \rightarrow e^+e^-b\bar{b}$, le cas particulier où l'un des électrons se trouve dans la région inter-cryostat est également analysé (cette analyse est notée ee^{icr}). Dans la cas où les leptons chargés sont des muons, ces derniers sont sélectionnés dans le sytème de muon (cette analyse est notée $\mu\mu$). Pour augmenter l'acceptance du signal dans le canal $ZH \rightarrow \mu^+\mu^-b\bar{b}$, le cas particulier où l'un des muons (qui échappe au détecteur de muon) n'est indentifié que grâce à sa trace, est également analysé (cette analyse est notée $\mu\mu^{trace}$). Ces quatre analyses ee, ee^{icr}, $\mu\mu$, et $\mu\mu^{trace}$ ont été analysées séparément pour les Runs IIa et IIb.

Les analyses ont été réalisées dans Vjets cafe, à partir des fichiers ROOT (n-tuples), cf. 2.2.5, que nous avons produits pour les données réelles et simulées, suivant l'ordre des étapes décrites ci-dessous, voir fig 7.1.

1. La pré-sélection (section 7.2.3 à 7.8), qui se compose de :
   - la sélection des évènements qui contiennent un boson Z. On demande donc que nos évènements aient au moins deux leptons chargés, et que la distribution de leur masse invariante soit compatible avec celle du boson Z. Les coupures de cette première étape de la pré-sélection sont différentes pour les quatre analyses ee, ee^{icr}, $\mu\mu$, et $\mu\mu^{trace}$, ainsi que pour les Runs IIa et IIb.
   - la sélection des évènements qui contiennent au moins deux jets parmis ceux sélectionnés ci-dessus. Les coupures de cette deuxième étape de la pré-sélection sont pratiquement identiques pour les quatre analyses. Le résultat dépend légèrement des Runs IIa et IIb, car les jets y ont été sélectionnés différemments ;

- l'étiquetage des jets de b des évènements ayant au moins deux jets. Cette étape sert à réduire le bruit de fond multijet provenant essentiellement des évènements $Z + jets$ légers, à ce stade de l'analyse ;
- durant cette première étape, les évènements simulés sélectionnés sont corrigés pour qu'ils reproduisent au mieux la réalité.

2. L'ajustement cinématique pour optimiser la valeur des variables cinématiques des leptons et des jets (section 7.9).

3. L'utilisation d'une technique de sélection avancée, qui permet une séparation accrue du signal et du bruit de fond, et donc d'augmenter la sensibilité de l'analyse. Cet outil produit une variable finale discriminante, qui sert à établir la limite (section 7.10).

4. Le calcul de la limite sur la section efficace de production du boson de Higgs sur la variable finale discriminante, avec un logiciel informatique adapté (chapitre 8).

Durant ma thèse, j'ai essentiellement travaillé au niveau de la pré-sélection dans les analyses ee et ee^{icr}. Ma contribution ayant été plus importante dans cette dernière, elle sera par conséquent plus détaillée par la suite. Les quatre analyses ont été réalisées en étroite collaboration comme pour, par exemple, déterminer des facteurs de normalisation dépendant des quatre échantillons, pour optimiser l'accord entre les données et la simulation (cf. 7.5). Je présenterai donc tout au long de mon exposé, les résultats des analyses ee et ee^{icr}, mais également les résultats de la combinaison des quatre analyses.

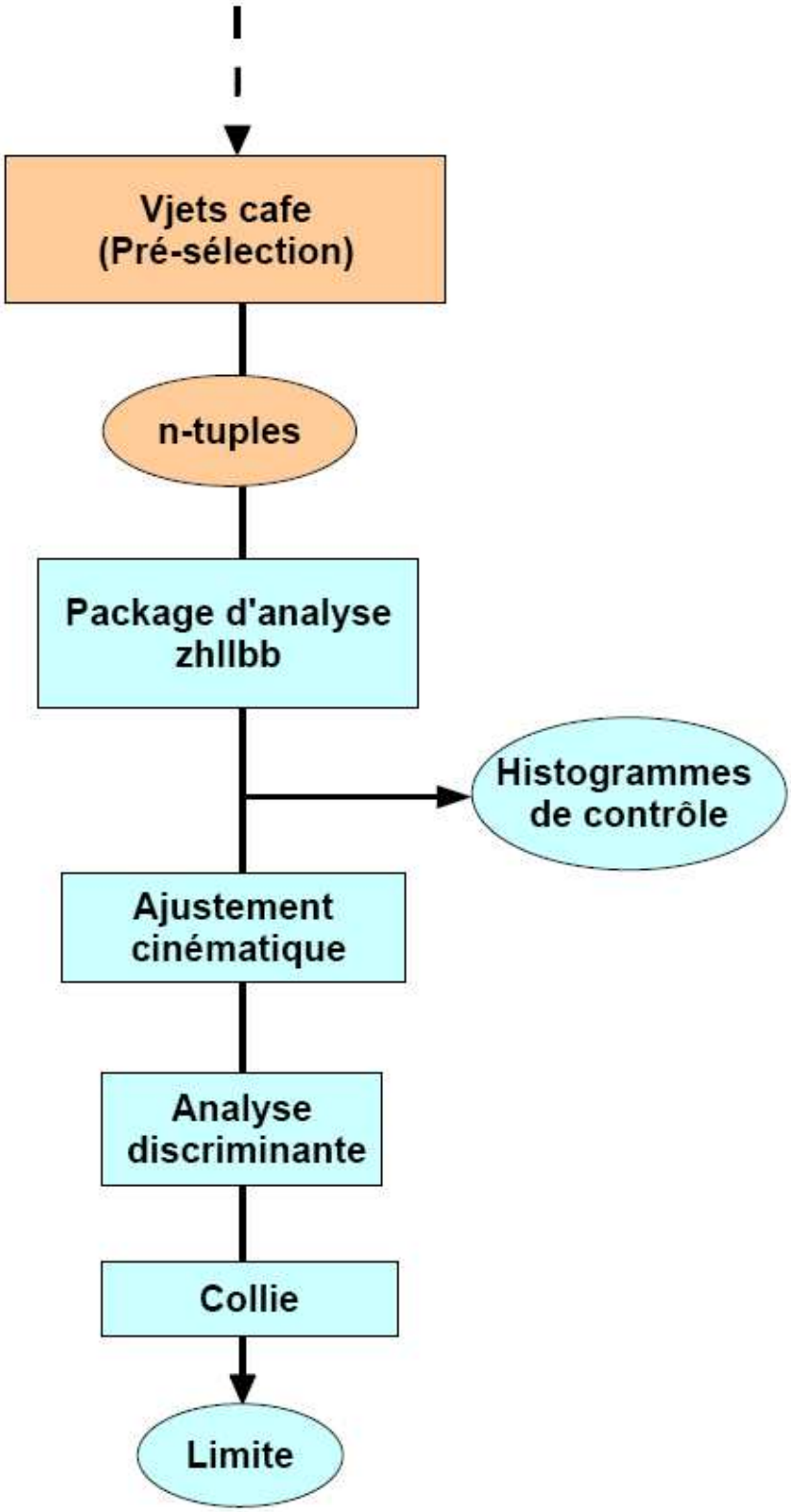


FIG. 7.1 – Chaîne de l'analyse $ZH \rightarrow l^+l^-b\bar{b}$.

7.2 Les données

7.2.1 Echantillons de données

Les données analysées correspondent à l'intégralité des données des Runs IIa et IIb1, et une partie des données du Run IIb2 enregistrées jusqu'au run numéro 247960. Pour plus de simplicité, dans la suite Runs IIb désignera l'ensemble des Runs IIb1 et IIb2 analysés. Les données ont été collectées par le détecteur DØ entre avril 2002 et février 2006 pour le Run IIa et entre juin 2006 et décembre 2008 pour les Runs IIb. Elles ont ensuite été reconstruites par le programme de reconstruction d0reco, puis mises dans le format CAF-trees (cf. 2.2.5).

L'analyse où les deux électrons sont sélectionnés dans le calorimètre électromagnétique (notée ee) a été réalisée en utilisant les skims 2EMhighpt (cf. 2.2.5) ; celle où au moins un des électrons est sélectioné dans la région inter-cryostat (notée ee^{icr}) a été réalisée en utilisant les skims EMinclusive, voir tableau 7.1. Le skim 2EMhighpt ne convient pas à l'analyse ee^{icr} car il contient au moins deux objets électromagnétiques de $\mathrm{p}_T > 12$ GeV. L'analyse ee^{icr} ne demande qu'un électron identifié comme un objet électromagnétique, et l'autre l'électron se trouvant dans la région inter-cryostat doit être identifié comme un objet tau, cf. 6.2. Le skim EMinclusive utilisé pour l'analyse ee^{icr} comporte des évènements avec au moins un objet électromagnétique de $\mathrm{p}_T > 4$ GeV.

Initialement le skim 2EMhighpt comporte 36 294 680 (93 103 997) évènements pour le Run IIa (les Runs IIb), alors que le skim EMinclusive en compte 307 447 272 (672 886 956) pour le Run IIa (les Runs IIb). La taille des skims EMinclusive a donc dû être à nouveau réduite pour diminuer les temps de calcul. Ainsi, nous avons fait des sous-skims des skims EMinclusive des Runs IIa et IIb, avec des coupures appropriées et très lâches, pour n'enlever que les évènements sans intérêts pour notre analyse. Les coupures sont les suivantes, on demande au moins :

- un objet électromagnétique d'impulsion transverse supérieure à 15 GeV, appartenant au CC ou au EC du calorimètre. L'objet électromagnétique doit avoir une isolation inférieure à 0.2, et avoir déposé plus de 90 % de son énergie dans le calorimètre ;
- un objet tau associé à une trace d'impulsion transverse supérieure à 10 GeV, et appartenant à l'ICR.

Le nombre total d'évènements constituant ainsi les sous-skims EMinclusive est fortement réduit à 190 998 (1 046 731) pour le Run IIa (les Runs IIb).

7.2.2 Qualité des données et luminosité

Les évènements bruyants (cf. 2.2.3.3), les mauvais runs (c'est-à-dire ayant un problème de fonctionnement des trajectographes, du calorimètre ou du système à muons), et les mauvais blocs de luminosité (cf. 2.2.2.2), définis par la collaboration [78], ont été enlevés avec le processeur "CafeDataQualityProcessor" du package infor-

Période	Skim
Run IIa (avril 2002 - février 2006)	CSG_CAF_<skim>_PASS3_p18.14.00
Run IIb1 (juin 2006 - août 2007)	CSG_CAF_<skim>_PASS2_p21.10.00
Run IIb2 (octobre 2007 - décembre 2008, jusqu'au run 247960)	CSG_CAF_<skim>_PASS4_p21.10.00_p20.12.00 CSG_CAF_<skim>_PASS4_p21.10.00_p20.12.01 CSG_CAF_<skim>_PASS4_p21.10.00_p20.12.02 CSG_CAF_<skim>_PASS4_p21.10.00_p20.12.04 CSG_CAF_<skim>_PASS4_p21.10.00_p20.12.05

TAB. 7.1 – Skims analysés pour les Runs IIa, IIb1 et IIb2. Les analyses ee et ee^{icr} utilisent respectivement les skims 2EMhighpt et EMinclusive.

matique "caf_dq". Ensuite la luminosité intégrée a été calculée avec l'ensemble des triggers JT_125TT (JT125_L3J125) pour le Run IIa (les Runs IIb). Ils requièrent au moins un jet d'énergie transverse supérieure à 125 GeV. Les ensembles de triggers JT_125TT et JT125_L3J125 n'ont de pas de prescale, (cf. 2.2.4), et ils permettent donc de calculer la luminosité simplement en mesurant le volume de données qu'ils ont accepté durant ces périodes. En pratique, on somme tous les blocs de luminosité correspondant aux données enregistrées, pour obtenir la luminosité intégrée totale de l'échantillon. On utilise pour cela un outil standard de calcul de luminosité "lumitool.py" [79] se trouvant dans le package informatique "dq_util". On trouve ainsi une luminosité de 1.074 ± 0.066 fb^{-1} (3.077 ± 0.188 fb^{-1}) pour le Run IIa (les Runs IIb).

7.2.3 Pré-sélection des évènements

L'étape de la pré-sélection consiste à sélectionner les évènements ayant au moins deux électrons dans le calorimètre électromagnétique provenant d'un boson Z, et au moins deux jets. Pour augmenter l'acceptance du signal, les évènements contenant deux électrons dont un appartient à la région inter-cryostat sont également analysés. Comme il n'y a pas de recouvrement entre le calorimètre et la région inter-cryostat, ces deux lots de données sont orthogonaux. Cette analyse est la première à utiliser les électrons de la région inter-cryostat.

7.2.3.1 Vertex d'interaction principal

Pour que les évènements sélectionnés aient une bonne qualité de reconstruction, leur vertex d'interaction principale (PV) doit se trouver dans la partie centrale du SMT, donc la distance suivant l'axe z entre le PV et le centre du détecteur ($z = 0$) doit être inférieure à 60 cm. Le PV doit également être associé à au moins trois traces.

7.2.3.2 Electrons ee

Le cas où les deux électrons sont reconstruits dans le calorimètre est noté ee. Les électrons sont sélectionnés à l'aide du processeur standard "ElectronSelector" se trouvant dans le package informatique "caf_util", à partir des skims 2EMhighpt réalisés par la collaboration. L'efficacité d'identification des deux électrons issus du Z étant proche de 100%, aucune demande explicite de trigger n'est faite pour garder le maximum d'évènements. Dans l'analyse ee, les évènements sélectionnés doivent avoir au moins deux électrons dont au moins un qui appartient au CC. On a donc deux échantillons : un où les deux électrons appartiennent au CC (noté CC-CC), et un autre où l'un des électrons appartient au CC et l'autre au EC (noté CC-EC). Le cas où les deux électrons appartiennent au EC (noté EC-EC) n'est pas prise en compte, car l'efficacité du signal dans cet échantillon est pratiquement 10 fois plus petite que celle des deux autres configurations CC-CC et CC-EC.
Pour avoir le maximum de signal tout en maîtrisant le bruit de fond, les évènements sont sélectionnés avec des coupures lâches (Loose). Les évènements du Run IIa (des Runs IIb) sont sélectionnés avec la version 12 (11) de la coupure Loose notée LV12 (LV11). Les coupures LV11 sont les suivantes (la majorité des variables ont été définies dans les sections 3.3, 5.3 et 5.6.1) :

- $p_T > 15$ GeV ;
- Iso < 0.1 ;
- EMfr > 0.95 ;
- HMx7 < 35 (CC), et HMx8 < 20 (EC) ;
- IsoHC4 < 3 (CC), et IsoHC4 < 2 (EC) ;
- TrkM > 0.001 ou EMHits > 0.4 (CC) ;
- Δz(PV, e)< 1 cm, la distance suivant l'axe z entre l'impulsion transverse de l'électron et le PV doit être inférieure à 1 cm ;
- NN-elec7 > 0.2 (CC), et NN-elec3 $> (0.4)$ (EC).

Les coupures LV12 sont pratiquement les mêmes que les coupures LV11 citées ci-dessus sauf : TrkM > 0.0 dans le CC ; et l'ajout de la coupure E/p < 4.0 dans le CC, où E représente l'énergie mesurée dans le calorimètre et p l'impulsion de la trace associée.

7.2.3.3 Electrons ee^{icr}

Le cas où l'un des deux électrons se trouve dans le calorimètre et l'autre dans la région inter-cryostat (ICR) est noté ee^{icr}. Ces évènements sont sélectionnés à partir des sous-skims EMinclusive que nous avons réalisés. On demande donc deux électrons dont un qui appartient au CC ou aux EC du calorimètre, et l'autre doit appartenir à l'ICR. L'électron du calorimètre, noté e, doit avoir déclenché l'ensemble de triggers Single EM, ce dernier requiert au moins un objet électromagnétique de $p_T > 20$ GeV. L'électron e doit également passer les coupures LV12 (LV11) définies pour le Run IIa (les Runs IIb), cf. section 7.2.3.2. L'électron de l'ICR, noté e^{icr}, est sélectioné avec

le processeur standard "ICRElectronSelector" du package informatique "caf_util". Comme il n'y a pratiquement pas de couche électromagnétique dans cette région du détecteur, l'électron e^{icr} est reconstruit comme un objet tau, (cf. section 6.2). Il doit par conséquent passer les coupures adéquates, moyennement sévères (Medium) suivantes (la majorités des variables ont été définies dans la section 6.6.1) :

- se trouver dans la région ICR $1.1 < |\eta_{det}| < 1.5$;
- avoir une $E_T > 20$ GeV et un $p_T > 20$ GeV ;
- avoir une trace associée telle que le $p_T^{trace} > 20$ GeV, et dont l'extrapolation se trouve dans la région ICR ;
- la trace doit avoir produit au moins un coup dans le SMT ;
- l'électron e^{icr} doit passer une coupure de réseau de neurones de moyenne sévérité, NN-$\tau > 0.7$;
- l'électron e^{icr} reconstruit comme un tau de type 1 doit avoir déposé au moins 10% de son énergie dans l'ICD ;
- l'électron e^{icr} identifié comme un tau de type 3 est plus souvent mal reconstruit que l'électron identifié comme un tau de type 1 ou 2, car il a au moins deux traces associées. Il doit par conséquent passer une coupure supplémentaire de likelihood : LHood-T3 > 0.15. Cette dernière rejette plus de 98% des évènements multijets ayant passés la coupure NN-τ.

7.2.3.4 Bosons Z

On reconstruit le boson Z à partir de ses produits de désintégration. L'étape de la sélection des bosons Z consiste donc à sélectionner deux électrons, ee ou ee^{icr}, dont la masse invariante soit compatible avec celle du Z, c'est-à-dire qu'elle doit se trouver dans l'intervalle de masse compris entre 60 GeV et 150 GeV.

7.2.3.5 Jets

Les jets sont reconstruits avec l'algorithme de cône Run II, dans un cône de rayon R = 0.5, et corrigés de l'échelle d'énergie des jets (JES), (cf. section 3.4). Au moins deux "bons" jets sont demandés dans chaque évènement, répondant aux critères suivants :

- le premier (second) jet le plus énergétique doit avoir un $p_T^{jet1} > 20$ GeV ($p_T^{jet2} > 15$ GeV) après correction ;
- se trouver dans la région $|\eta_{det}| < 2.5$;
- chaque jet doit être distant des deux électrons de $\Delta R > 0.5$, cf Eq. 3.1 ;
- être associé à au moins deux traces issues du vertex primaire (VertexConfirmedJets). Cette condition n'est demandée que pour les évènements des Runs IIb. A haute luminosité, le nombre d'interactions augmente et cette condition supprime des jets qui n'appartiennent pas à l'interaction où les leptons sont issus.

7.3 Les évènements simulés

Pour avoir un modèle permettant de décrire les données, on simule les processus du SM ayant le même état final que notre signal. Ces processus sont appelés bruits de fond SM. Puis on compare la somme de leur contribution aux données. On simule également le signal ZH. Le bruit de fond instrumental est déterminé à partir des données, cf. 7.4.

Les bruits de fond SM : les principaux bruits de fond SM dans cette analyse sont les suivants :

- les productions $Z + jets$ légers, où le Z se désintègre en e^+e^- et les $jets$ légers proviennent d'un quark léger (u, d, c) ou d'un gluon. Ce bruit de fond est dominant avant l'étiquetage des jets de b.
- les productions $Z + jets$ lourds, quand le Z se désintègre en e^+e^- et les $jets$ lourds proviennent d'un quark b ou c. Ce bruit de fond est dominant après l'étiquetage des jets de b.
- les productions di-bosons ZZ, WW, et WZ, accompagnées éventuellement par des jets, où la majorité du bruit de fond provient des cas où un des bosons se désintègre en deux jet de b, et l'autre se désintègre en électron et neutrino.
- les productions $t\bar{t}$, lorsque les t se désintègrent en un jet de b accompagné d'un W qui donne un électron et un neutrino.

Les différents bruits de fond SM simulés sont résumés dans les tableaux 7.2 (7.3) pour le Run IIa (les Runs IIb).

Le signal : étant donné que la masse du Higgs n'est pas connue, les évènements de signal ZH ont été générés pour plusieurs masses comprises entre 100 GeV et 150 GeV, par pas de 5 GeV. Les évènements de signal simulés sont résumés dans le tableau 7.4 pour les Runs IIa et IIb.

Processus (M_Z GeV)	$\sigma(\times$ BR)	req-id's	Nombre d'évènements
Z+0lp excl. (15 - 75)	337 pb	92112/3/4	580k
Z+0lp excl. (75 - 130)	132 pb	92152/3/4/5/6, 92287/8/9, 92290/1, 92527/28/29/30/31	2980k
Z+0lp excl. (130 - 250)	0.891 pb	92072/3	300k
Z+1lp excl. (15 - 75)	40.1 pb	92115/6/7 91539, 91541/3	480k
Z+1lp excl. (75 - 130)	40.8 pb	92157/58/59/60/61, 92272/3/4/5/6	1910k
Z+1lp excl. (130 - 250)	0.373 pb	92074	190k
Z+2lp excl. (15 - 75)	10.0 pb	92118	190k
Z+2lp excl. (75 - 130)	10.0 pb	92162/3/4/5/6	990k
Z+2lp excl. (130 - 250)	95.4 fb	92075	100k
Z+3lp incl. (15 - 75)	2.76 pb	92119	100k
Z+3lp incl. (75 - 130)	3.15 pb	92167/68/69/70/71	990k
Z+3lp incl. (130 - 250)	31.9 fb	92076	100k
Z+2b+0lp excl. (15 - 75)	0.546 pb	84892	200k
Z+2b+0lp excl. (75 - 130)	0.395 pb	73952	200k
Z+2b+0lp excl. (130 - 250)	3.49 fb	85452	110k
Z+2b+1lp excl. (15 - 75)	0.192 pb	84912	100k
Z+2b+1lp excl. (75 - 130)	0.193 pb	73953	100k
Z+2b+1lp excl. (130 - 250)	1.71 fb	85455	10k
Z+2b+2lp incl. (15 - 75)	81.3 fb	84915	100k
Z+2b+2lp incl. (75 - 130)	97.3 fb	73954	50k
Z+2b+2lp incl. (130 - 250)	0.969 fb	85458	50k
Z+2c+0lp excl. (15 - 75)	4.08 pb	85112	200k
Z+2c+0lp excl. (75 - 130)	0.900 pb	74418	200k
Z+2c+0lp excl. (130 - 250)	7.43 fb	85461	100k
Z+2c+1lp excl. (15 - 75)	1.04 pb	85132	100k
Z+2c+1lp excl. (75 - 130)	0.478 pb	74415	120k
Z+2c+1lp excl. (130 - 250)	4.57 fb	85464	50k
Z+2c+2lp incl. (15 - 75)	0.347 pb	85135	100k
Z+2c+2lp incl. (75 - 130)	0.289 pb	74412	50k
Z+2c+2lp incl. (130 - 250)	2.80 fb	85467	50k

Processus	$\sigma(\times$ BR)	req-id's	Nombre d'évènements
$t\bar{t}$ ($\rightarrow l^+l^-$) + 0lp excl.	0.352 pb	94272/3/4/5	770k
$t\bar{t}$ ($\rightarrow l^+l^-$) + 1lp excl.	0.143 pb	94276/7/8	480k
$t\bar{t}$ ($\rightarrow l^+l^-$) + 2lp incl.	0.071 pb	94292/3	290k
WW incl.	11.6 pb	38488/89/90, 38939/40/41/42/43, 42208/09/10/11,79214,79553	2100k
WZ incl.	3.25 pb	38491, 42212, 79212, 79554	400k
ZZ incl.	1.33 pb	38492, 42213, 79213, 79552	300k

TAB. 7.2 – Les différents processus du bruit de fond SM, pour le Run IIa. σ est la section efficace et BR est le rapport de branchement, req-id's est leur identifiant pour la simulation.

Processus (M_Z GeV)	$\sigma(\times$ BR)	req-id's	Nombre d'évènements
Z+0lp excl. (15 - 75)	337 pb	70273, 87252/3/4, 91538, 91540/2, 93635/8/9	1580k
Z+0lp excl. (75 - 130)	132 pb	67713, 72772/3, 93632, 93634	1010k
Z+0lp excl. (130 - 250)	0.891 pb	87262/3	350k
Z+1lp excl. (15 - 75)	40.0 pb	70274, 87255/6/7, 91539, 91541/3	1580k
Z+1lp excl. (75 - 130)	40.1 pb	67714, 72792/3	570k
Z+1lp excl. (130 - 250)	0.376 pb	87264	180k
Z+2lp excl. (15 - 75)	9.95 pb	70275, 87258/9, 91544/6	550k
Z+2lp excl. (75 - 130)	9.98 pb	67715, 72812	270k
Z+2lp excl. (130 - 250)	102 fb	87265	160k
Z+3lp incl. (15 - 75)	2.77 pb	70276, 87260/1, 91545/7	540k
Z+3lp incl. (75 - 130)	3.30 pb	67716, 72832	130k
Z+3lp incl. (130 - 250)	33.1 fb	87266, 91112	300k
Z+2b+0lp excl. (15 - 75)	0.518 pb	89372	170k
Z+2b+0lp excl. (75 - 130)	0.401 pb	68978	200k
Z+2b+0lp excl. (130 - 250)	3.40 fb	86575	90k
Z+2b+1lp excl. (15 - 75)	0.207 pb	89375	85k
Z+2b+1lp excl. (75 - 130)	0.173 pb	68981	93k
Z+2b+1lp excl. (130 - 250)	1.67 fb	89376	45k
Z+2b+2lp incl. (15 - 75)	78.3 fb	89374	82k
Z+2b+2lp incl. (75 - 130)	107 fb	68984	44k
Z+2b+2lp incl. (130 - 250)	0.938 fb	89377	44k
Z+2c+0lp excl. (15 - 75)	4.09 pb	86951	180k
Z+2c+0lp excl. (75 - 130)	0.901 pb	87892	180k
Z+2c+0lp excl. (130 - 250)	7.50 fb	86578, 87893	180k
Z+2c+1lp excl. (15 - 75)	1.03 pb	87895, 91548	180k
Z+2c+1lp excl. (75 - 130)	0.506 pb	71093	90k
Z+2c+1lp excl. (130 - 250)	4.33 fb	87896, 91575	90k
Z+2c+2lp incl. (15 - 75)	0.382 pb	87898, 91549	170k
Z+2c+2lp incl. (75 - 130)	0.286 pb	87899	50k
Z+2c+2lp incl. (130 - 250)	2.67 fb	87900, 91577	90k

Processus	$\sigma(\times$ BR)	req-id's	Nombre d'évènements
$t\bar{t}$ ($\rightarrow l^+l^-$) + 0lp excl.	0.352 pb	87299, 87300/1/2	750k
$t\bar{t}$ ($\rightarrow l^+l^-$) + 1lp excl.	0.142 pb	87313/4/5	450k
$t\bar{t}$ ($\rightarrow l^+l^-$) + 2lp incl.	0.068 pb	87322/3	280k
WW incl.	11.6 pb	86772/3/9, 86780	710k
WZ incl.	3.25 pb	86774/5/6/7	630k
ZZ incl.	1.33 pb	86778, 86787/8	540k

TAB. 7.3 – Les différents processus du bruit de fond SM, pour les Runs IIb. σ est la section efficace et BR est le rapport de branchement, req-id's est leur identifiant pour la simulation.

Run IIa

Processus	$\sigma(\times$ BR)	req-id's	Nombre d'évènements
$ZH \to l^+l^-bb$, $M_H = 100$ GeV	13.70 fb	96752, 96799	394k
$ZH \to l^+l^-bb$, $M_H = 105$ GeV	11.6 fb	96753, 96800	403k
$ZH \to l^+l^-bb$, $M_H = 110$ GeV	9.60 fb	94754, 96801	396k
$ZH \to l^+l^-bb$, $M_H = 115$ GeV	7.97 fb	96755, 96802	400k
$ZH \to l^+l^-bb$, $M_H = 120$ GeV	6.35 fb	96756, 96803	405k
$ZH \to l^+l^-bb$, $M_H = 125$ GeV	5.01 fb	96757, 96804	388k
$ZH \to l^+l^-bb$, $M_H = 130$ GeV	3.75 fb	96758, 96805	386k
$ZH \to l^+l^-bb$, $M_H = 135$ GeV	2.74 fb	96759, 96806	387k
$ZH \to l^+l^-bb$, $M_H = 140$ GeV	1.88 fb	96760, 96807	376k
$ZH \to l^+l^-bb$, $M_H = 145$ GeV	1.23 fb	96761, 96808	387k
$ZH \to l^+l^-bb$, $M_H = 150$ GeV	0.746 fb	96762, 96809	388k

Runs IIb

Processus	σ ($\times$BR)	req-id's	Nombre d'évènements
$ZH \to l^+l^-bb$, $M_H = 100$ GeV	13.70 fb	87432, 89872	309k
$ZH \to l^+l^-bb$, $M_H = 105$ GeV	11.6 fb	70312, 87433	270k
$ZH \to l^+l^-bb$, $M_H = 110$ GeV	9.60 fb	87434, 89873	309k
$ZH \to l^+l^-bb$, $M_H = 115$ GeV	7.97 fb	63733, 87435	271k
$ZH \to l^+l^-bb$, $M_H = 120$ GeV	6.35 fb	87436, 89874	311k
$ZH \to l^+l^-bb$, $M_H = 125$ GeV	5.01 fb	70313, 87437	270k
$ZH \to l^+l^-bb$, $M_H = 130$ GeV	3.75 fb	87438, 89875	311k
$ZH \to l^+l^-bb$, $M_H = 135$ GeV	2.74 fb	70314, 87439	271k
$ZH \to l^+l^-bb$, $M_H = 140$ GeV	1.88 fb	87440, 89876	309k
$ZH \to l^+l^-bb$, $M_H = 145$ GeV	1.23 fb	70315, 87441	270k
$ZH \to l^+l^-bb$, $M_H = 150$ GeV	0.746 fb	87442, 89877	306k

TAB. 7.4 – Le signal ZH pour plusieurs masses comprises entre 100 GeV et 150 GeV, pour les Runs IIa (en haut) et IIb (en bas). σ est la section efficace et BR est le rapport de branchement, req-id's est leur identifiant pour la simulation.

7.3.1 Génération des évènements simulés

Les évènements de signal ZH et de bruits de fond WW, WZ, ZZ ont été générés avec PYTHIA [80], un générateur au premier ordre (LO : Leading Order) de la théorie des perturbations.

Les processus $Z + jets$ et $t\bar{t}$ ont été générés avec ALPGEN [81]. En utilisant ALPGEN, on peut générer les évènements en précisant le nombre n ($n = 0, 1, ...$) de parton(s), qui accompagne le Z. Ainsi, on peut avoir une statistique comparable pour des valeurs de n différentes. On l'interface à PYTHIA qui fournit la modélisation de la fragmentation des quarks. On applique un algorithme de "matching" [82] pour éviter un double comptage des partons générés par ALPGEN et ensuite par PYTHIA.
Pour augmenter la statistique, les échantillons $Z + nlp$, $Z + b\bar{b} + nlp$ et $Z + c\bar{c} + nlp$, où nlp représente le nombre de parton(s) léger(s), sont simulés séparément. De nouveau, pour éviter le double comptage, les évènements contenant des quarks b et c ont été enlevés des échantillons $Z + nlp$; les évènements contenant des quarks c ont été enlevés des échantillons $Z + b\bar{b} + nlp$ [83].

7.3.2 Corrections des évènements simulés

La simulation ne reproduit pas parfaitement les données, par exemple, la présence de matériaux morts mal connue dans le détecteur n'est pas exactement prise en compte dans la simulation, et cela se traduit par des désaccords plus ou moins importants entre les données et la simulation. Les évènements simulés sont donc corrigés. L'ensemble des corrections appliquées à la simulation sont décrites dans les prochains paragraphes :

Correction de l'énergie des électrons du calorimètre : La résolution en énergie des électrons dans les évènements simulés est meilleure que celle des données. De plus la résolution a été améliorée avec une calibration basée sur la H-Matrix et l'angle ϕ de l'électron. L'énergie des électrons dans les données et la simulation a été corrigée avec avec le processeur "phimodCorr" du package informatique "EMresolution_cafe" [84]. Puis l'énergie des électrons des évènements simulés a été dégradée (Smearing) avec le processeur "EMmcsmear" du même package informatique. Le Smearing a été effectué avec une fonction de Crystal Ball [85]. Cette dernière est composée d'une Gaussienne convoluée à une fonction puissance, et elle permet de modéliser les processus en tenant compte de la perte d'énergie.

Correction de l'énergie des électrons de l'ICR : la mesure de l'énergie des électrons e^{icr} est moins bonne par rapport à celle des électrons e, à cause de la quasi-abscence des couches électromagnétiques du calorimètre dans l'ICR. L'énergie des e^{icr} a donc été calculée à partir de l'impulsion transverse de la trace qui leur est associée p_T^{trace}. Cette dernière étant meilleure dans la simulation que dans les données, elle a été dégradée avec une fonction composée d'une gaussienne et d'une Landau, [86] [87].

Correction de l'efficacité d'identification des électrons du calorimètre : les différences observées dans les courbes d'efficacité d'identification des électrons du calorimètre entre les données et la simulation, sont corrigées avec des facteurs d'échelle qui tiennent compte de ces différences. Le facteur global de correction est le produit des deux termes suivants (cf. chapitre 5) : la correction de l'efficacité des critères de pré-sélection de l'électron paramétrée en fonction de η_{det}, et celle de l'efficacité de sélection en fonction de η_{det} et de ϕ. Ces corrections sont appliquées évènement par évènement avec le processeur "EMCorr" du package informatique "caf_eff_utils".

Correction de l'efficacité d'identification des électrons de l'ICR : les désaccords observés entre les données et la simulation dans les efficacités de reconstruction des électrons de l'ICR ont été corrigés. Le facteur global de correction est le produit des trois termes suivants, (cf. chapitre 6) : la correction de l'efficacité d'avoir une trace dans l'ICR, la correction de l'efficacité d'avoir un objet tau associé à une trace dans l'ICR, et enfin la correction de l'efficacité qu'un objet tau dans l'ICR, associé à une trace, passe une coupure de réseau de neurones NN-τ. Les corrections ont été appliquées avec le processeur "ICRCorr" du package informatique "emid_eff_utils".

Correction de l'impulsion transverse du boson Z : la distribution de l'impulsion transverse du boson Z est mal modélisée par les générateurs ALPGEN+PYTHIA, quand le p_T du Z est inférieur à 100 GeV. On a donc repondéré (re-weighting en anglais) [88] le p_T des évènements simulés, en tenant compte des différences entre les données et la simulation au niveau du générateur, voir fig. 7.2. Les paramètres de la correction utilisés correspondent à la version p20Alpgen_njet15 d'ALPGEN, qui traite séparément les évènements avec 0, 1 et au moins 2 jets, pour des jets ayant une $E_T > 15$ GeV. Comme les évènements simulés des Runs IIa et IIb ont été générés avec la même version et les mêmes paramètres d'ALPGEN, les mêmes corrections leurs ont été appliquées. Les corrections ont été effectuées avec le processeur "ZPtReweighting" du package informatique "caf_mc_util".

Correction de l'énergie des jets : la résolution en énergie des jets des évènements simulés étant meilleure que celle des données, elle a été corrigée avec une technique appelée JSSR (pour Jets Smearing, Shifting and Removing) [89] [90], développée dans la collaboration DØ afin d'ajuster leur énergie à celle des données. Cette technique est basée sur le déséquilibre, noté ΔS, en impulsion transverse entre le jet et le boson, dans des évènements $\gamma + jet$ ou $Z + jet$, des données réelles et simulées, dans lesquels le boson et le jet doivent être dos-à-dos. La variable ΔS s'exprime de la manière suivante :

$$\Delta S = \frac{p_T^{jet} - p_T^{\gamma/Z}}{p_T^{\gamma/Z}}. \tag{7.1}$$

La comparaison de ΔS entre les données réelles et simulées est paramétrée par l'impulsion transverse du boson, qui est meilleure que celle du jets, puis ajustée pour avoir les paramètres de la correction. Les jets simulés ayant une meilleure résolution que celle des données, cette opération va dégrader (Smearing) et décaler (Shifting)

l'impulsion des jets simulés, qui sera alors proche de celle des jets des données, voir fig. 7.3. Les jets simulés ayant un $p_T < 15$ GeV seront rejetés (Removing). Ces corrections ont été appliquées avec le processeur "JetSSR" du package informatique "caf_mc_util".

Correction en η et ΔR des jets : en général, les distributions en η et ΔR des jets ne sont pas bien reproduites par la simulation dans les échantillons $Z + jets$. Ces variables ont donc été corrigées au niveau du générateur, avec des facteurs de corrections déterminés par la collaboration. Les corrections ont été calculées successivement avec un ajustement de la distribution des rapports entre les données réelles et simulées pour η_{det} et ΔR, respectivement avec une fonction polynomiale du quatrième et du troisième ordre [91], voir fig. 7.4. Ces corrections ont été appliquées aux échantillons $Z + nlp$, $Z + b\bar{b} + nlp$, et $Z + c\bar{c} + nlp$.

Correction associée à la confirmation de vertex : l'efficacité associée à la confirmation de vertex est différente entre les données et la simulation. La collaboration a donc certifié des facteurs de correction pour cette coupure [92]. Les facteurs de correction ont été calculés à partir des échantillons $\gamma + jets$ et $Z + jets$, et ils ont été paramétrés en fonction de la position du vertex primaire par rapport à l'axe z, de la pseudo-rapidité et de l'impulsion transverse du jet. Cette correction n'a été appliquée qu'aux évènements simulés des Runs IIb.

Correction de la section efficace : en général, la section efficace de production des évènements simulés est calculée au premier ordre (LO) avec les générateurs. Pour prédire le nombre total d'évènements aux ordres plus élevés, les évènements sont pondérés avec un facteur dit k-facteur et noté kf. Le kf d'un processus est le rapport entre sa section efficace calculée aux ordres supérieurs, typiquement au second (NLO) ou au troisième (NNLO) ordre de la théorie des perturbations, et sa section efficace calculée à l'ordre LO (kf = N(N)LO / LO) [93].

– La section efficace de production des processus $Z + jets$ légers a été calculée avec ALPGEN à l'ordre des logarithmes dominants (LL : Leading Log) de la théorie des perturbations, puis normalisée à l'ordre NNLO. Comme ce facteur d'échelle n'a pas été calculé à partir de la section efficace à l'ordre LO, il est appelé kf'. Le kf' vaut ainsi 1.30 et il a été appliqué à tous les processus $Z + jets$ légers.

– Le kf (NLO/LO) a été calculé pour les saveurs lourdes $Z + b\bar{b}$ et $Z + c\bar{c}$, et pour les processus $Z + jets$ légers avec MCFM [94], un générateur dont les calculs sont faits au deuxième ordre de la théorie des perturbations. Ensuite, MCFM a déterminé les facteurs de correction associés à la section efficace de production des saveurs lourdes (noté HF), en faisant le rapport du kf de chaque saveur lourde par le kf des jets légers : HF = kf($Z + b\bar{b}/c\bar{c}$) / kf($Z + nlp$). Ainsi, la section efficace des processus $Z + jets$ lourds générée par ALPGEN est corrigée avec ce facteur supplémentaire. La correction totale pour ces processus vaut donc :

$$\text{kf' * HF}_{b\bar{b}} = 1.30 * 1.52 = 1.96, \text{ pour les } Z + b\bar{b} + nlp$$

$$\text{kf' * HF}_{c\bar{c}} = 1.30 * 1.67 = 2.15, \text{ pour les } Z + c\bar{c} + nlp.$$

– La section efficace de production des processus $t\bar{t}$ et des di-bosons est également corrigée avec un kf (NLO/LO). Les sections efficaces LO (resp. NLO) ont été calculées avec PYTHIA et ALPGEN (resp. MCFM). Les kf valent 1.03 pour les ZZ, 1.06 pour les WZ, 1.01 pour les WW, et 1.43 pour les $t\bar{t}$.

– La section efficace de production du signal ZH a été corrigée comme conseillé en [95]. Ces corrections tiennent compte des corrections électro-faibles par exemple.

Correction des paramètres d'ALPGEN : après les corrections appliquées à η_{det} et ΔR des jets simulés, on observe encore des désaccords entre les données et la simulation. Ces différences proviennent des paramètres d'ALPGEN utilisés dans la simulation comme les facteurs de renormalisation, la variable utilisée pour déterminer la valeur de la constante de couplage à chaque vertex, ou encore les différents paramètres servant au matching. Ces paramètres d'ALPGEN ont donc été corrigés comme décrit en [91].

Correction du profil de la luminosité instantanée : le profil de la luminosité instantanée des évènements de biais minimum (cf. 2.2.4) dans la simulation, n'est pas identique à celui des données. Les évènements simulés ont donc été repondérés avec le processeur "LumiReWeighting" du package informatique "lumi_profiles" [96], voir fig. 7.5.

Correction du vertex d'interaction principal : durant un store, la forme du point de collision change et devient plus large à la fin du store, à cause de la croissance de l'émittance des faisceaux. La distribution du vertex d'interaction principal des évènements simulés est modélisée avec une fonction Gaussienne de largeur constante, mais cette fonction n'est pas adéquate. Ainsi, le vertex d'interaction principal du boson Z des évènements simulés est repondéré avec le processeur "BeamWeightProc" du package informatique "beamposition" [97], voir fig. 7.6

Correction du trigger EM : dans l'analyse ee^{icr}, l'électron e doit avoir déclenché l'ensemble de triggers Single EM. L'effet de cet ensemble de triggers a été déterminé dans la sélection des données, et il a été pris en compte lors de la simulation avec les processeurs "cafTriggerEfficiency", "Probability_nEM" et "cafTriggerOutput" du package informatique "caf_trigger" [98].

7.4 Détermination du bruit de fond multijet

Le bruit de fond multijet (ou QCD) dans cette analyse, correspond aux évènements ayant un ou plusieurs jets mal identifiés comme un électron e dans le calorimètre, ou comme un électron e^{icr} dans l'ICR, (on dit également fake e, ou fake e^{icr}). Cette mauvaise identification est liée aux imperfections du détecteur, et aux divers coupures de sélection qui n'ont pas une efficacité non nulle pour le bruit de

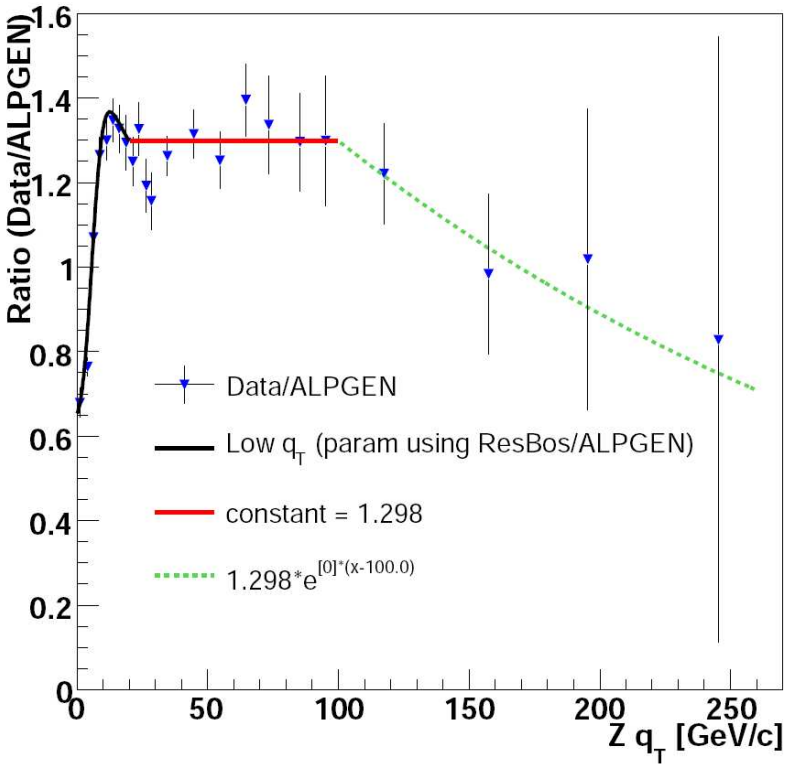


FIG. 7.2 – Rapport entre les données et la simulation de l'impulsion transverse du boson Z.

fond, ce qui est inévitable si on veut garder une efficacité relativement élevée pour le signal. Comme le bruit de fond multijet n'est pas correctement reproduit par la simulation, il est déterminé à partir des données selon les conditions suivantes : les critères de sélection doivent être proches de ceux utilisés pour les données, et la statistique doit être suffisante. Une détermination correcte du bruit de fond multijet se fait en deux étapes, d'abord on détermine la forme de la distribution du bruit de fond multijet, voir sous-sections 7.4.1 et 7.4.2, ensuite on calcule le nombre des évènements multijets, voir section 7.5.

7.4.1 Les évènements multijets dans l'analyse ee

Le bruit de fond QCD dans l'analyse ee est directement déterminé à partir du skim 2EMhighpt. Les évènements multijets doivent passer les mêmes coupures que les données lors de la pré-sélection, cf. 7.2.3.2. Les seules différences concernent les coupures suivantes : Iso < 0.2, EMfr > 0.9, HMx7 > 35 (CC) ou HMx8 > 20 (EC), les coupures NN-elec7 et NN-elec3 ont été supprimées, et enfin les évènements ayant plus d'un électron de qualité Loose ont été enlevés.
Etant donné que l'ensemble des termes de triggers favorise les objets EM ayant un faible HMx7 ou HMx8, pour ne pas introduire de biais dû à l'inversion de la coupure sur ces variables, on corrige le bruit de fond QCD en p_T et η, de sorte que sa distribution soit en accord avec la distribution de référence, composée d'électrons

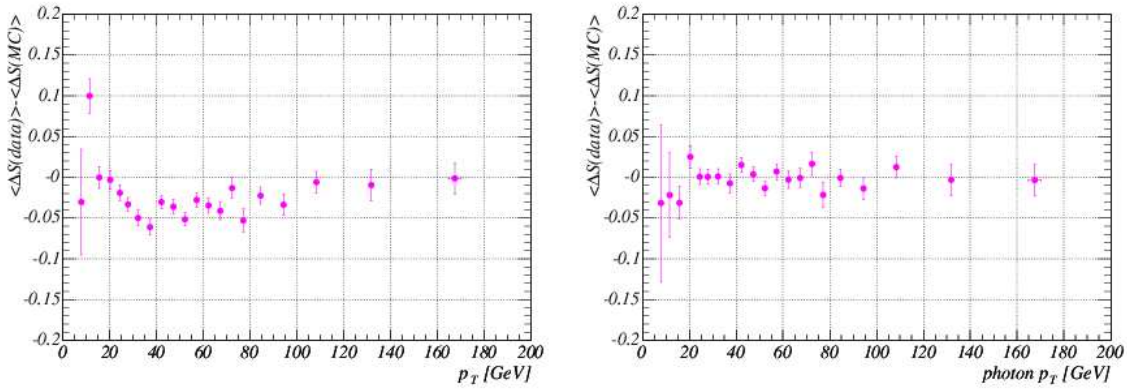


FIG. 7.3 – Différence de la variable ΔS entre les données et la simulation, avant correction (à gauche), et après correction (à droite), dans la partie centrale du calorimètre.

de mêmes charges, non biaisée et déterminée séparément comme expliqué dans la note [99]. Les facteurs de correction correspondent au rapport entre la distribution de référence et celle du bruit de fond multijet, en fonction du p_T et η du premier et du deuxième électron les plus énergétiques.

7.4.2 Les évènements multijets dans l'analyse ee^{icr}

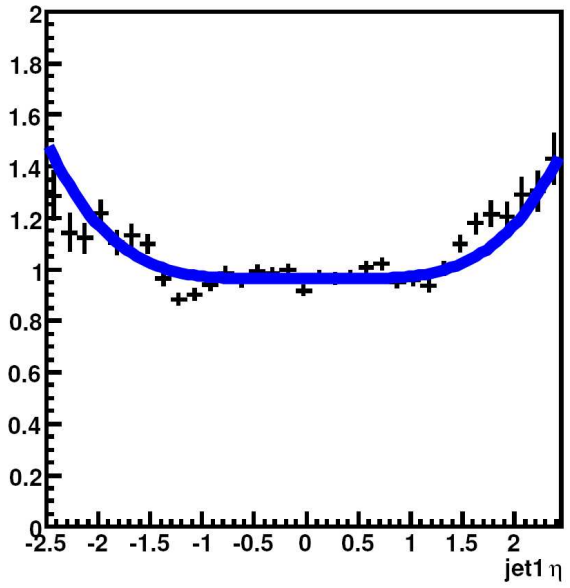
Le bruit de fond multijet dans l'analyse ee^{icr} est directement déterminé à partir du skim EMinclusive que nous avons réalisé. La détermination des évènements multijets dans l'analyse ee^{icr} se fait en sélectionnant un électron e et un fake e^{icr}. Les électrons e doivent passer les mêmes coupures que les données lors de la présélection, cf. 7.2.3.2, alors que le fake e^{icr} est déterminé dans la région de bruit de fond de la variable de sortie des réseaux de neurones NN-τ et NN-elecτ adapté aux objets taus, soit respectivement $0.2 <$ NN-$\tau < 0.4$ et NN-elec$\tau < 0.03$.

7.5 Normalisation des bruits de fond

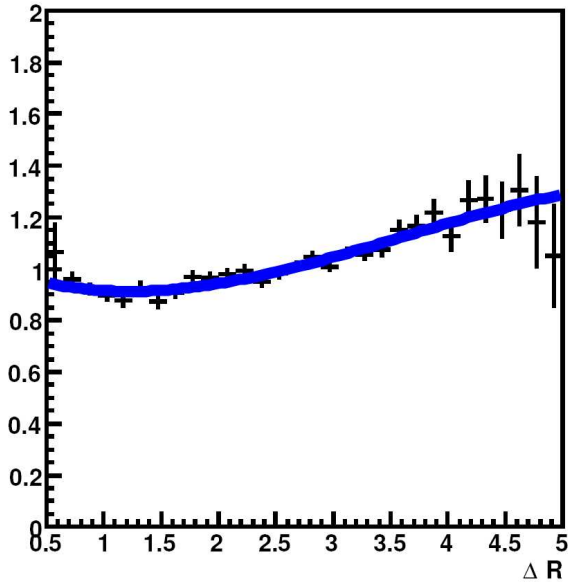
Les évènements simulés ont été normalisés en fonction de la luminosité L, de la section efficace de production σ, et de la correction de l'efficacité ϵ_{SF} (rapport entre l'efficacité des données et de la simulation), avec des poids ($N = L\ .\ \epsilon_{SF}\ .\ \sigma$). Pour déterminer le nombre d'évènements multijets et pour améliorer la modélisation des différents bruits de fond, des facteurs de normalisation ont été calculés pour chaque échantillon de bruit de fond, à partir d'un ajustement réalisé sur la distribution de la masse invariante du boson Z des données. L'ajustement permet d'obtenir le meilleur accord possible entre les données et les bruits de fond. L'ajustement a été effectué simultanément pour les quatre canaux (ee, ee^{icr}, $\mu\mu$, et $\mu\mu^{trace}$) et pour les différentes multiplicités des jets. Il consiste à minimiser le χ^2 suivant :

$$\chi^2 = \sum_i \sum_j \sum_m \frac{\left[D_m^{ij} - \alpha^{ij} \,.\, Q_m^{ij} - k_L \,.\, k_\epsilon^i \,.\, (k_Z^j \,.\, Z_m^{ij} + O_m^{ij})\right]^2}{D_m^{ij}} + \sum_k \frac{(k_X - 1)^2}{\sigma_X^2} \quad (7.2)$$

où m est le bin de la masse, i est l'un des quatre canaux, et j est le nombre de jets. D est le nombre des données, Z est le nombre des évènements Z+jets, O est le nombre des évènements dibosons et $t\bar{t}$, et Q est le nombre des évènements multijets. k_L, k_ϵ^i, et k_Z^j sont respectivement les facteurs de correction de la luminosité, des facteurs d'échelle (SF) dans chaque canal, et de la section efficace de production des évènements $Z + jets$ pour chaque multiplicité de jet. Enfin, α^{ij} sont des facteurs libres permettant de déterminer le nombre des évènements multijets pour chaque canal et chaque multiplicité de jet. Comme le nombre de paramètres à déterminer directement des données est trop important, une contrainte externe de la forme $(k_X - 1)^2/\sigma_X^2$ est nécessaire, k_X est la notation générique des k_L, k_ϵ^i, k_Z^j et les σ_X sont leurs erreurs respectives.
La section efficace de production du processus $Z + 0\ jet$ est bien connue et donc son paramètre k_Z^0 est fixé à 1. Le facteur k_L ne peut pas être déterminé indépendemment des facteurs k_ϵ^i et k_L, il a donc aussi été fixé à 1. On a relâché cette contrainte plus tard pour déterminer l'erreur systématique.
Le bruit de fond multijet domine aux basses masses du Z ($M_Z < 60$ GeV). L'ajustement est appliqué en deux étapes, d'abord, il est réalisé avec des paramètres libres sur l'intervalle de masse compris entre 40 GeV et 150 GeV. Puis le facteur du bruit de fond multijet, α^{ij} est fixé et l'ajustement est réalisé sur l'intervalle de masse compris entre 60 GeV et 150 GeV.
L'ajustement est effectué séparément pour les Runs IIa et IIb, et pour les topologies di-électrons CC-CC et EC-CC. Les résultats sont regroupés dans les tableaux 7.5 (7.6) pour les Runs IIa et IIb pour les quatre analyses (pour l'analyse ee CC-EC). Les erreurs représentent seulement les erreurs statistiques. Les erreurs systématiques sont données en 8.2.1. Le résultat de la normalisation est illustré sur la figure 7.7 pour les analyses ee et ee^{icr}.

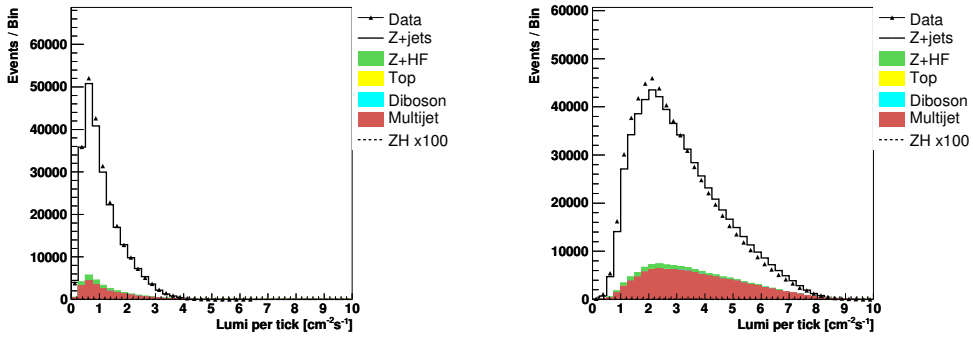


(a) jet_1^{η} *correction*



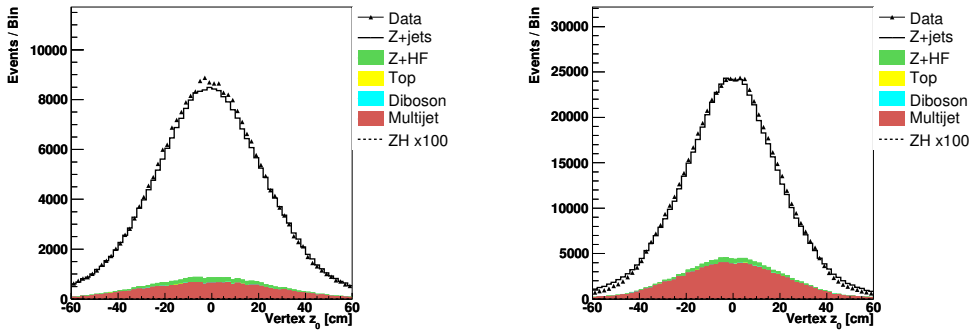
(b) $\Delta R(jet_1, jet_2)$ *correction*

FIG. 7.4 – Corrections appliquées au jet le plus énergétique en fonction de η 7.4(a), et les corrections appliquées à la distance dans le plan (η, ϕ) entre les deux jets les plus énergétiques 7.4(b).



(a) *RunIIa* (b) *RunIIb*

FIG. 7.5 – Distribution du profil de la luminosité instantanée, après correction, pour le Run IIa 7.5(a) et les Run IIb 7.5(a).



(a) *RunIIa* (b) *RunIIb*

FIG. 7.6 – Distribution du vertex primaire du boson Z, après correction, pour le Run IIa 7.6(a) et les Runs IIb 7.6(a).

Run IIa

i	$\mu\mu$	ee (CC-CC)	$\mu\mu^{trace}$	ee^{icr}
k_ϵ^i	0.956 ± 0.0038	1.06 ± 0.0043	0.995 ± 0.0105	1.03 ± 0.0071
α^{i0}	3.01 ± 0.265	0.934 ± 0.0121	1	1.4 ± 0.07
α^{i1}	0.238 ± 0.0363	0.781 ± 0.0265	1	0.567 ± 0.099
α^{i2}	0.00928 ± 0.00378	0.593 ± 0.0246	1	0.622 ± 0.125

Runs IIb

i	$\mu\mu$	ee (CC-CC)	$\mu\mu^{trace}$	ee^{icr}
k_ϵ^i	0.894 ± 0.0023	0.96 ± 0.0024	0.928 ± 0.0068	0.951 ± 0.0042
α^{i0}	1.89 ± 0.0953	0.67 ± 0.003	1	0.673 ± 0.0179
α^{i1}	0.198 ± 0.0215	0.608 ± 0.0087	1	0.398 ± 0.0263
α^{i2}	0.0112 ± 0.00255	0.546 ± 0.0115	1	0.324 ± 0.0331

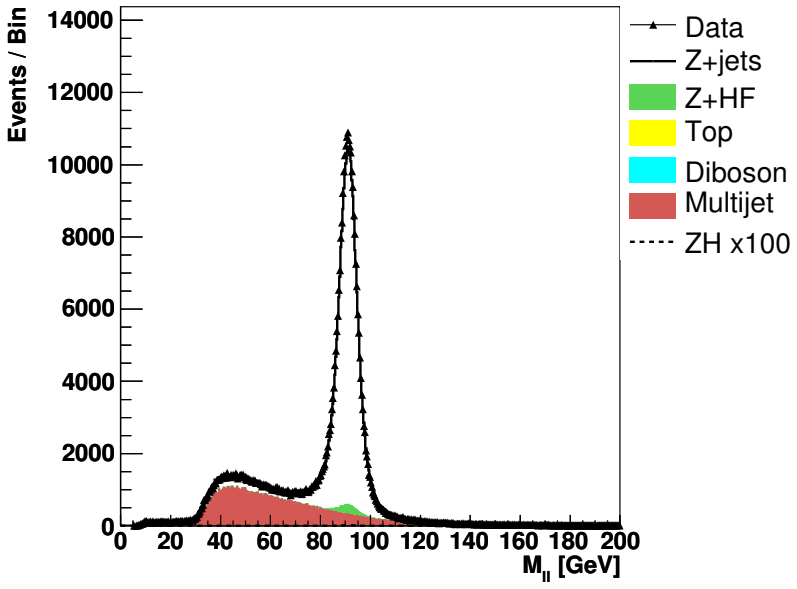
	Run IIa	Runs IIb
k_Z^0	1	1
k_Z^1	0.957 ± 0.0071	0.924 ± 0.0046
k_Z^2	1.11 ± 0.012	1.04 ± 0.009

TAB. 7.5 – Résultat de l'ajustement combiné pour les quatre analyses, avec ee (CC-CC), et pour les Runs IIa et IIb.

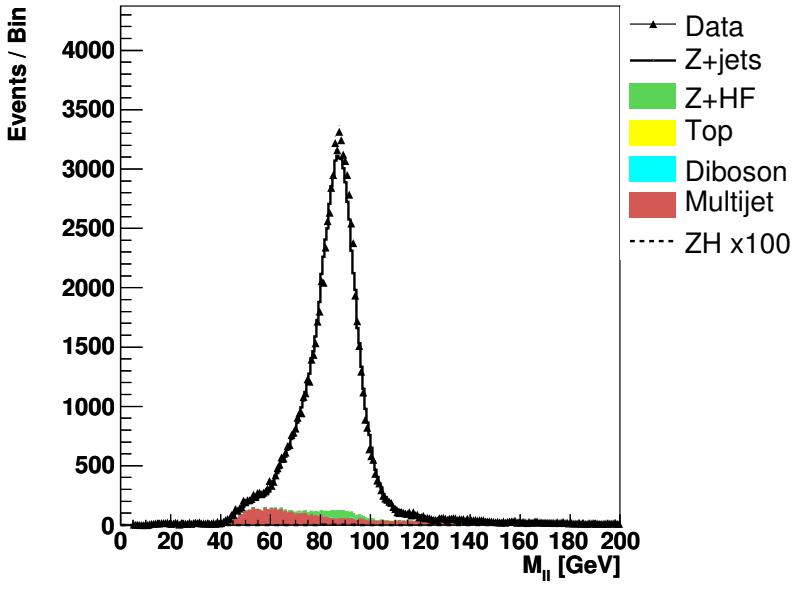
i	Run IIa ee (CC-EC)	Runs IIb ee (CC-EC)
k_ϵ^i	1.02 ± 0.00541	0.924 ± 0.00345
α^{i0}	0.872 ± 0.0103	0.673 ± 0.00352
α^{i1}	0.769 ± 0.0255	0.601 ± 0.0089
α^{i2}	0.6 ± 0.0249	0.519± 0.0118

	Runs IIa et IIb
k_Z^0	1
k_z^1	0.95 ± 0.00961
k_Z^2	1.2 ± 0.0211

TAB. 7.6 – Résultat de l'ajustement combiné pour l'analyse ee (CC-EC) pour les Runs IIa et IIb.



(a) M^{ee}

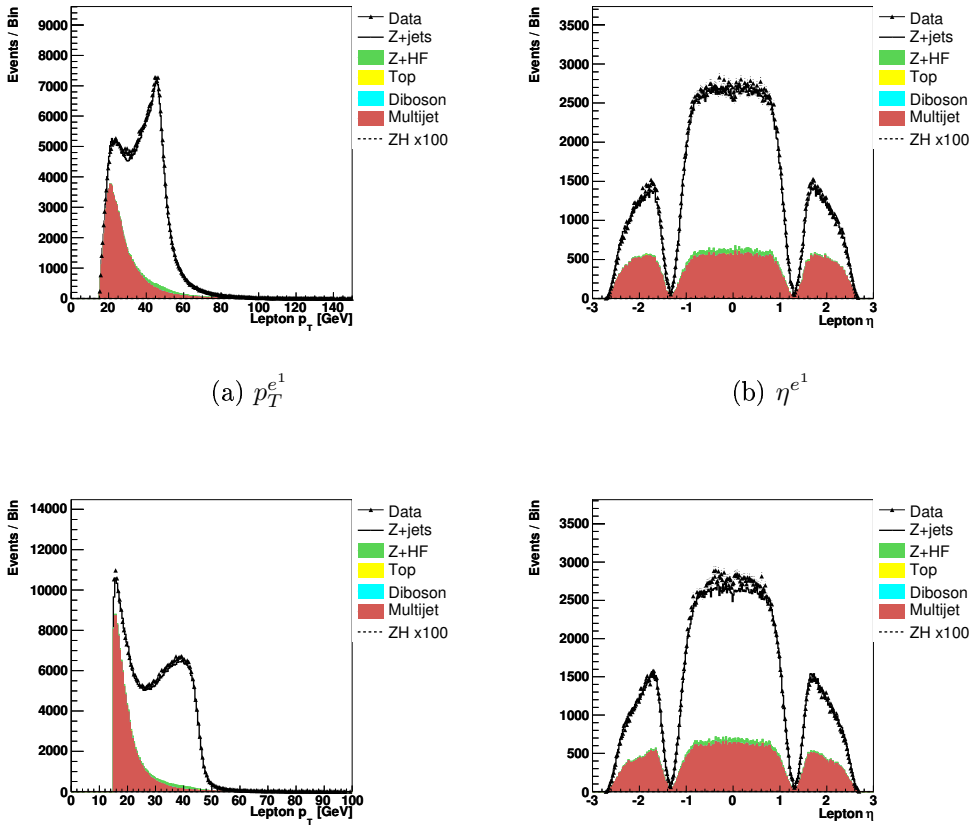


(b) $M^{ee^{icr}}$

FIG. 7.7 – Distribution de la masse invariante des deux électrons pour l'analyse ee 7.7(a), et ee^{icr} 7.7(b), après la normalisation de tous les bruits de fond aux données.

7.6 Distributions de contrôle avant étiquetage des jets de b

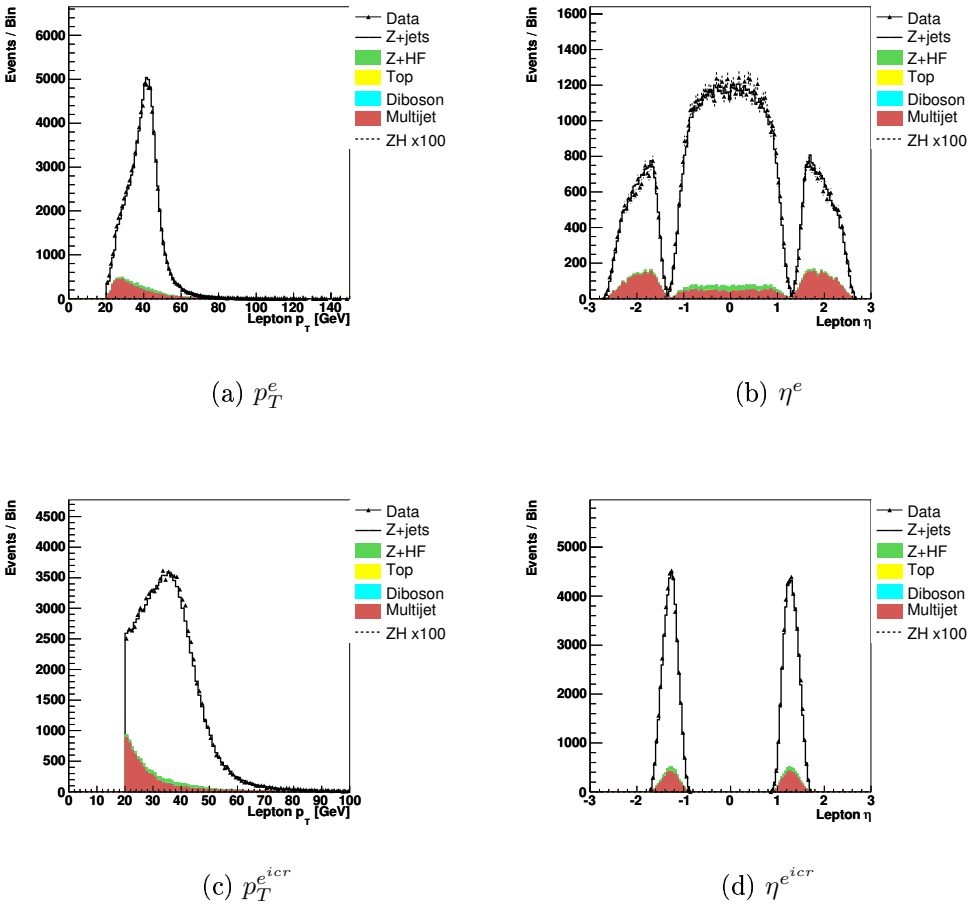
Après les étapes de la pré-sélection (cf. 7.2.3), des différentes corrections appliquées à la simulation (cf. 7.3.2), et de la sélection du bruit de fond multijet (cf. 7.4), les données peuvent être comparées à la simulation.
Les tables 7.7 et 7.8 regroupent le nombre d'évènements sélectionnés pour les analyses ee et ee^{icr} respectivement. On y trouve entre autres, pour les Runs IIa et IIb, le nombre des données, des différents bruits de fond, ainsi que celui du signal ZH pour une masse de 115 GeV, après avoir appliqué les coupures de pré-sélection, cf. 7.2.3 et 7.4.
Les figures 7.8 et 7.9 montrent la distribution de p_T et de η pour les deux électrons de plus haute impulsion transverse, respectivement pour le cas ee et ee^{icr}. La figure 7.10 montre la distribution de la masse invariante et de l'impulsion transverse du boson Z, pour le cas ee et ee^{icr}. Comme les analyses ee et ee^{icr} étudiées séparément présentent la même topologie pour les jets, il est judicieux de combiner les résultats pour avoir plus de statistique et donc pour mieux juger de l'accord entre les données et la simulation. Ainsi, la figure 7.11 montre la distribution de l'impulsion transverse et de η des deux jets les plus énergétiques, pour le cas où les analyses ee et ee^{icr} sont combinées. La figure 7.12 montre la distribution de la distance entre les deux jets ΔR, et celle de leur masse invariante pour le cas où les analyses ee et ee^{icr} sont combinées. L'ensemble de nos distributions montrent qu'après correction, notre simulation décrit correctement nos données.



(a) $p_T^{e^1}$ (b) η^{e^1}

(c) $p_T^{e^2}$ (d) η^{e^2}

FIG. 7.8 – Distribution de l'impulsion transverse p_T et de la distribution angulaire η, pour l'électron de plus haute impulsion transverse, respectivement 7.9(a) et 7.9(b), et pour le deuxième électron de plus haute impulsion transverse, respectivement 7.9(c) et 7.9(d), dans l'analyse ee. Aucune coupure sur la masse du Z ou sur le nombre de jets n'est appliquée.



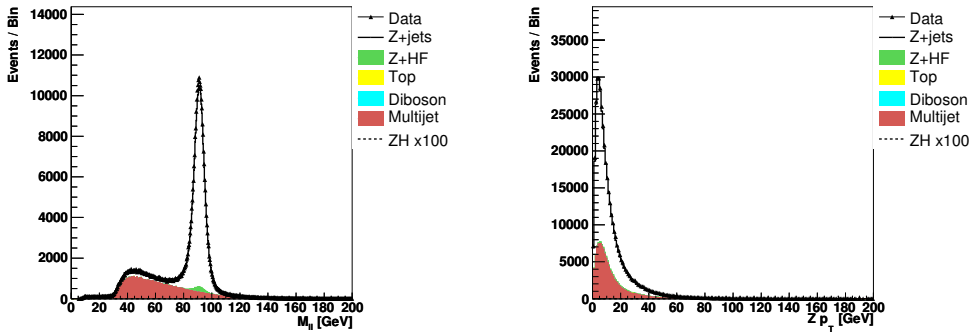
(a) p_T^e

(b) η^e

(c) $p_T^{e^{icr}}$

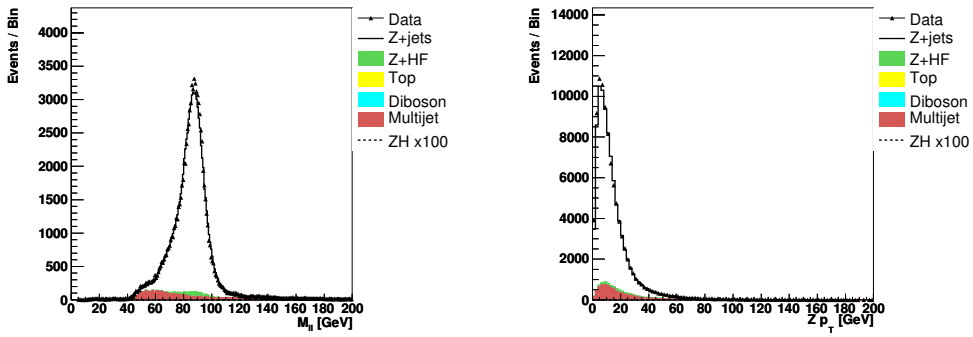
(d) $\eta^{e^{icr}}$

FIG. 7.9 – Distribution de p_T et de η pour l'électron du calorimètre, respectivement 7.9(a) et 7.9(b), et pour l'électron de l'ICR, respectivement 7.9(c) et 7.9(d), dans l'analyse ee^{icr}. Aucune coupure sur la masse du Z ou sur le nombre de jets n'est appliquée.



(a) M^{ee}

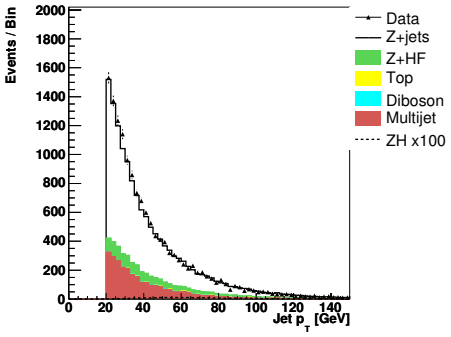
(b) p_T^{ee}



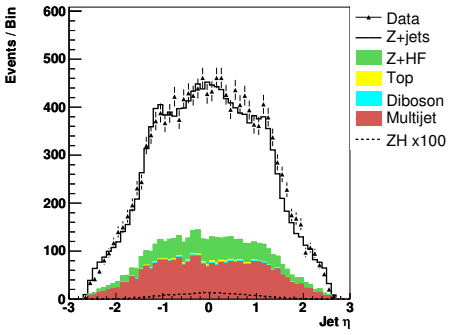
(c) $M^{ee^{icr}}$

(d) $p_T^{ee^{icr}}$

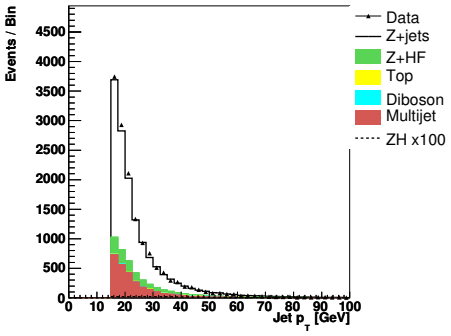
FIG. 7.10 – Distribution de la masse invariante et de l'impulsion transverse des deux électrons pour l'analyse ee, respectivement 7.10(a) et 7.10(b), et pour l'analyse ee^{icr}, respectivement 7.10(c) et 7.10(d). Aucune coupure sur la masse du Z ou sur le nombre de jets n'est appliquée.



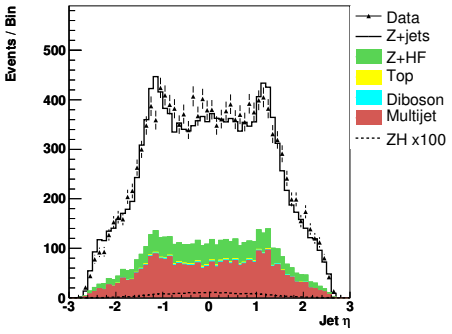
(a) $p_T^{jet^1}$ (ee, ee^{icr})



(b) η^{jet^1} (ee, ee^{icr})

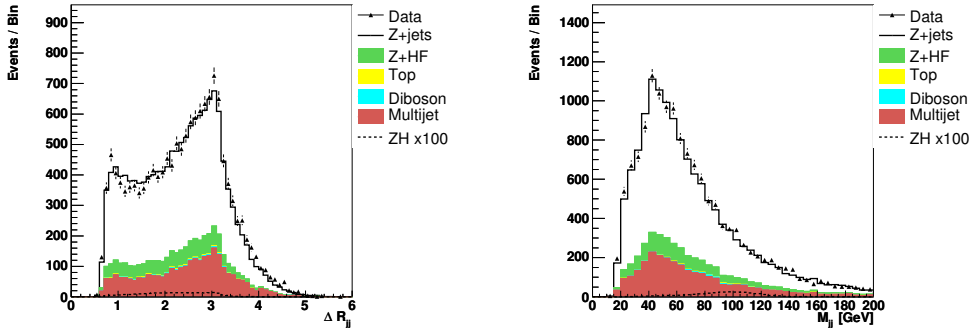


(c) $p_T^{jet^2}$ (ee, ee^{icr})



(d) η^{jet^2} (ee, ee^{icr})

FIG. 7.11 – Distribution de p_T et de η du jet le plus énergétique respectivement 7.11(a) et 7.11(b), et du second jet le plus énergétique respectivement 7.11(c) et 7.11(d), pour le cas où les analyses ee et ee^{icr} sont combinées. On demande au moins deux jets dans les évènements et la masse du Z doit être comprise entre 60 GeV et 150 GeV.



(a) $\Delta R^{jet^{1,2}}$ $(ee + ee^{icr})$ (b) $M^{jet^{1,2}}$ $(ee + ee^{icr})$

FIG. 7.12 – Distribution de la distance entre les deux jets 7.12(a), et de la masse invariante des deux jets 7.12(b), pour le cas où les analyses ee et ee^{icr} sont combinées. On demande au moins deux jets dans les évènements et la masse du Z doit être comprise entre 60 GeV et 150 GeV.

7.7 Etiquetage des jets de b

7.7.1 Principe de fonctionnement du b-tagging

Pour réduire les bruits de fond (cf. 7.3), et augmenter la sensibilité de notre signal, les jets de b sont étiquetés avec une méthode appelée "b-tagging". Cette technique est basée sur le temps de vie relativement long du quark b. Ainsi, les jets de b peuvent être différenciés des jets de quarks légers à l'aide des indications suivantes : les traces provenant de la désintégration du quark b sont issus du vertex secondaire, et la distance d'approche minimale entre la trace chargée et le vertex principal (dit paramètre d'impact) est élevée, voir figure 7.13.
Notons que le "b-tagging" n'est appliqué qu'aux jets respectant les critères d'étiquetabilité [100] [101], dans ce cas, le jet est dit étiquetable (ou taggable). Pour être étiquetable, le jet doit, entre autres, avoir une impulsion transverse supérieure à 15 GeV, être dans le région $|\eta^{jet}| < 2.5$, et être associé à au moins deux traces de bonnes qualités.

En général, l'étiquetage des jets de b est effectué avec un réseau de neurones adapté (dit NN-tagger) [102] [103]. D'un point de vue général, le réseau de neurones est un modèle de calcul qui combine de manière statistique plusieurs variables d'entrée dans une variable de sortie. Il est utilisé pour construire un discriminant qui permet une séparation entre un signal et un bruit de fond. Le NN-tagger utilisé prend comme variable d'entrée les résultats de trois algorithmes qui utilisent les informations liées à la longue durée de vie des quarks b : L'algorithme SVT (Secondary Vertex Tagger) [104] [105] qui est basé sur la reconstruction des vertex secondaires, et les algorithmes JLIP (Jet LIfetime Probability) [106] [107] et CSIP (Counting Signed Impact Parameter) [108] [109] qui utilisent les paramètres d'impact des traces des jets. A la sortie du NN-tagger, le jet est caractérisé avec un nombre compris entre 0 et 1, et

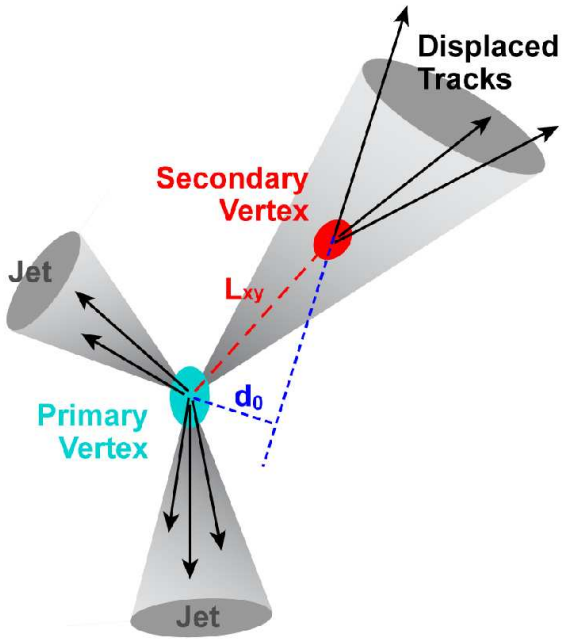


FIG. 7.13 – Schéma des propriétés caractéristiques d'un jet de b. Contrairement aux jets légers, les traces reconstruites du jet de b proviennent du vertex secondaire (Secondary Vertex). Ce dernier est distant du vertex premier (Primary Vertex) de L_{xy}. d_0 est le paramètre d'impact.

le jet ressemblera d'autant plus à un jet de b (resp. jet léger) que ce nombre sera proche de 1 (resp. 0). On définit donc des points d'opération, c'est-à-dire une valeur requise de la variable discriminante qui correspond à la probabilité que le jet soit étiqueté. Un jet est étiqueté Loose (ou L : coupure lâche) si la variable de sortie du NN-tagger pour ce jet est supérieure à 0.20. Le jet sera étiqueté Very Tight (ou VT : coupure très serrée) si cette même variable est supérieure à 0.85. Notons que l'efficacité d'étiquetage du point d'opération (noté OP : Operating Point) VT (L) est d'environ 48% (72%) et que le taux de mauvaise identification est d'environ 0.5% (5%), cf. figure 7.14.

7.7.2 Le b-tagging dans l'analyse ZH

Dans cette analyse où le Higgs se désintègre en $b\bar{b}$, les évènements $Z + jets$ légers constituent le bruit de fond dominant lors de la pré-sélection. Les jets de b ont donc été étiquetés avec la méthode du b-tagging décrite ci-dessus. Deux topologies ont été utilisées : la première consiste à sélectionner les évènements avec au moins deux jets de b étiquetés chacun avec le point d'opération Loose (noté 2L) ; la deuxième consiste à sélectionner, parmis les évènements qui ne l'ont pas été précédemment, ceux ayant un seul jet de b étiqueté avec le point d'opération Very Tight (noté 1VT). Par la suite, les autres évènements qui n'ont pas été sélectionnés seront rejetés, et les deux échantillons 1VT et 2L, indépendants, seront analysés séparément.

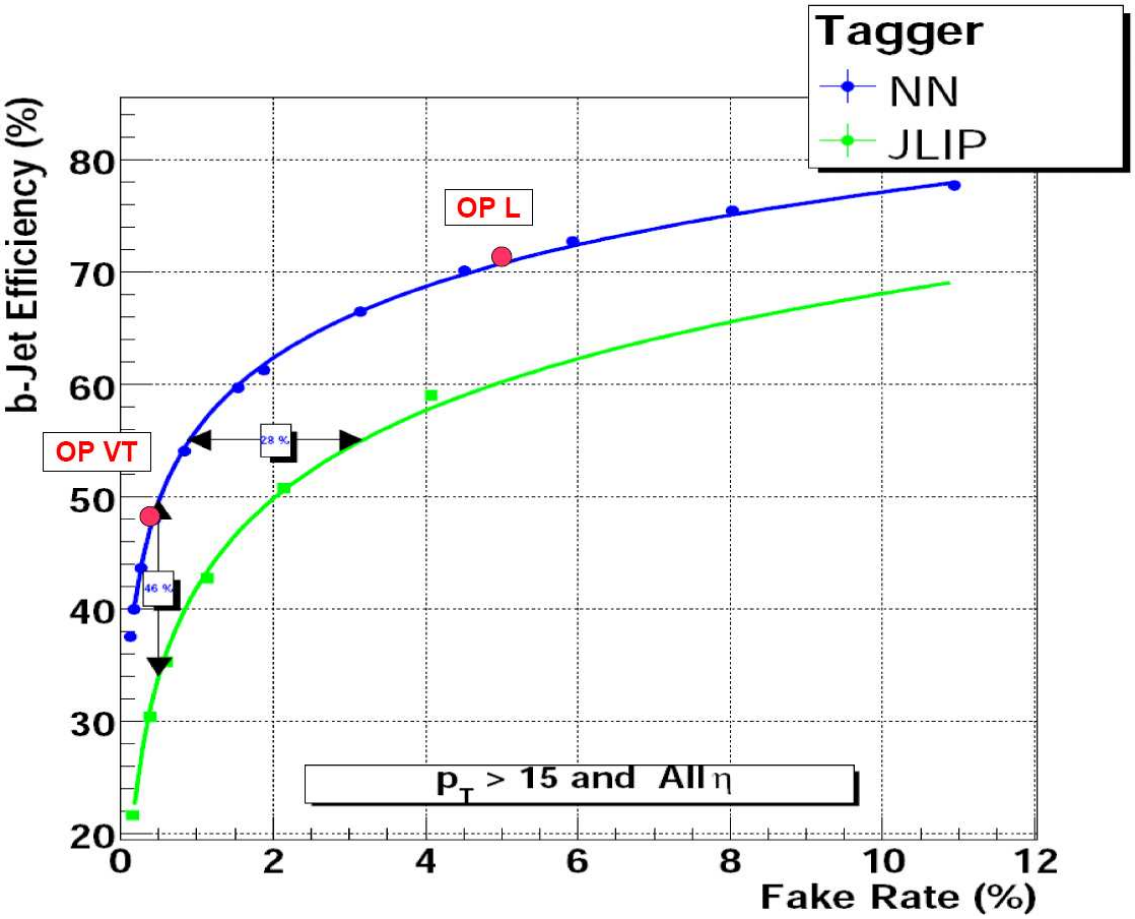


FIG. 7.14 – Efficacité d'étiquetage des jets (b-Jet Efficiency) de b en fonction du taux de mauvaise identification (Fake Rate), pour l'algorithme d'étiquetage JLIP seul (courbe du bas, en verte) et le réseau de neurones (courbe du haut, en bleue), pour tous les jets de $p_T > 15$ GeV se trouvant dans toutes les régions du détecteur. Les points d'opération Loose (OP L) et Very Tight (OP VT) sont également représentés.

Les données

Pour étiqueter les jets des données, la procédure du NN-tagger à directement été appliquée sur les jets taggables. Lors de la reconstruction de la masse du boson de Higgs dans les données ($H \rightarrow b\bar{b}$), si les évènements n'ont qu'un seul jet étiqueté Very Tight, alors le jet non étiqueté de plus haute impulsion transverse est utilisé comme second jet. Si les évènements ont plus de deux jets étiquetés Loose, alors le choix porte sur les deux jets étiquetés Loose de plus haute impulsion transverse.

La simulation et le bruit de fond multijet

Comme la statistique des évènements simulés et de bruits de fond multijets est faible, au lieu d'utiliser le NN-tagger, on utilise une fonction qui calcule la probabilité que les jets taggables soient étiquetés, puis le jet est pondéré par cette probabilité. Ainsi aucun évènement n'est rejeté au moment de l'étiquetage. Cette fonction est appelée TRF (ou Tag Rate Fonction, pour fonction de taux d'étiquetage), et elle est para-

métrée en fonction de p_T et η du jet. Il peut ainsi y avoir plusieurs façons de choisir deux jets étiquetés Loose, ou un seul jet étiqueté Very Tight dans les évènements. On peut démontrer que [110] :

– soit L_i (L_j) la probabilité que le jet i (j) soit étiqueté Loose, alors pour une combinaison $1 \leq i < j \leq n$, où n est le nombre de jets dans l'évènement et i et j représentent l'ordre des indices du p_T des jets, la probabilité de sélectionner une combinaison de deux jets Loose est donnée par :

$$p(i,j) = \frac{L_i L_j}{(1-L_i)(1-L_j)} \prod_{k=1}^{j} (1-L_k). \tag{7.3}$$

– la probabilité de sélectionner une combinaison de deux jets tel que un seul jet soit étiqueté Very Tight est donnée par :

$$p(i,j) = \left(\prod_{k=1}^{n} (1-L_k) \right) \left(\frac{T_j}{1-L_j} + \delta_{j2} \cdot \frac{T_1}{1-L_1} \right) \cdot \delta_{i1}, \tag{7.4}$$

T_i est la probabilité que le jet soit étiqueté Very Tight et δ_{ij} vaut 1 quand $i = j$ et 0 autrement.

7.7.3 Correction de l'étiquetabilité des jets

L'étiquetabilité des jets (ou taux d'étiquetage des jets) correspondant au rapport entre le nombre de jets étiquetables et le nombre total de jets, a été calculé pour les données et la simulation. Comme le taux de jet étiquetable est meilleur dans cette dernière, il est corrigé avec un facteur d'échelle pour que sa valeur soit la même que celle des données. Le facteur de correction correspond au rapport entre le taux d'étiquetage des jets des données et celui de la simulation, et il a été paramétré en fonction de p_T et η du jet, ainsi que de la position suivant l'axe z du vertex primaire.

7.8 Distributions de contrôle après étiquetage des jets de b

Après les étapes de la pré-sélection (cf. 7.2.3), des différentes corrections appliquées à la simulation (cf. 7.3.2), de la sélection du bruit de fond multijet (cf. 7.4), et après l'étiquetage des jets de b en deux lots 1VT et 2L (cf. 7.7.2), nous contrôlons de nouveau l'accord entre les distributions des données et de la simulation.
Les tableaux 7.7 et 7.8 montrent le nombre d'évènements sélectionnés pour différentes coupures, respectivement pour l'analyse ee et ee^{icr}, pour les Runs IIa et IIb. La figure 7.15 montre la distribution de la masse invariante des deux jets pour ces deux lots 1VT et 2L. Pour avoir plus de statistique et donc mieux juger de l'accord entre les données et la simulation, le cas où les quatre analyses (cf. 7) ee, ee^{icr}, $\mu\mu$, et $\mu\mu^{trk}$ ont été combinées sert d'illustration.

Analyse ee Run IIa

	Inclusive	2j-multijet	2j-pretag	1VT	2L
Data	99129	4424	3771	88	40
All bkg	98266 ± 150	4224 ± 27	3644 ± 20	84.48 ± 0.57	44.39 ± 0.31
Multijet	16877 ± 117	1061 ± 24	618 ± 17	13.02 ± 0.40	4.46 ± 0.14
Zjj	79181 ± 93	2723 ± 12	2610 ± 10	18.899 ± 0.065	8.946 ± 0.029
$Zb\bar{b}$	556.7 ± 2.4	117.0 ± 1.2	112.7 ± 1.2	30.23 ± 0.36	16.06 ± 0.25
$Zc\bar{c}$	1504.1 ± 5.4	265.5 ± 2.4	253.5 ± 2.3	15.98 ± 0.18	7.34 ± 0.10
ZZ	30.13 ± 0.31	15.18 ± 0.22	14.86 ± 0.22	1.193 ± 0.036	1.152 ± 0.041
WZ	40.24 ± 0.56	17.78 ± 0.37	17.27 ± 0.37	0.642 ± 0.022	0.2195 ± 0.0078
WW	46.49 ± 0.54	2.66 ± 0.13	1.82 ± 0.11	0.0401 ± 0.0049	0.0159 ± 0.0023
$t\bar{t}$	29.53 ± 0.18	22.88 ± 0.16	15.77 ± 0.13	4.471 ± 0.040	6.200 ± 0.058
ZH (115)	0.8804 ± 0.0059	0.6817 ± 0.0053	0.6719 ± 0.0052	0.1888 ± 0.0015	0.2491 ± 0.0022

Analyse ee Runs IIb

	Inclusive	2j-multijet	2j-pretag	1VT	2L
Data	307005	10018	8155	229	149
All bkg	304738 ± 405	9801 ± 59	7963 ± 47	213.6 ± 1.3	136.98 ± 0.88
Multijet	90183 ± 266	3586 ± 47	2037 ± 31	48.80 ± 0.79	26.13 ± 0.40
Zjj	208673 ± 304	5258 ± 35	5027 ± 35	37.97 ± 0.26	29.77 ± 0.19
$Zb\bar{b}$	1481.9 ± 6.4	258.3 ± 2.9	249.3 ± 2.8	70.21 ± 0.90	41.29 ± 0.67
$Zc\bar{c}$	3992 ± 15	560.8 ± 5.9	534.8 ± 5.7	40.47 ± 0.50	21.12 ± 0.29
ZZ	82.20 ± 0.88	33.13 ± 0.57	32.14 ± 0.56	2.667 ± 0.094	2.60 ± 0.11
WZ	111.1 ± 1.5	39.59 ± 0.89	38.86 ± 0.89	1.606 ± 0.058	0.697 ± 0.024
WW	131.8 ± 2.9	6.38 ± 0.63	4.12 ± 0.49	0.147 ± 0.039	0.078 ± 0.027
tt	83.22 ± 0.52	58.98 ± 0.43	40.61 ± 0.36	11.69 ± 0.11	15.30 ± 0.15
ZH (115)	2.475 ± 0.020	1.754 ± 0.017	1.725 ± 0.016	0.4924 ± 0.0049	0.6337 ± 0.0066

TAB. 7.7 – Nombre d'évènements dans l'analyse ee pour les Runs IIa (en haut) et IIb (en bas) pour : les données (Data), la somme tous les bruits de fond (All bkg), le bruit de fond QCD (Multijet), les différents bruits de fond séparément, et le signal ZH simulé pour une masse de 115 GeV. "Inclusive" correspond au nombre des évènements après que les coupures de la pré-sélection aient été appliquées ; il n'y a pas de condition sur le nombre de jets, ni sur la masse du boson Z. "2j-multijet" correspond au nombre des évènements de la pré-sélection ayant au moins deux jets ; il n'y a pas de condition sur la masse du boson Z. "2j-pretag" correspond au nombre des évènements de la pré-sélection ayant au moins deux jets, et on demande que $60 < M_{ee} < 150$ GeV. "1VT" est le nombre des évènements n'ayant qu'un seul jet étiqueté Very Tight et aucun autre étiqueté Loose. "2L" est le nombre des évènements ayant au moins deux jets étiquetés Loose.

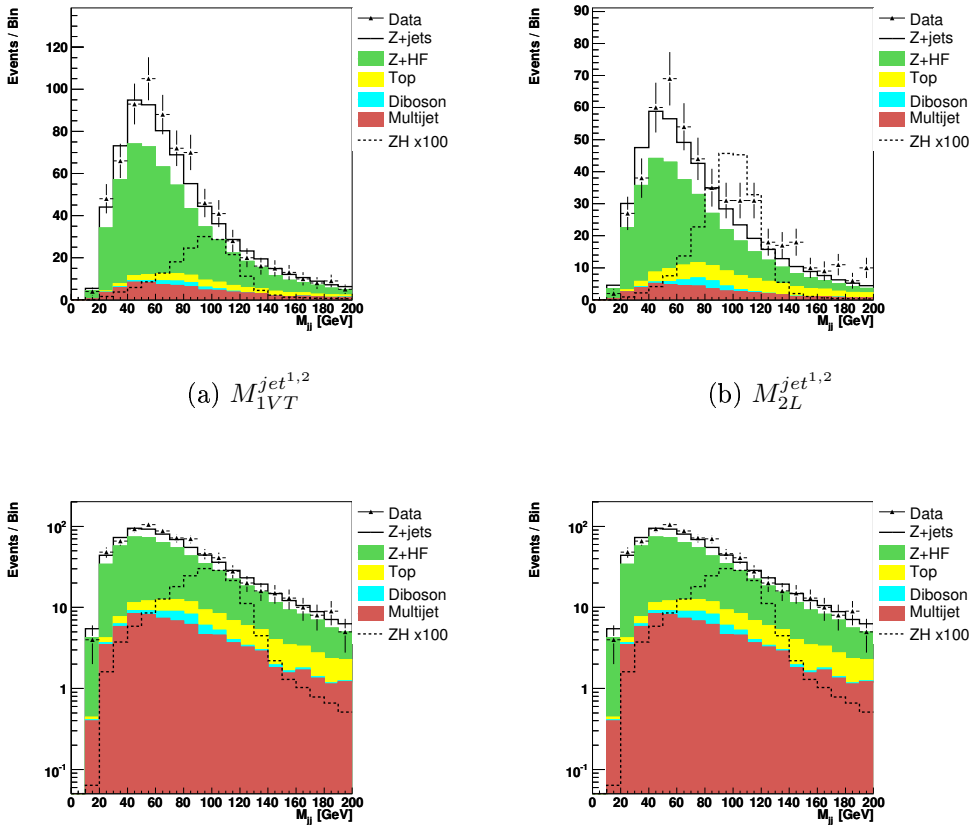
Analyse ee^{icr} Run IIa

	Inclusive	2j-multijet	2j- pretag	1VT	2L
Data	27311	1050	935	26	3
All bkg	26689 ± 65	951.9 ± 8.9	880.5 ± 7.6	18.52 ± 0.21	9.68 ± 0.14
Multijet	1854 ± 48	80.2 ± 7.1	47.9 ± 5.5	0.627 ± 0.077	0.565 ± 0.064
Zjj	24239 ± 44	756.0 ± 5.3	723.5 ± 5.1	4.874 ± 0.031	2.209 ± 0.014
$Zb\bar{b}$	149.6 ± 1.2	30.50 ± 0.58	29.14 ± 0.56	7.52 ± 0.17	3.44 ± 0.10
$Zc\bar{c}$	406.9 ± 2.6	70.5 ± 1.2	67.3 ± 1.2	3.989 ± 0.086	1.753 ± 0.048
ZZ	8.40 ± 0.16	4.20 ± 0.12	4.06 ± 0.12	0.325 ± 0.020	0.349 ± 0.027
WZ	11.89 ± 0.29	4.85 ± 0.19	4.67 ± 0.19	0.1626 ± 0.0093	0.0537 ± 0.0029
WW	12.07 ± 0.27	0.645 ± 0.063	0.428 ± 0.052	0.0183 ± 0.0050	0.0127 ± 0.0067
tt	6.249 ± 0.082	4.905 ± 0.072	3.532 ± 0.061	1.006 ± 0.019	1.303 ± 0.026
ZH (115)	0.2039 ± 0.0028	0.1606 ± 0.0025	0.1545 ± 0.0024	0.04342 ± 0.00072	0.05330 ± 0.00097

Analyse ee^{icr} Runs IIb

	Inclusive	2j-multijet	2j- pretag	1VT	2L
Data	65119	1893	1676	48	37
All bkg	64329 ± 168	1712 ± 19	1567 ± 18	41.10 ± 0.51	27.69 ± 0.37
Multijet	6416 ± 63	237.8 ± 8.8	155.4 ± 7.1	4.11 ± 0.20	2.17 ± 0.10
Zjj	56510 ± 156	1258 ± 16	1209 ± 16	9.21 ± 0.12	7.306 ± 0.091
$Zb\bar{b}$	351.9 ± 2.9	58.7 ± 1.3	56.3 ± 1.3	15.47 ± 0.39	9.54 ± 0.31
$Zc\bar{c}$	957.6 ± 6.5	126.7 ± 2.6	120.6 ± 2.5	9.01 ± 0.21	4.71 ± 0.13
ZZ	20.14 ± 0.42	7.76 ± 0.26	7.36 ± 0.26	0.593 ± 0.042	0.698 ± 0.062
WZ	28.90 ± 0.71	9.76 ± 0.41	9.31 ± 0.40	0.370 ± 0.022	0.1506 ± 0.0084
WW	29.9 ± 1.2	1.53 ± 0.33	0.96 ± 0.26	0.021 ± 0.011	0.0065 ± 0.0025
tt	15.22 ± 0.21	11.14 ± 0.18	8.13 ± 0.15	2.323 ± 0.045	3.117 ± 0.064
ZH (115)	0.5204 ± 0.0085	0.3678 ± 0.0072	0.3530 ± 0.0071	0.1009 ± 0.0021	0.1286 ± 0.0028

TAB. 7.8 – Nombre d'évènements dans l'analyse ee pour les Runs IIa (en haut) et IIb (en bas) pour : les données (Data), la somme tous les bruits de fond (All bkg), le bruit de fond QCD (Multijet), les différents bruits de fond séparément, et le signal ZH simulé pour une masse de 115 GeV. "Inclusive" correspond au nombre des évènements après que les coupures de la pré-sélection aient été appliquées ; il n'y a pas de condition sur le nombre de jets, ni sur la masse du boson Z. "2j-multijet" correspond au nombre des évènements de la pré-sélection ayant au moins deux jets ; il n'y a pas de condition sur la masse du boson Z. "2j-pretag" correspond au nombre des évènements de la pré-sélection ayant au moins deux jets, et on demande que $60 < M_{ee} < 150$ GeV. "1VT" est le nombre des évènements n'ayant qu'un seul jet étiqueté Very Tight et aucun autre étiqueté Loose. "2L" est le nombre des évènements ayant au moins deux jets étiquetés Loose.



(a) $M_{1VT}^{jet^{1,2}}$

(b) $M_{2L}^{jet^{1,2}}$

(c) $M_{1VT}^{jet^{1,2}}$

(d) $M_{2L}^{jet^{1,2}}$

FIG. 7.15 – Distribution de la masse invariante des deux jets dans l'échantillon 1VT 7.15(a) et 7.15(c), et dans l'échantillon 2L 7.15(b) et 7.15(d), dans le cas où les quatre analyses ee, ee^{icr}, $\mu\mu$ et $\mu\mu^{trk}$ ont été combinées. Les figures du haut (resp. bas) sont en échelle linéaire (resp. logarithmique). Le signal ZH simulé pour une masse de 115 GeV, et multiplié par un facteur 100, est représenté en pointillé.

7.9 Ajustement cinématique

7.9.1 Principe

Une imparfaite détermination de l'énergie des leptons et des jets entraîne de l'énergie transverse manquante dans les évènements. Pour résoudre ce problème, et avoir la meilleure information possible sur les variables cinématiques, on fait un ajustement cinématique sur la mesure de l'énergie et des angles des leptons et des jets, avant et après l'étape du b-tagging. L'ajustement cinématique consiste à minimiser un χ^2, en permettant aux énergies et aux angles des leptons et des jets de fluctuer sous les contraintes suivantes :

- la masse des deux électrons doit être égale à 91.2 $\pm$ 2.5 GeV ;
- les sommes des impulsions du Z et des deux jets, suivant les axes x et y doivent être égale à 0 $\pm$ 7 GeV.

Pour satisfaire ces contraintes, on utilise la résolution angulaire, ainsi que la fonction de résolution de l'énergie des leptons et des jets pour contraindre l'ajustement.

La résolution angulaire des leptons et des jets
La résolution sur les angles des leptons et des jets est montrée dans le tableau 7.9.

La résolution en énergie du calorimètre
La résolution du calorimètre en fonction de l'énergie mesurée E en GeV, s'exprime de la manière suivante :

$$\frac{\sigma(E)}{E} = \sqrt{\left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2} \tag{7.5}$$

N est un terme de bruit relié à l'électronique d'acquisition, des effets de l'empilement des évènements, et de la radioactivité de l'uranium ; C est un terme constant de calibration qui tient compte des non-uniformités mécaniques et électroniques ; S est un terme d'échantillonage, il représente les fluctuations d'énergie déposée, et rend compte de la non-uniformité de l'échantillonnage du calorimètre, due à des effets d'irrégularités géométriques.
Les paramètres C, S et N ont été estimés pour les Runs IIa et IIb pour la partie électromagnétique (électrons et photons), et pour la partie hadronique (jets), pour les différentes régions du détecteur.
Pour la partie électromagnétique [111], les termes sont identiques pour les deux Runs, C_{em} vaut respectivement 0.020, 0.027 et 0.032 dans les CC, EC nord et EC sud du calorimètre, et $N = 0.4$. La valeur typique de S pour un électron de 45 GeV pour $\eta_{det} = 0$ est de 0.157.
Pour la partie hadronique [112], les paramètres sont regroupés dans le tableau 7.10 pour les Runs IIa et IIb.

	η	ϕ
Lepton	0.002	0.001
Jet	0.08	0.08

TAB. 7.9 – Résolution angulaire des leptons et des jets.

Run IIa

	N_{had}	S_{had}	C_{had}
$\lvert\eta_{det}\rvert < 0.8$	3.261	1.228	0.048
$0.8 < \lvert\eta_{det}\rvert < 1.6$	4.392	1.242	0.088
$1.6 < \lvert\eta_{det}\rvert < 2.4$	3.842	1.283	0.028
$2.4 < \lvert\eta_{det}\rvert < 3.6$	5.082	1.302	0.0

Runs IIb

	N_{had}	S_{had}	C_{had}
$\lvert\eta_{det}\rvert < 0.8$	3.550	1.144	0.062
$0.8 < \lvert\eta_{det}\rvert < 1.6$	6.322	0.814	0.110
$1.6 < \lvert\eta_{det}\rvert < 2.4$	2.920	1.302	0.0
$2.4 < \lvert\eta_{det}\rvert < 3.6$	6.075	1.108	0.063

TAB. 7.10 – Les coefficients N_{had}, S_{had}, et C_{had} de la fonction de résolution de l'énergie transverse des jets pour le Run IIa (en haut) et les Runs IIb (en bas), en fonction de la région du détecteur.

La résolution en énergie de l'électron dans l'ICR

La résolution en énergie de l'électron dans l'ICR est déterminée à partir de l'impulsion transverse p_T et de la résolution angulaire de la trace de l'électron. Elle s'exprime en fonction de η_{det} de la façon suivante :

$$\sigma(\frac{1}{p_T}) = \begin{cases} \sigma_0 + \frac{\sigma_1}{p_T} & |\eta_{det}| \leq 1.28 \\ \sqrt{(\sigma_0 + \frac{\sigma_1}{p_T})^2 + ((c_0 + \frac{c_1}{p_T}) * (|\eta_{det}| - |\eta_0|))^2} & |\eta_{det}| > 1.28 \end{cases}$$

η_0 vaut 1.28, σ_0 est la résolution de base, tandis que σ_1 prend en compte la dépendance sur $1/p_T$. Pour $|\eta_{det}| > 1.3$, il y a moins de couches de silicium dans le SMT, ce qui entraîne une plus grande perte de résolution : les termes c_i prennent en compte cet effet. Les coefficients sont les mêmes pour les Runs IIa et IIb et ne dépendent pas de la région du détecteur. σ_0 et σ_1 valent respectivement 0.002 et 0.010 ; c_0 et c_1 valent respectivement 0.007 et 0.049.

7.9.2 Résultats

Comme on peut le voir sur la figure 7.16, après l'ajustement cinématique, les bruits de fond sont déplacés vers les petites valeurs, ce qui permet une augmentation du rapport signal sur bruit de fond.

7.10 Analyse discriminante

Après l'étape de la pré-sélection, du b-tagging, et de l'ajustement cinématique, pour davantage augmenter la séparation du signal et du bruit de fond à l'aide des différentes informations cinématiques, on utilise un outil appelé Random Forest (RF), du package informatique nomé TMVA (pour Toolkit for Multivariate Data Analysis) [113]. Le RF est un ensemble d'arbres de décision de type binaire, construits de façon aléatoire. Il sert à classifier les évènements en "signal" ou "bruit de fond" à l'aide de leurs informations cinématiques.

7.10.1 Construction des arbres

Le RF utilisé dans cette analyse se compose de 200 arbres. Chaque arbre a été construit (ou entraîné) de la manière suivante, (voir fig. 7.17) :
soit un lot d'entraînement composé d'évènements simulés de signal et de bruit de fond, ils forment la racine de l'arbre. Partant d'une liste de 20 variables, cf 7.10.2, tous les évènements sont triés selon 10 variables choisies au hasard dans cette liste. Pour chacune des 10 variables choisies, l'algorithme évalue la valeur de décision qui donne la meilleure séparation des évènements en deux nœuds fils, un contenant des évènements de signal (noté P pour Pass) et l'autre des évènements de bruit de fond (noté F pour Fail). Le critère de division utilisé est le facteur de *Gini* [114] [115] défini par $Gini = 2\ .\ SB/(S+B)$, où S et B sont respectivement le nombre d'évènements de signal et de bruit de fond contenus dans le nœud.
L'algorithme est itéré aux deux nœuds fils de manière récursive, et la division se termine quand le nombre d'évènements disponible dans un nœud est inférieur à 20, ou quand le niveau de l'arbre excède 15 étages. Dans ce cas le nœud est appelé feuille. L'entraînement de l'arbre se poursuit jusqu'à ce que chaque nœud devienne une feuille.
Chaque feuille est ensuite étiquetée signal (ou bruit de fond) si sa pureté définie par $P = S/(S+B)$, est supérieure (ou inférieure) à 0.5. Les 200 arbres du RF ont été entraînés de la même façon, mais avec des combinaisons de 10 variables différentes choisies au hasard dans la liste. On injecte ensuite la totalité des évènements de données et de bruit de fond multijets, ainsi qu'une partie (la moitié) des évènements simulés à la racine de l'arbre. Cette opération est effectuée sur les 200 arbres. Suivant ses informations cinématiques, les évènements rejoindront une feuille, et il sera attribué à l'évènement la pureté de cette feuille. La pureté totale de l'évènement correspondra alors à la moyenne des puretés calculées pour cet évènement par chaque arbre, et pondérée de leur performance. La performance d'un arbre est définie par $Perf = N * (1-N)$, N étant égal au nombre relatif des évènements qui ont correctement été sélectionnés pendant l'entraînement. Si un arbre classe un grand nombre d'évènements en bruit de fond, alors qu'il s'agit d'évènements de signal (et réciproquement), cet arbre aura une mauvaise performance.

Le RF a été entraîné séparément pour les Runs IIa et IIb, pour les 11 masses de Higgs comprises entre 100 GeV et 150 GeV (par pas de 5 GeV), et pour les deux lots d'évènements 1VT et 2L. Le RF a également été entraîné séparément pour les

quatre analyses ee, ee^{icr}, $\mu\mu$, et $\mu\mu^{trace}$, ainsi que pour différentes combinaisons des canaux et des Runs. Le RF a été entraîné avec un quart des échantillons simulés de signal ZH et des bruits de fond SM ; un autre quart des évènements simulés a servi à le tester et à déterminer son efficacité, cf. fig. 7.18 ; et la moitié restante des évènements simulés ainsi que la totalité des données et de bruit de fond QCD ont permis de produire les distributions qui serviront à établir les limites.
Le RF fournit une variable finale discriminante (RF_{out}) comprise entre 0 et 1 pour les données et la simulation. Ces dernières ressembleront d'autant plus à du signal (bruit de fond) que leurs valeurs dans la distribution de la variable finale discriminante sera proche de 1 (0). La RF_{out} permet ainsi de rechercher le signal du boson de Higgs. La RF_{out} est utilisée par COLLIE pour une analyse statistique, afin de déterminer la section efficace de production du boson de Higgs, voir chapitre 8.

7.10.2 Variables utilisées

Les 10 variables qui ont servi à entraîner le RF ont été choisies au hasard parmis les 20 variables suivantes :

les variables des jets : la distribution de la masse invariante des deux jets (M^{j1j2}), la somme de l'impulsion transverse des deux jets (p_T^{j1j2}), et l'impulsion transverse des deux jets (p_T^{j1} et p_T^{j2}). Les angles $\Delta\phi$ et $\Delta\eta$ entre les deux jets. Enfin, les variables M^{j1j2}, p_T^{j1} et p_T^{j2}, après la correction de l'ajustement cinématique. En général, les variables qui sont corrélées ne servent pas à l'entraînement du RF, mais comme ces variables des jets ont des distributions suffisamment différentes avant et après l'ajustement cinématique, (cf. fig. 7.16), elles ont été utilisées.

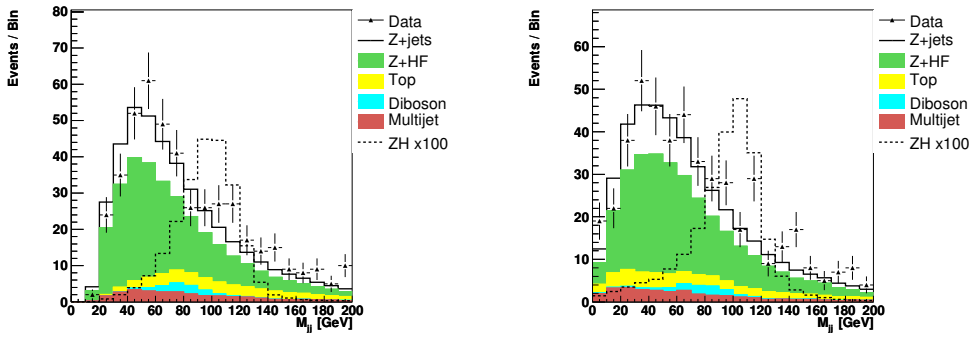
les variables associées au boson Z : l'impulsion transverse du boson Z, les angles $\Delta\phi$ et $\Delta\eta$ entre les deux leptons, ainsi que la colinéarité entre les deux leptons qui est le cosinus de l'angle entre les deux leptons.

les variables associées au système Z+jets : la somme de l'impulsion transverse des deux leptons et des deux jets ($\mathrm{p}_T(l1l2, j1j2)$) après la correction de l'ajustement cinématique, ainsi que l'angle $\Delta\phi$ entre les leptons et les deux jets $\Delta\phi$(Z, $j1j2$).

les variables topologiques : $\cos(\theta^*)$, θ^* est l'angle entre la direction du faisceau de proton et le Z dans le repère où l'impulsion transverse est nulle, il permet de différencier les Z produits en association avec deux jets de b issus du Higgs de spin nul, des autres Z associés à des jets de b issus de gluon de spin entier [116] [117]. La masse invariante $\mathrm{M}(l1l2, j1j2)$ et la somme scalaire de l'énergie $\mathrm{H}_T(l1l2, j1j2)$ associée aux deux leptons $l1$ et $l2$ et aux deux jets. Enfin, la masse invariante et la somme de l'impulsion transverse de tous les jets présents dans les évènements sont aussi utilisées.

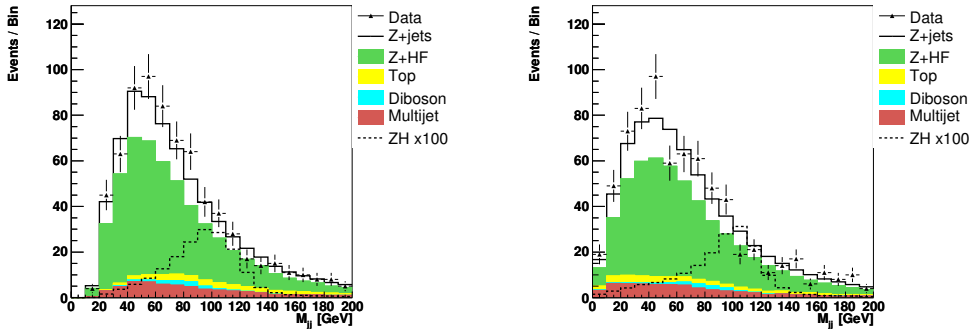
Il est très important d'avoir un bon accord entre les données et la simulation des variables servant à entraîner le RF. On comprend donc que l'étape de la pré-sélection est cruciale pour cette partie de l'analyse, et par conséquent pour l'établissement de

la limite. D'autres distributions non encore montrées dans les chapitres précédents comme la colinéarité entre les deux électrons, la masse invariante de tous les jets et celle du système $(l1l2, j1j2)$ ou encore la somme de l'énergie scalaire du système $(l1l2, j1j2)$, pour le cas où les analyses ee et ee^{icr} sont combinées, sont montrées en exemple sur la figure 7.19.



(a) $M_{2L}^{jet^{1,2}}$ (pre-fit)

(b) $M_{2L}^{jet^{1,2}}$ (post-fit)



(c) $M_{1VT}^{jet^{1,2}}$ (pre-fit)

(d) $M_{1VT}^{jet^{1,2}}$ (post-fit)

FIG. 7.16 – Distribution de la masse invariante des deux jets, pour le lot 2L avant 7.16(a) et après 7.16(b) ajustement cinématique, et pour le lot 1VT avant 7.16(c) et après 7.16(d) ajustement cinématique. Les quatre analyses ee, ee^{icr}, $\mu\mu$ et $\mu\mu^{trk}$ ont été combinées. Le signal ZH simulé pour une masse de 115 GeV, et multiplié par un facteur 100, est représenté en pointillé.

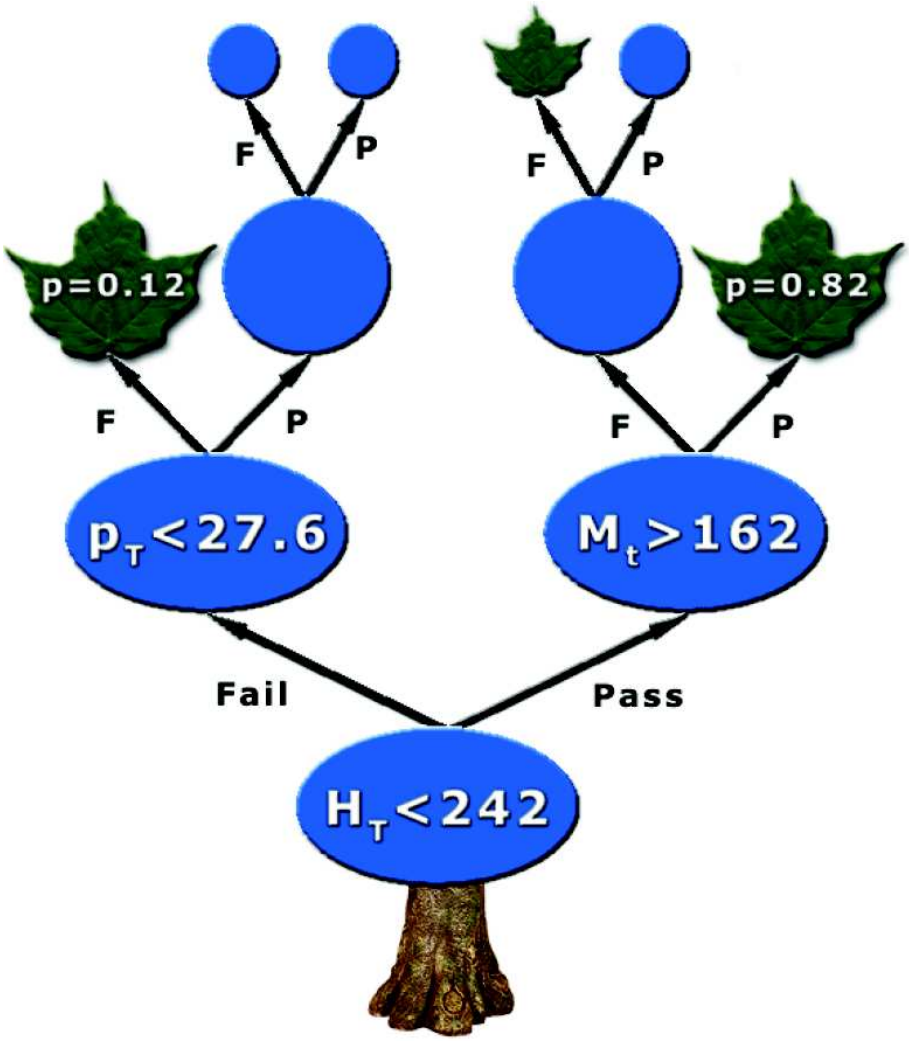


FIG. 7.17 – Schéma d'un des arbres du RF. Les nœuds sont représentés par des cercles (en bleus) et les nœuds terminaux sont représentés par des feuilles (en vertes).

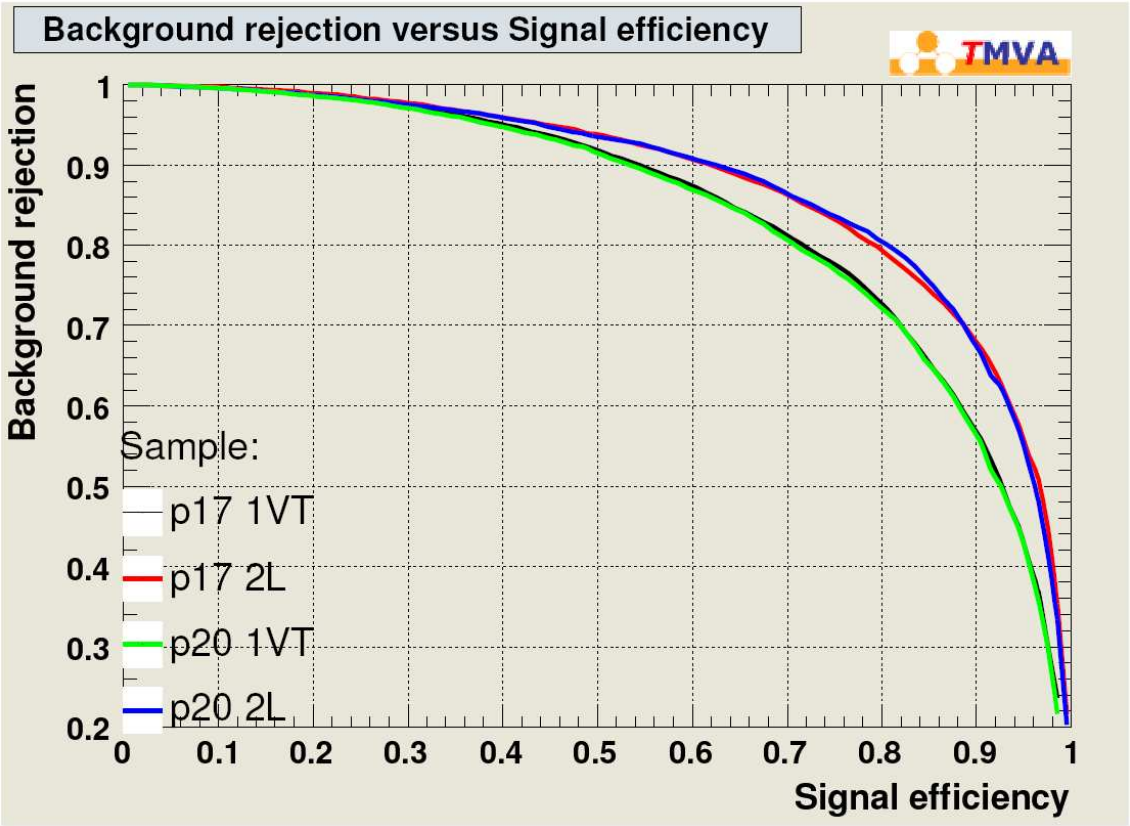
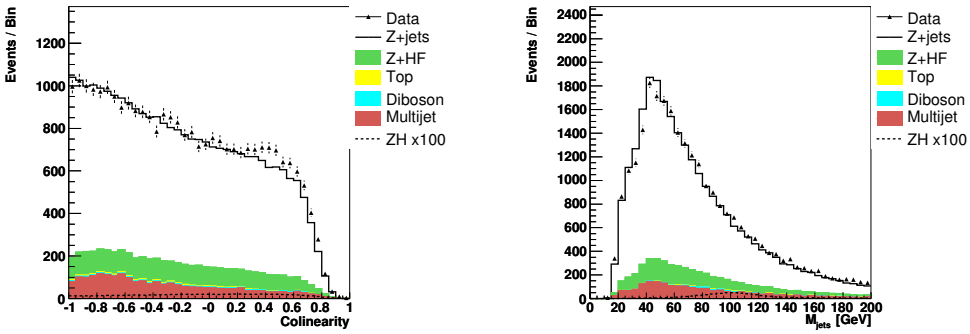
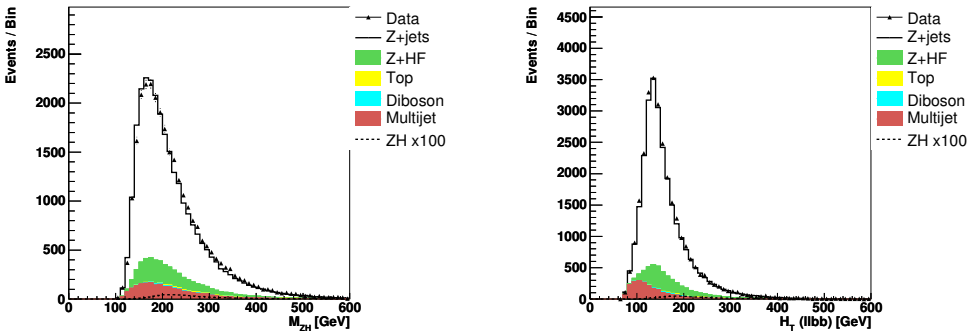


FIG. 7.18 – Efficacité du signal en fonction du bruit de fond pour le RF. Les deux courbes au-dessus rouge et bleue sont respectivement les efficacités pour les échantillons 2L pour les Runs IIa (p17 2L) et IIb (p20 2L). Les deux courbes d'efficacité aux-dessous noire et verte sont respectivement pour les échantillons 1VT des Runs IIa (p17 1VT) et IIb (p20 1VT).



(a) $Colinearity$ $(ee + ee^{icr})$

(b) M^{jets} $(ee + ee^{icr})$



(c) $M^{l1l2,j1j2}$ $(ee + ee^{icr})$

(d) $H_T^{l1l2,j1j2}$ $(ee + ee^{icr})$

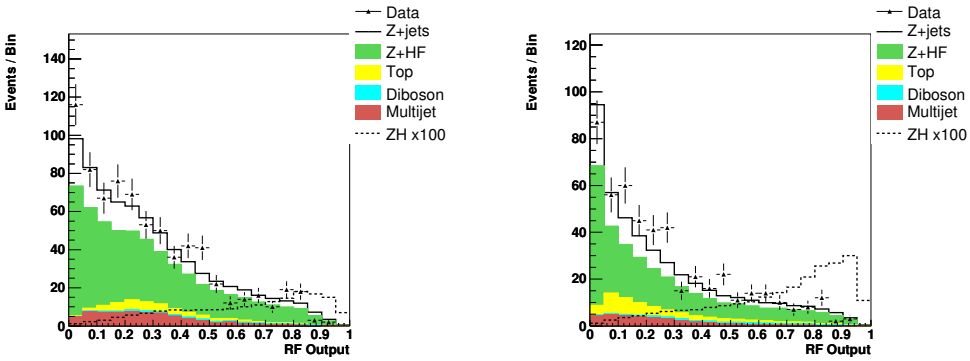
FIG. 7.19 – Distribution de la colinéarité entre les deux électrons 7.19(a), de la masse invariante de tous les jets 7.19(b) et du système $(l1l2, j1j2)$ 7.19(c), et enfin la distribution de la somme de l'énergie scalaire associée au système $(l1l2, j1j2)$ 7.19(d). Les analyses ee et ee^{icr} ont été combinées, de plus, on demande au moins deux jets dans les évènements et la masse du Z doit être comprise entre 60 GeV et 150 GeV.

7.10.3 Variable discriminante finale

La figure 7.20 montre la variable de sortie du RF pour les lots 1VT et 2L, pour le cas où les quatre analyses sont combinées et pour l'analyse ee^{icr} seule. La figure 7.21 montre la distribution de la différence entre les données et les bruits de fond, obtenue avec le RF pour un Higgs de 115 GeV, les quatre analyses ainsi que les RunIIa et IIb ont été combinés. On voit que nos données sont compatible avec le bruit de fond. Comme il n'y a pas d'excès de données dans la région du signal, la variable discriminante du RF est utilisée pour l'établissement de la limite.

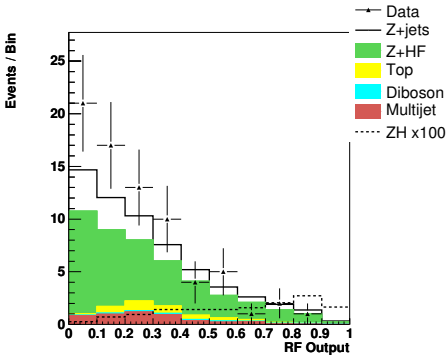
7.11 Conclusion

Ce chapitre décrit la recherche du boson de Higgs du modèle standard produit en association avec un boson Z, où le Higgs se désintègre en paire de jets $b\bar{b}$ et le Z en électron-positron. Les grandes étapes de l'analyse qui sont la pré-sélection des données et des bruits de fond, l'étiquetage des jets de b, les différentes corrections appliquées aux évènements simulés, et ainsi que la sélection finale des évènements ont été détaillées. Enfin, les distributions de contrôle entre les données et la simulation, présentées à chaque étape de l'analyse, ont montrées un bon accord.

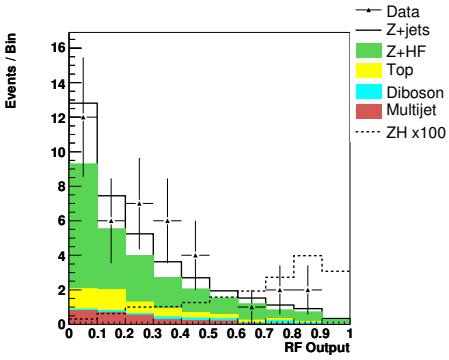


(a) RF_{out}^{1VT} *all*

(b) RF_{out}^{2L} *all*



(c) RF_{out}^{1VT} ee^{icr}



(d) RF_{out}^{2L} ee^{icr}

FIG. 7.20 – Variables de sorties du RF pour un Higgs de masse 115 GeV, pour les quatre analyses ee, ee^{icr}, $\mu\mu$ et $\mu\mu^{trk}$ combinées, pour les lots 1VT 7.20(a) et 2L 7.20(b), et pour l'analyse ee^{icr} seule, pour les lots 1VT 7.20(c) et 2L 7.20(d).

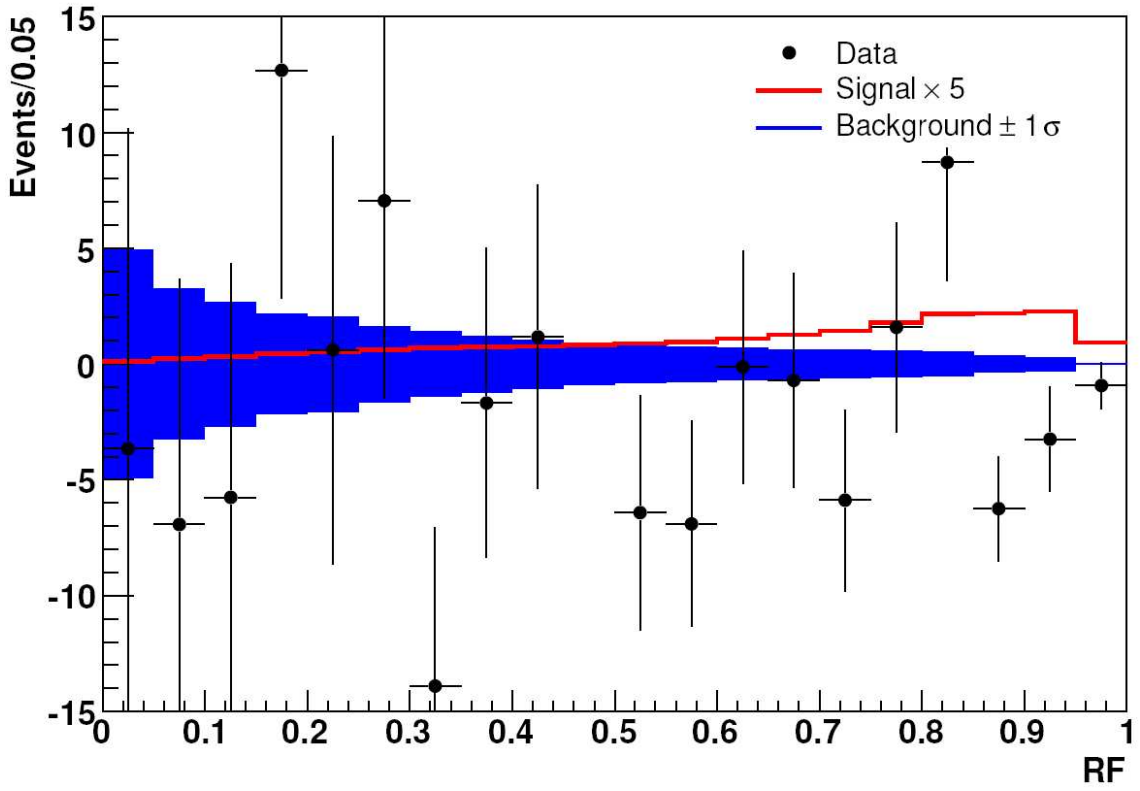


FIG. 7.21 – Distribution de la différence entre les données et les bruits de fond en fonction de RF_{out} pour un Higgs de 115 GeV. Les points sont les données obtenues après la soustraction des bruits de fond. La bande bleue représente l'erreur systématique après traitement, cf. 8.2.2, et la ligne rouge représente la distribution du signal ZH multipliée par un facteur 5.

Chapitre 8

Limites sur la section efficace de production du boson de Higgs

Contents

8.1 La statistique de test 141
8.1.1 La méthode CL_s 141
8.1.2 Le niveau de confiance CL_s 144
8.2 Les erreurs systématiques 144
8.2.1 Erreurs systématiques dans l'analyse ZH 144
8.2.2 Traitement des erreurs systématiques 149
8.3 Calcul de la limite 149
8.3.1 Implémentation de la limite 149
8.3.2 Limite obtenue 149
8.4 Conclusion 150
Résumé 155

La comparaison entre les données et les différents bruits de fond, n'a montrée aucun excès d'évènements compatible avec la présence d'un signal ZH dans les données. Ainsi, dans ce chapitre nous calculons la limite supérieure sur la section efficace de production du boson de Higgs en fonction de sa masse dans le canal $ZH \rightarrow e^+e^-b\bar{b}$. Nous étudierons la statistique de test utilisée, nous verrons également le traitement des erreurs systématiques, et enfin les limites supérieures établies sur la section efficace de production du boson de Higgs seront présentées.

8.1 La statistique de test

8.1.1 La méthode CL_s

La limite a été calculée à partir de la méthode des niveaux de confiance CL_s (Confidence Levels) [118], implémentée dans un package informatique appelé COLLIE (COnfidence Level LImit Evaluator) [119] [120] [121]. Cette méthode a également été utilisée par les collaborations LEP [122] [123].

Soit un échantillon final où s représente le nombre d'évènements de signal, b le nombre de bruits de fond, et d le nombre d'événements observés, alors la statistique de test Q est le rapport des probabilités (likelihood ratio test statistic) suivantes :

$$Q = \frac{L(s+b)}{L(b)}, \tag{8.1}$$

dans cette équation $L(s+b)$ est la probabilité poissonienne que l'échantillon testé contienne du signal et du bruit de fond, et $L(b)$ celle qu'il ne contienne que du bruit de fond. En considérant $L(s+b)$ et $L(b)$ comme des expériences de comptage, l'expression de Q prend la forme adéquate :

$$Q = \frac{e^{-(s+b)}(s+b)^d/d!}{e^{-b}(b)^d/d!}. \tag{8.2}$$

Lorsque l'on recherche un signal dans une variable finale pour différents canaux i, Q est calculée dans chaque bin j des histogrammes suivant la combinaison suivante :

$$Q = \prod_{i=1}^{canaux} \prod_{j=1}^{bins} \frac{e^{-(s_{ij}+b_{ij})}(s_{ij}+b_{ij})^{d_{ij}}/d_{ij}!}{e^{-b_{ij}}(b_{ij})^{d_{ij}}/d_{ij}!} \tag{8.3}$$

$$Q = \prod_{i=1}^{canaux} \prod_{j=1}^{bins} e^{-(s_{ij})} \left(\frac{s_{ij}+b_{ij}}{b_{ij}}\right)^{d_{ij}}. \tag{8.4}$$

Pour rendre l'expression de Q linéaire, et donc simplifier l'ajout des canaux et des bins, on utilise le log-likelihood ratio (LLR), noté χ et défini par :

$$\chi = -2ln(Q) = 2 \sum_{i}^{canaux} \sum_{j}^{bins} (s_{ij} - d_{ij}\ ln(1 + s_{ij}/b_{ij})). \tag{8.5}$$

Soient $dP_b/d\chi$ et $dP_{s+b}/d\chi$, les densités de probabilité des LLR pour les hypothèses bruit de fond seul (LLR_b) et signal plus bruit de fond (LLR_{s+b}) respectivement. Elles sont calculées à partir de pseudo-expériences où les d_{ij} sont tirés au hasard selon une statistique poissonienne de valeur moyenne b et $s+b$ respectivement.

s et b dépendent de paramètres dit de nuisances. Les paramètres de nuisances dans cette analyse sont les paramètres qui interviennent dans le calcul de la limite, comme par exemple la luminosité ou les efficacités. Les paramètres de nuisances comportent des erreurs appelées erreurs systématiques. On peut tenir compte de ces erreurs en convoluant leurs distributions avec les lois de poisson. Ces dernières se trouvent élargies et on perd de la sensibilité. La dégradation sur le résultat final sera d'autant plus importante que les erreurs sur b et $s+b$ seront importantes. Il

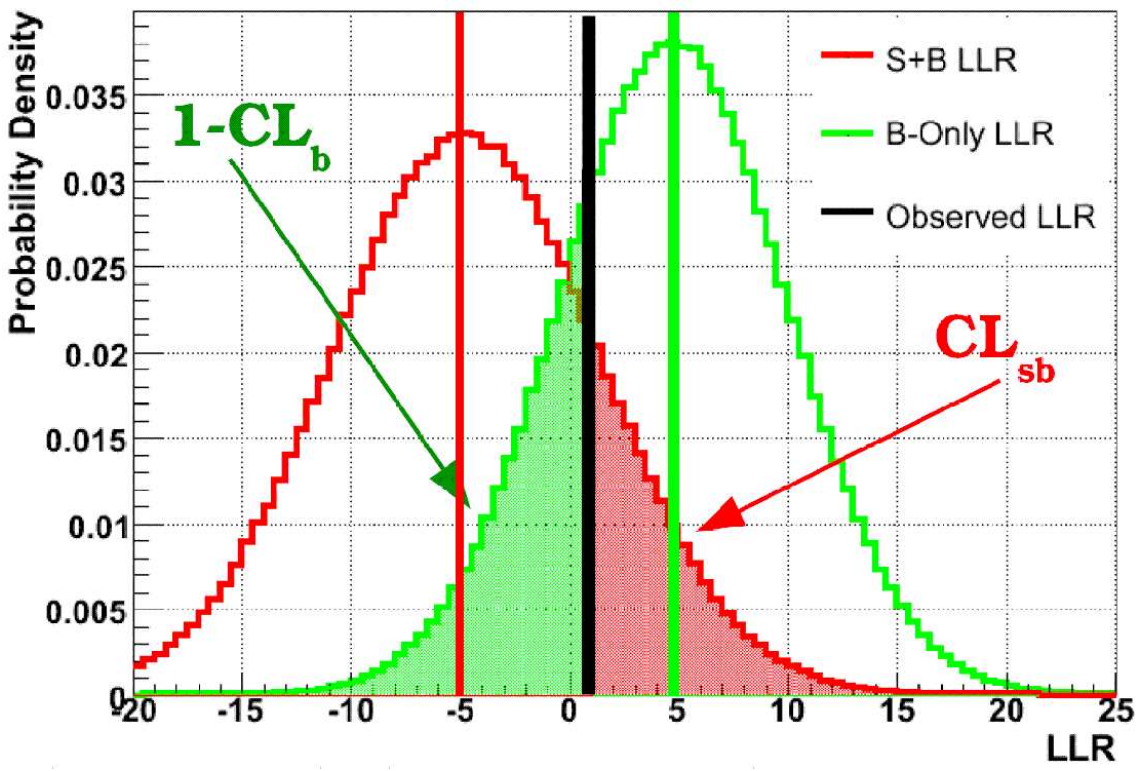


FIG. 8.1 – Exemple de distributions des LLR pour les hypothèses signal plus bruit de fond (distribution de gauche en rouge) et bruit de fond seul (distribution de droite en verte). Le trait rouge vertical (à gauche) représente la valeur centrale du LLR pour l'hypothèse signal plus bruit de fond, le trait noir vertical (au milieu) représente la valeur de l'observation, et le trait vert vertical (à droite) représente la valeur centrale de l'hypothèse bruit de fond seul.

est donc important de contrôler ces erreurs et notamment les erreurs systématiques. Les erreurs systématiques prises en compte dans cette analyse sont listées en 8.2.1, et leurs effets ont été minimisés comme expliqué en 8.2.2.

La sensibilité est caractérisée par la distance entre les médianes passant par le sommet des deux distributions LLR_b et LLR_{s+b}, divisée par l'écart type de la distribution de LLR_b. Le test sera d'autant plus performant que ce rapport sera grand, et on juge qu'une "bonne" sensibilité doit avoir un rapport ≥ 2, voir fig. 8.1. Pour quantifier les fluctuations du LLR observés (LLR_{obs}), des bandes représentant des écarts de 1 (2) déviation(s) standard(s) (notée σ) sont déterminées autour de la valeur calculée dans l'hypothèse bruit de fond seulement. La présence d'un signal est caractérisée par une déviation négative de 2 ou 3 σ de LLR_{obs} par rapport à la médiane de LLR_b.

8.1.2 Le niveau de confiance CL_s

Le niveau de confiance est calculé à partir des LLR. Soit CL_{s+b} l'hypothèse signal plus bruit de fond et CL_b l'hypothèse bruit de fond seul, alors :

$$CL_{s+b} = P_{s+b}(\chi \geq \chi_d) = \int_{\chi_d}^{\infty} \frac{dP_{s+b}}{d\chi} d\chi \tag{8.6}$$

$$CL_b = P_b(\chi \geq \chi_d) = \int_{\chi_d}^{\infty} \frac{dP_b}{d\chi} d\chi. \tag{8.7}$$

Enfin, le niveau de confiance CL_s, est le rapport des deux précédents termes, avec par convention le CL_{s+b} au numérateur [124] :

$$CL_s = \frac{CL_{s+b}}{CL_b}. \tag{8.8}$$

8.2 Les erreurs systématiques

8.2.1 Erreurs systématiques dans l'analyse ZH

Nous distinguons deux types d'erreurs systématiques : les erreurs dites "plates" ou "lisses" (flat systematics) sont les erreurs des quantités comme les facteurs de normalisation, ou les sections efficaces, etc ..., qui ne changent les distributions de RF qu'avec un facteur constant ; et les erreurs dites de "forme" (shape systematics), sont les erreurs des quantités qui sont susceptibles de changer la forme des distributions de RF. COLLIE traite ces deux types d'erreur d'une manière différente.
On distingue également les erreurs "corrélées" et "indépendantes" : les erreurs corrélées sont les erreurs des paramètres qui sont communs pour les différents canaux comme la luminosité, les sections efficaces, ou les facteurs de normalisation k_Z^j ; et les erreurs indépendantes sont par exemple les facteurs d'échelle des efficacités k_ϵ^i, ou les coefficients du bruit de fond multijet α_{ij}, et elles sont appliquées indépendamment aux canaux.

Erreurs systématiques lisses

– Les estimations des erreurs systématiques communes sont montrées dans le tableau 8.1. Dans la dernière colonne est indiquée la fraction par rapport au bruit de fond total. Par exemple, une erreur de 20% qui affecte de 10% le bruit de fond correspond à 2%. σ_ϵ^{COR} est essentiellement l'erreur sur le facteur $k_L * k_Z^0$, voir Eq. 7.2. L'erreur r(A/B) (r(EC) pour la topologie CC-EC) tient compte d'une erreur dont l'origine est la luminosité accrue pour les Runs IIb et qui fait que les valeurs de k_Z^2 déterminées séparément pour les Runs IIa et IIb, sont différentes au-delà de l'erreur statistique.
– Les estimations des erreurs systématiques appliquées séparément dans les différents canaux sont montrées dans le tableau 8.2. L'intervention des erreurs systématiques dans les différents canaux est montrée dans le tableau 8.3. On y montre

Systematic	Uncertainty (%)	bkg (%)
σ_ϵ^{COR}	6	0.7
k_Z^2 (Run IIa)	1.1	0.9
k_Z^2 (Run IIb)	0.9	0.7
k_Z^2 (CC-EC)	1.8	
r(A/B)	7	0.8
r(EC)	13	
$Z + HF$ xsec	20	10.9
$t\bar{t}$ xsec	10	0.9
Diboson xsec	7%	0.2%
ZH xsec	6	

TAB. 8.1 – Erreurs systématiques en pourcentage, sur la normalisation et les sections efficaces, communes pour tous les canaux.

Run IIa

i	ee (bkg) (%)	ee^{icr} (bkg) (%)	$\mu\mu$ (bkg) (%)	$\mu\mu^{trace}$ (bkg) (%)
$\sigma_\epsilon^{i,IND}$	2 (1.8)	4 (3.6)	2 (2.0)	3 (2.4)
Multijet	20 (2.5)	60 (6.1)	50 (1.1)	50 (10)

Runs IIb

i	ee (bkg) (%)	ee^{icr} (bkg) (%)	$\mu\mu$ (bkg) (%)	$\mu\mu^{trace}$ (bkg) (%)
$\sigma_\epsilon^{i,IND}$	2 (1.6)	4 (1.9)	2 (1.9)	3 (1.9)
Multijet	20 (4.2)	20 (11)	30 (0.9)	20 (7.0)

TAB. 8.2 – Erreurs systématiques en pourcentage, sur la normalisation pour chaque canaux, et pour les Runs IIa (en haut) et IIb (en bas).

également si l'erreur est appliquée indépendamment ou si elle est corrélée entre les Runs IIa et IIb.

Erreurs systématiques de forme

Les erreurs systématiques sur la forme de la distribution du RF sont estimées en faisant varier les paramètres expérimentaux de $\pm 1\sigma$, et en mesurant l'effet de cette variation sur la variable finale. L'amplitude du changement de cette dernière donne l'erreur systématique. Les valeurs en moyenne trouvées pour les Runs IIa et IIb sont regroupées dans les tableaux 8.4 pour l'échantillon 1T, et 8.5 pour l'échantillon 2L. On considère :

- L'identification des jets (Jet ID), l'impact des incertitudes de l'échelle d'énergie des jets (JES), et la résolution en énergie des jets (RES) ;
- Z_{P_T} : la repondération de l'impulsion transverse du boson Z pour les processus $Z + jets$;

	Multijet	$Z+LP$	$Z+HF$	Diboson	$t\bar{t}$	ZH	Run IIa/IIb
Multijet	X						Independent
k_Z^2		X	X				Independent
$\sigma_\epsilon^{i,IND}$		X	X	X	X	X	Independent
σ_ϵ^{COR}				X	X	X	Correlated
r(A/B)				X	X	X	Independent
r(EC)				X	X	X	Independent
$Z+HF$ xsec			X				Correlated
Diboson xsec				X			Correlated
$t\bar{t}$ xsec					X		Correlated
ZH xsec						X	Correlated

TAB. 8.3 – Applicabilité des erreurs systématiques sur la normalisation et les sections efficaces pour chaque échantillon, et en fonction de la corrélation entre les Runs IIa et IIb.

- PDF : l'effet du choix des fonctions de densité de partons (PDF : Parton Density Function) des jets sur l'acceptance du signal a été mesuré [125] [126] ;

- TRF HF : la valeur attribuée à l'efficacité d'étiquetage des jets de b ;
- TRF LF : la valeur attribuée à l'efficacité d'étiquetage des jets légers ;
- Taggability : facteur de correction de la taggabilité des jets ;
- Vjet RW : erreur systématique de la renormalisation des processus $Z+jets$ simulés ;
- Trig : la valeur attribuée au facteur de correction du trigger ;
- MLM : erreur systématique associée au matching ;
- Alpgen Scale : la repondération de la simulation Alpgen pour corriger la cinématique des jets est appliquée au processus $Z+jets$;
- Alpgen UE : erreur systématique associée aux paramètres utilisés dans la simulation pour générer les interactions des partons spectateurs.

Run IIa

	ZH (115) (%)	$Z+LP$ (%)	$Z+bb$ (%)	$Z+c\bar{c}$ (%)	Diboson	$t\bar{t}$	Multijet
JES	2.4	1.0	3.5	1.7	3.5	4.0	
RES	3.3	3.9	8.9	7.8	5.4	5.5	
Jet ID	1.0	0.6	0.5	0.4	0.6	0.7	
ZpT		2.2	2.3	2.1			
PDF	0.9	0.6	1.7	0.9	0.2	0.7	
TRF HF	1.0		2.3	4.0		0.4	
TRF LF		16			3.8		1.6
taggability	2.4	2.2	2.5	2.5	2.4	2.5	
Vjet RW		0.9	0.9	0.9			
$\mu\mu$ Trig	1.9	1.9	2.0	2.1	2.2	1.9	
$\mu\mu^{trace}$ Trig	4.8	4.6	4.8	4.9	3.8	5.4	
MLM		0.1					
Alpgen Scale		0.04	0.1	0.04			
Alpgen UE		0.1	0.1	0.07			

Runs IIb

	ZH (115) (%)	$Z+LP$ (%)	$Z+bb$ (%)	$Z+c\bar{c}$ (%)	Diboson	$t\bar{t}$	Multijet
JES	1.9	0.9	2.4	2.8	3.1	3.8	
RES	1.3	2.2	4.9	3.3	1.5	2.4	
Jet ID	2.9	0.2	1.2	0.2	5.5	3.5	
ZpT		2.4	2.6	2.3			
PDF	1.0	0.7	1.6	1.0	0.2	0.3	
TRF HF	0.03		2.3	3.6		0.4	
TRF LF		19			4.0		4.2
taggability	0.6	0.4	0.4	0.5	0.7	0.7	
Vjet RW		0.9	0.9	0.9			
$\mu\mu$ Trigger	1.8	1.7	1.8	1.8	1.8	1.7	
$\mu\mu^{trace}$ Trigger	3.6	3.6	3.3	3.6	3.3	3.4	
MLM		0.03					
Alpgen Scale		0.01	0.07	0.04			
Alpgen UE		0.08	0.1	0.07			

TAB. 8.4 – Erreurs systématiques en pourcentage, sur les différents paramètres pour chaque échantillon du lot 1VT, pour les Runs IIa (en haut) et IIb (en bas).

Run IIa

	ZH (115) (%)	$Z+LP$ (%)	$Z+bb$ (%)	$Z+c\bar{c}$ (%)	Diboson	$t\bar{t}$	Multijet
JES	3.1	1.0	3.0	3.2	3.7	3.4	
RES	3.5	4.9	12	10	3.8	3.9	
Jet ID	1.1	0.7	0.7	0.8	0.6	0.8	
ZpT		2.2	2.1	1.9			
PDF	0.7	0.6	1.5	0.6	0.3	0.7	
TRF HF	5.6		4.6	5.3		6.5	
TRF LF		4.9			4.5		0.4
taggability	4.9	3.9	4.6	4.1	4.5	5.5	
Vjet RW		0.8	0.9	0.9			
$\mu\mu$ Trig	2.0	2.1	2.1	2.1	2.2	2.0	
$\mu\mu^{trace}$ Trig	3.8	5.0	4.9	4.7	4.4	6.4	
MLM		0.2					
Alpgen Scale		0.1	0.01	0.06			
Alpgen UE		0.1	0.1	0.3			

Runs IIb

	ZH (115) (%)	$Z+LP$ (%)	$Z+bb$ (%)	$Z+c\bar{c}$ (%)	Diboson	$t\bar{t}$	Multijet
JES	2.7	0.4	2.5	3.0	5.1	3.1	
RES	2.0	1.3	7.9	3.7	4.2	0.9	
Jet ID	3.6	0.2	0.5	0.7	2.8	4.4	
ZpT		2.5	2.3	2.1			
PDF	0.7	0.7	1.5	1.3	0.1	0.4	
TRF HF	5.0		4.1	4.5		6.7	
TRF LF		4.8			4.4		1.0
taggability	0.9	0.8	0.9	0.8	1.0	1.2	
Vjet RW		0.9	1.0	0.9			
$\mu\mu$ Trig	1.8	2.1	1.9	1.8	2.2	1.8	
$\mu\mu^{trace}$ Trig	3.5	4.4	3.8	3.0	3.7	3.7	
MLM		0.1					
Alpgen Scale		0.06	0.02	0.02			
Alpgen UE		0.01	0.1	0.2			

TAB. 8.5 – Erreurs systématiques en pourcentage, sur les différents paramètres pour chaque échantillon du lot 2L, pour les Runs IIa (en haut) et IIb (en bas).

8.2.2 Traitement des erreurs systématiques

Le test statistique Q (cf. 8.1) dépend des erreurs systématiques. Ces dernières élargissent les fonctions de densité de probabilité et réduisent donc les densités de probabilité. Ceci a pour effet de réduire le pouvoir discriminant entre signal et bruit. Pour minimiser l'impact des erreurs systématiques, on utilise la méthode dite de Profile Likelihood [127] [128] qui consiste à déterminer le meilleur ajustement entre les bruits de fond et les données. L'ajustement va maximiser un test de vraisemblance (ou likelihood) entre les données et les bruits de fond, en étant paramétré en fonction des paramètres de nuisances et de leur incertitudes, ce qui va réduire leurs erreurs et ainsi l'élargissement des distributions des LLR.

8.3 Calcul de la limite

8.3.1 Implémentation de la limite

Les limites sont calculées à partir de deux estimateurs de CL_s, CL_s^{exp} et CL_s^{obs}, des équations (8.6) et (8.7). CL_s^{exp} (limite attendue) est le cas où la valeur de la coupure χ_d correspond à la médiane du LLR dans l'hypothèse bruit de fond seul, et CL_s^{obs} (limite observée) est le cas où la valeur de la coupure χ_d correspond à l'observation. Comme dit précédemment, la détermination de la limite se fait avec l'implémentation de la méthode CL_s qui se trouve dans COLLIE. Ce dernier calcule la limite supérieure sur la section efficace de production du boson de Higgs à 95% CL, c'est à dire la limite qui correspond à $CL_s \ \leq 5\%$.

Le signal s est relié à la section efficace de production théorique σ_{SM}, à la luminosité L et à son acceptance ϵ_s par : $s = R \ . \ \sigma_{SM} \ . \ L \ . \ \epsilon_s$, où R est un facteur. Pour déterminer la limite à 95% CL, R est augmenté jusqu'à ce que CL_s diminue et atteigne 5% . R est alors la limite supérieure à 95% CL (noté R_{95}) de la section efficace en unité de la section efficace du SM.

8.3.2 Limite obtenue

Etant donné qu'aucun excès de données compatible avec la présence d'un signal du Higgs n'a été observé dans la variable de sortie du RF pour les deux lots 1VT et 2L, ces derniers ont servi de variable d'entrée dans COLLIE, pour établir la limite supérieure sur la section efficace de production du Higgs, dans l'unité de la section efficace de production du SM. Des limites ont été calculées pour la combinaison des Runs IIa et IIb, pour les 11 masses du Higgs comprises entre 100 GeV et 150 GeV par pas de 5 GeV, pour les analyses ee^{icr} et ee seules et combinées, ainsi que pour les quatre analyses ee^{icr}, ee, $\mu\mu$ et $\mu\mu^{trace}$ combinées.

Le tableau 8.6 montre les limites attendues et observées pour les différentes masses du Higgs, pour différentes combinaisons des quatre analyses. Les figures 8.2,

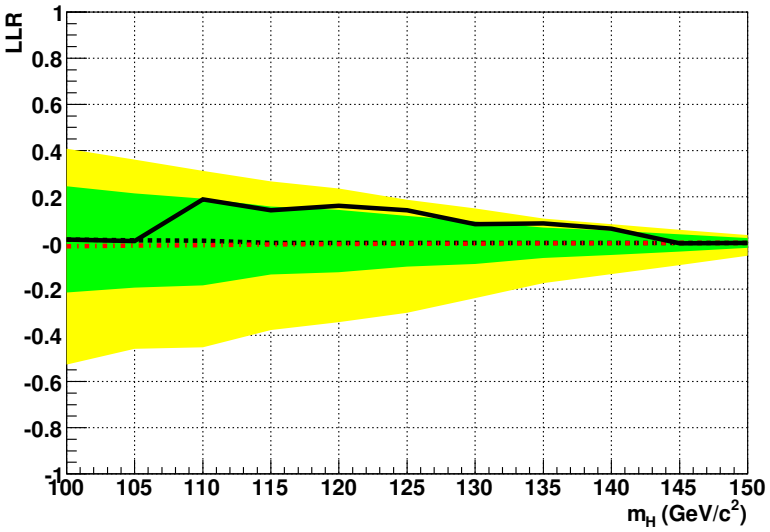
	ee^{icr}		ee		ee^{icr}, ee		$ee^{icr}, ee, \mu\mu, \mu\mu^{trace}$	
M_H GeV	Exp/SM	Obs/SM	Exp/SM	Obs/SM	Exp/SM	Obs/SM	Exp/SM	Obs/SM
100	22.75	21.11	9.18	6.87	8.17	6.05	5.30	2.99
105	26.52	24.82	10.07	8.58	9.09	7.36	5.92	3.52
110	29.55	21 51	10.89	9.08	9.85	7.36	6.44	3.87
115	35.47	27.17	12.54	10.08	11.50	8.22	7.52	4.83
120	39.87	28.02	14.97	17.07	13.30	12.66	8.71	8.65
125	49.06	33.95	17.10	14.39	15.60	11.34	10.24	8.03
130	61.84	49.50	22.62	20.35	20.50	15.58	13.08	11.03
135	86.27	58.47	28.21	24.16	25.87	18.28	16.98	16.44
140	114.47	79.22	40.93	26.23	36.32	20.27	23.70	18.94
145	161.16	151.29	58.52	41.81	53.07	37.73	33.86	35.78
150	270.84	257.79	88.90	69.75	81.65	59.88	53.21	48.86

TAB. 8.6 – Limite observée (Obs) et attendue (Exp) en unité de la section efficace du SM, en fonction de la masse du Higgs, dans les analyses séparées (ee^{icr} et ee), dans les deux analyses combinées (ee^{icr}, ee) et pour les quatre analyses combinées ($ee^{icr}, ee, \mu\mu, \mu\mu^{trace}$). Les Runs IIa et IIb ont été combinés pour chaque canal.

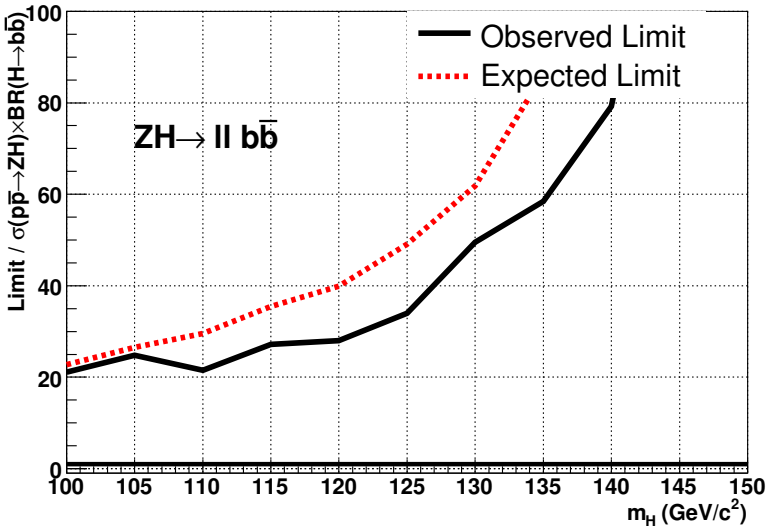
8.3, 8.4, et 8.5 montrent la courbe de R_{95} en fonction de la masse du Higgs ainsi que la distribution des LLR associées aux différentes hypothèses, respectivement pour les analyses ee^{icr} et ee séparées et combinées, et enfin pour les quatre analyses combinées (cf. 8.1 et 8.3.1). Aucune interpolation n'est effectuée entre ces points de masse, et les points pour lesquels la limite est calculée sont simplement reliés par une droite.

8.4 Conclusion

Ce chapitre a décrit la méthode utilisée pour établir la limite sur la section efficace de production du boson de Higgs, ainsi qu'une méthode permettant de réduire l'impact des erreurs systématiques sur la limite. Les erreurs systématiques utilisées ont été énumérées, et enfin la limite sur la section efficace de production du boson de Higgs, pour une gamme de masse comprise entre 100 GeV et 150 GeV a été montrée. Pour un Higgs de 115 GeV, la limite observée (resp. attendue) pour l'analyse ee^{icr} est de $R_{95}^{obs} = 27.17$ (resp. $R_{95}^{exp} = 35.47$), pour l'analyse ee elle est de $R_{95}^{obs} = 10.08$ (resp. $R_{95}^{exp} = 12.54$). Pour la combinaison de ces deux analyses elle est de $R_{95}^{obs} = 8.22$ (resp. $R_{95}^{exp} = 11.50$), et enfin pour la combinaison des quatre analyses elle est de $R_{95}^{obs} = 4.83$ (resp. $R_{95}^{exp} = 7.52$). L'analyse ee^{icr} seule permet une amélioration de près de 9% sur la limite attendue établie dans le canal ee seul.

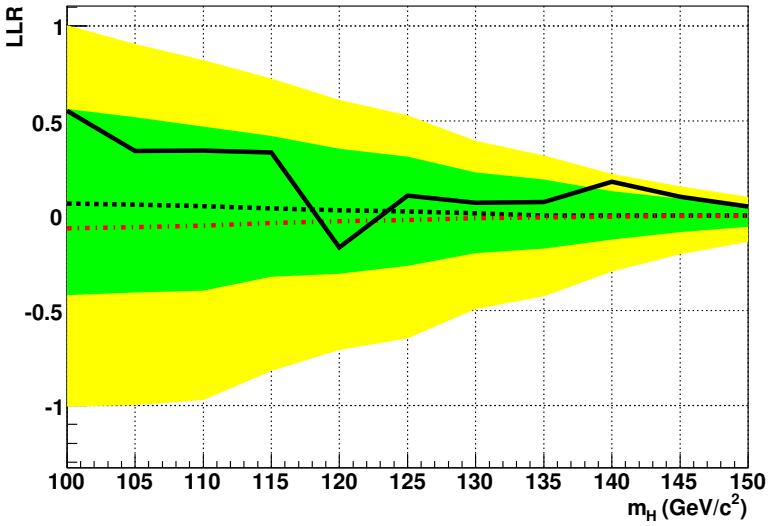


(a) LLR ee^{icr}

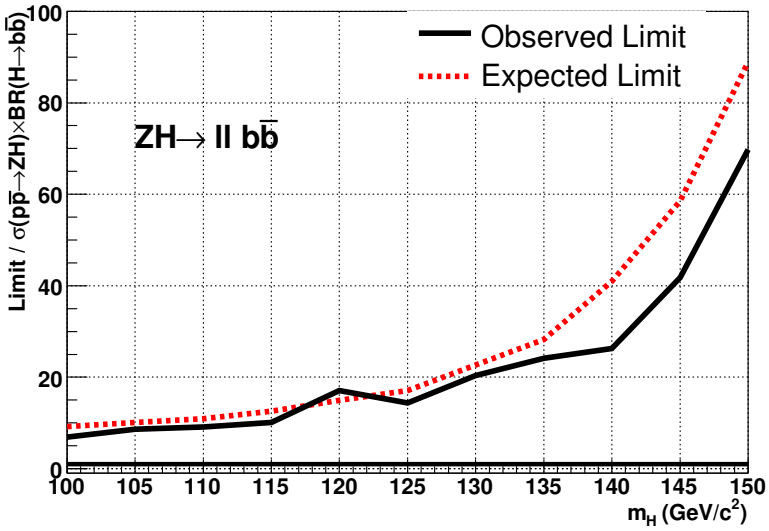


(b) Limit ee^{icr}

FIG. 8.2 – Limite sur la section efficace de production du processus $ZH \rightarrow l^+l^-b\bar{b}$ dans l'analyse ee^{icr}. 8.2(a) distribution du LLR en fonction de la masse du Higgs : la ligne noire pleine représente la valeur du LLR observé, la ligne rouge pointillée la valeur du LLR attendu dans l'hypothèse signal + bruit de fond, la ligne noire pointillée la valeur du LLR attendu dans l'hypothèse bruit de fond seul, la bande verte (foncée) et jaune (claire) représentent respectivement l'intervalle à 1 et 2 σ autour de la valeur médiane du LLR dans l'hypothèse bruit de fond seul. 8.2(b) représente le facteur R_{95} en fonction de la masse du Higgs, la ligne pleine noire (resp. en pointillée rouge) représente la limite observée (resp. attendue).

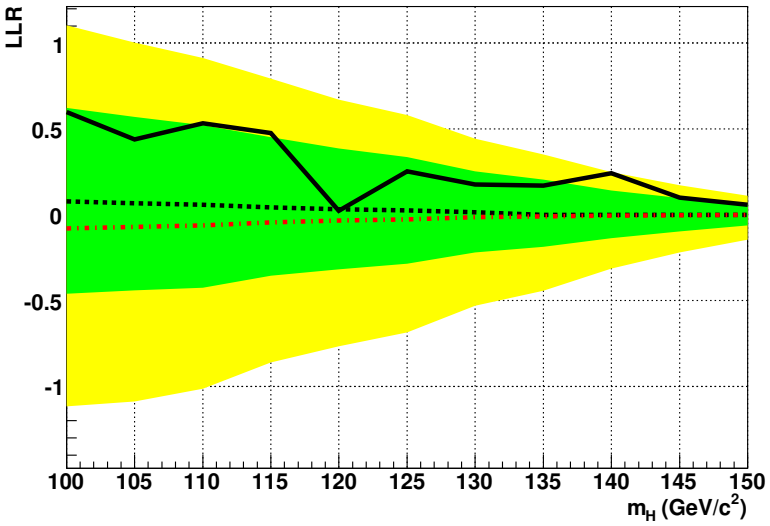


(a) LLR *ee*

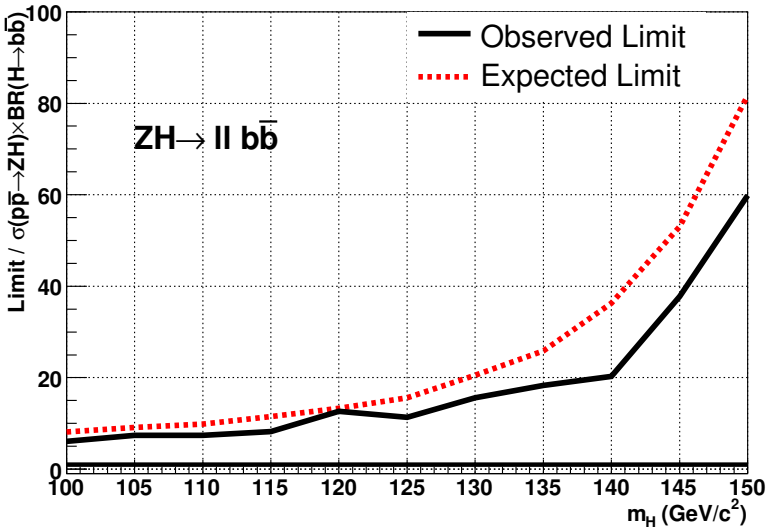


(b) Limite *ee*

FIG. 8.3 – Limite sur la section efficace de production du processus $ZH \rightarrow l^+l^-b\bar{b}$ dans l'analyse *ee*. 8.3(a) distribution du LLR en fonction de la masse du Higgs : la ligne noire pleine représente la valeur du LLR observé, la ligne rouge pointillée la valeur du LLR attendu dans l'hypothèse signal + bruit de fond, la ligne noire pointillée la valeur du LLR attendu dans l'hypothèse bruit de fond seul, la bande verte (foncée) et jaune (claire) représentent respectivement l'intervalle à 1 et 2 σ autour de la valeur médiane du LLR dans l'hypothèse bruit de fond seul. 8.3(b) représente le facteur R_{95} en fonction de la masse du Higgs, la ligne pleine noire (resp. en pointillée rouge) représente la limite observée (resp. attendue).

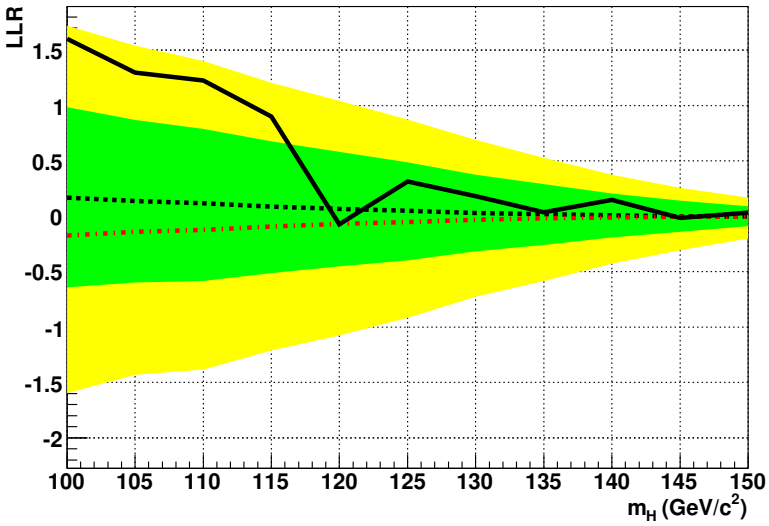


(a) LLR combine ee, ee^{icr}

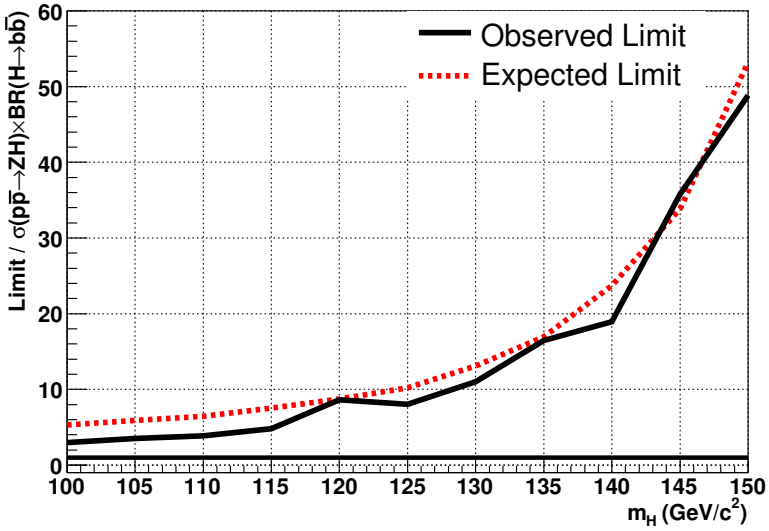


(b) Limite combine ee, ee^{icr}

FIG. 8.4 – Limite sur la section efficace de production du processus $ZH \rightarrow l^+l^-b\bar{b}$ dans la cas où les analyses ee^{icr} et ee ont été combinées. 8.4(a) distribution du LLR en fonction de la masse du Higgs : la ligne noire pleine représente la valeur du LLR observé, la ligne rouge pointillée la valeur du LLR attendu dans l'hypothèse signal + bruit de fond, la ligne noire pointillée la valeur du LLR attendu dans l'hypothèse bruit de fond seul, la bande verte (foncée) et jaune (claire) représente respectivement l'intervalle à 1 et 2 σ autour de la valeur médiane du LLR dans l'hypothèse bruit de fond seul. 8.4(b) représentent le facteur R_{95} en fonction de la masse du Higgs, la ligne pleine noire (resp. en pointillée rouge) représente la limite observée (resp. attendue).



(a) LLR combine ee, ee^{icr}, $\mu\mu$, $\mu\mu^{trace}$



(b) Limit combine ee, ee^{icr}, $\mu\mu$, $\mu\mu^{trace}$

FIG. 8.5 – Limite sur la section efficace de production du processus $ZH \rightarrow l^+l^-b\bar{b}$ dans la cas où les quatre analyses ee^{icr}, ee, $\mu\mu$ et $\mu\mu^{track}$ ont été combinées. 8.5(a) distribution du LLR en fonction de la masse du Higgs : la ligne noire pleine représente la valeur du LLR observé, la ligne rouge pointillée la valeur du LLR attendu dans l'hypothèse signal + bruit de fond, la ligne noire pointillée la valeur du LLR attendu dans l'hypothèse bruit de fond seul, la bande verte (foncée) et jaune (claire) représentent respectivement l'intervalle à 1 et 2 σ autour de la valeur médiane du LLR dans l'hypothèse bruit de fond seul. 8.5(b) représente le facteur R_{95} en fonction de la masse du Higgs, la ligne pleine noire (resp. en pointillée rouge) représente la limite observée (resp. attendue).

Résumé

Ce mémoire décrit les différents travaux auxquels j'ai participés durant ma thèse. Cette dernière s'est déroulée suivant deux grandes étapes. D'abord j'ai effectué des travaux d'intérêt général sur le calorimètre, puis j'ai participé à la recherche du boson de Higgs au Tevatron.

Le chapitre 1 a introduit le modèle standard de la physique des particules et le mécanisme de Higgs. On a vu la cohérence entre le SM et les observations expérimentales, et l'importance de la recherche du bosons de Higgs dans ce modèle. Il est la clé manquante à ce dernier car il permettrait d'expliquer l'origine de la masse des particules.

Le chapitre 2 a présenté le complexe d'accélération du Tevatron et du détecteur DØ à Fermilab. Les différentes étapes allant de la fabrication des faisceaux pour la collision, jusqu'à la reconstruction des évènements à analyser, en passant par la description des différentes parties du détecteur, ont été décrites. Nous avons également vu les moyens mis en œuvre pour la recherche du boson de Higgs.

Le chapitre 3 a introduit les objets physiques qui ont été utilisés tout au long de cette thèse. Les objets électromagnétiques et les jets, déterminant dans la recherche du boson de Higgs de cette thèse, ont fait l'objet d'une attention particulière.

Le chapitre 4 a résumé mon premier travail d'intérêt général. J'ai fait une étude sur l'énergie mesurée dans le calorimètre, et j'ai déterminé des facteurs de correction. L'étude a montrée que les facteurs de correction sont faibles et stables dans le temps. Cette analyse a fait l'objet d'une note interne [59] dont je suis un des principaux auteurs.

Le chapitre 5 a résumé mon deuxième travail d'intérêt général. J'ai participé à la certification de l'identification des électrons dans le calorimètre. Cette étude a permis de déterminer des coupures appropriées pour identifier les électrons dans les différentes régions du calorimètre à haute luminosité. Cette analyse a fait l'objet d'une note interne [66] dont je suis un des principaux auteurs.

Le chapitre 6 décrit l'étude que j'ai faite pour déterminer les efficacités d'identification des électrons ainsi que des facteurs de correction, dans la région inter-cryostat du calorimètre. J'ai montré que les coupures spécifiques utilisées permettaient d'avoir une excellente efficacité de détection et que les facteurs de correction étaient faibles. Cette étude a également fait l'objet d'une note interne [86] dont je suis un des principaux auteurs.

Le chapitre 7 détaille l'analyse sur la recherche du Higgs produit en association avec un boson Z dans les données du détecteur DØ correspondant à une luminosité intégrée de 4.2 fb^{-1}. J'ai travaillé dans l'analyse $ZH \rightarrow e^+e^-b\bar{b}$, au niveau de la pré-sélection des électrons dans le calorimètre, et en particulier dans le cas où un des électrons se trouve dans la région inter-cryostat. Cette analyse a été réalisée en collaboration avec l'analyse $ZH \rightarrow \mu^+\mu^-b\bar{b}$. De bons accords entre les données et la simulation ont été observés lors de la pré-sélection, et cette étape, cruciale, a servi de base à la suite de l'analyse. Aucun excès de données compatible avec le signal recherché n'a été observé dans l'analyse discriminante finale, et des limites sur la section efficace de production du boson de Higgs sur l'intervalle de masse compris entre 100 GeV et 150 GeV, ont par conséquent été établies pour les différents canaux, comme

décrit au **chapitre 8**. Ces analyses ont montrées que pour un Higgs de 115 GeV, la limite observée (resp. attendue) pour l'analyse ee^{icr} seule est de $R_{95}^{obs} = 27.17$ (resp. $R_{95}^{exp} = 35.47$), pour l'analyse ee seule elle est de $R_{95}^{obs} = 10.08$ (resp. $R_{95}^{exp} = 12.54$), pour ces deux analyses combinées elle est de $R_{95}^{obs} = 8.22$ (resp. $R_{95}^{exp} = 11.50$), et enfin pour la combinaison des quatre analyses, elle est de $R_{95}^{obs} = 4.83$ (resp. $R_{95}^{exp} = 7.52$). L'ajout des électrons de la région inter-cryostat a par conséquent amélioré la limite attendue dans l'analyse ee de près de 9%.
Ces résultats ont fait l'objet de plusieurs notes internes [110] [129] [130] [131]. Ils ont été présentés à la conférence Moriond en 2009 [77], et sont actuellement en cours d'approbation par la collaboration DØ pour être publiés au journal Physical Review Letters.

Bibliographie

[1] "Introduction to Elementary Particles" D. Griths, John Wiley & Sons, 1987.

[2] "Gauge Theories of the Strong, Weak, and Electromagnetic Interactions", C. Quigg, West-view Press, 1983.

[3] "Neutrino Masses and Mixing : Evidence and Implications", MC. Gonzalez-Garcia & Y. Nir, hep-ph/0202058 v2 2 Aug 2002.

[4] `http://en.wikipedia.org/wiki/Binary_pulsar`.

[5] "A Model of Leptons", S. Weinberg, Phys. Rev. Lett. 19, 1264-1266 (1967).

[6] S. L. Glashow, Phys. Rev. D2, 1285 (1970).

[7] "The Higgs Hunter's Guide", S. Dawson, J. F. Gunion, H. E. Haber, G. Kane, Frontiers in Physics, Volume 80.

[8] P. Higgs, Phys. Rev. Lett. 12, 132 (1964).

[9] P. Higgs, Phys. Rev. 145, 1156 (1966).

[10] R. Brout, F. Englert, Phys. Rev. Lett. 13, 321 (1964).

[11] Guralnik, Hagen and Kibble Phys.Rev.Letters 13 (1964) 585-587.

[12] N. Cabibbo, L. Maiani, G. Parisi and R. Petronzio, Nucl. Phys. B158, 295 (1979).

[13] G. Altarelli and G. Isidori, Phys. Lett. B 337 (1994) 141.

[14] T. Hambye and K. Riesselmann, Phys. Rev. D 55 (1997) 7255 [arXiv :hepph/9610272].

[15] G. Abbiendi et al., Phys. Lett. B., 565 61, 2003.

[16] R. Barate, Phys. Lett. B565, 61 (2003).

[17] The TEVNPH Working Group, FERMILAB-PUB-08-270-E (2008).

[18] Combined CDF and DØ Upper Limits on Standard Model Higgs-Boson Production with 2.1 - 5.4 fb-1 of Data, The TEVNPH Working Group for the CDF and D0 Collaborations, CDF Note 9998, DØ Note 5983.

[19] `http://www.fnal.gov/`.

[20] `http://public.web.cern.ch/public/`.

[21] G. P. Jackson, Beams-doc-172-v1, 2002.

[22] "Prospectus For an Electron Cooling System for the Recycler", J.A. MacLachlan, FERMILAB-TM-2061.

[23] "The DØ Run II Luminosity Monitor", C. Miao, Fermilab-Conf-98/395-E.

[24] DØ Collaboration, S. Abachi et al., Observation of the Top Quark, Phys. Rev. Lett. 74, 2632 (1995).

[25] CDF Collaboration, F. Abe et al., Observation of the Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab, Phys. Rev. Lett. 74, 2626 (1995).

[26] The D0 Layer 0 silicon group, D0 Layer 0 Conceptual Design Report, 2004, DØ Note 4415.

[27] "DØ Sillicon Tracker for Run IIa", E. Kajfasz, FERMILAB-CONF-01-382.
"Quelques expériences gravitant autour du problème de génération de masses", E. Kajfasz, Thèse de HDR, CPPM, mai 2002.

[28] "Design Report for the Central Fiber Tracker", the DØ coll., FERMILAB-PUB-98-095-E, hep-ex/9803023.

[29] `http://www-d0.fnal.gov/Run2Physics/WWW/results/final/NIM/N10B/`

[30] `http://www-d0.fnal.gov/Run2Physics/WWW/results/final/NIM/N09B/`

[31] `http://www-d0.fnal.gov/solenoid/`.

[32] "Technical Design Report for the Central Preshower", M. Adams et al., DØ note 3014.

[33] "Technical Design Report for the Forward Preshower", A. Gordeev et al., DØ note 3445.

[34] "Level 1 Trigger with pseudo AND/OR Terms", D. Edmunds et al., DØ note 3683.

[35] "B Physics with the DØ detector Upgrade in Run II", A. Lucotte, Nucl. Instr. Meth. A446(169-175), 2000.

[36] "The DØ Detector", S. Abachi et al., Nucl. Instr. and Methods, A338 (1994) 185.
"The DØ Upgrade, S. Abachi et al. FERMILAB-CONF-95-177-E."

[37] "Technical Design Report for the Upgrade of the ICD for DØ Run II", K. De et al. DØ note 2686.

[38] "DØ Calorimeter Upgrades for Tevatron Run II", Leslie Groer for the DØ Collaboration, DØ note 4240.

[39] "Technical Design Report for the Central Muon System", B. Baldin et al., DØ note 3365(1997).

[40] "Technical Design Report for the Forward Muon Tacking", G. Alexeev et al., DØ note 3366.

[41] M. Abolins et al, IEEE Trans. Nucl. Science 51 issue 3, 340 (2004).

[42] J. Bystricky et al, IEEE Trans. Nucl. Science 51 issue 3, 351 (2004).

[43] GEANT - CERN program library long writeup W5013, M. Goossens et al., `http://wwwinfo.cern.ch/asdoc/pdfdir/geant.pdf`.

[44] `http://www-d0.fnal.gov/computing/MonteCarlo/simulation/d0sim.html`.

[45] `http://root.cern.ch/drupal/`.

[46] http://www-d0.fnal.gov/Run2Physics/cs/caf/.

[47] "HTF : histogramming method for finding tracks. The algorithm description", A. Khanov, 2000, DØ Note 3778.

[48] "Central track extrapolation through the D0 detector", G. Hesketh, 2003, DØ Note 4079.

[49] "Primary Vertex Reconstruction by Means of Adaptive Vertex Fitting", A. Schwartzman and C. Tully, 2005, DØ note 4819.

[50] http://www-d0.fnal.gov/phys_id/emid/d0_private/emid_intro.html.

[51] "MC studies of jet algorithms in the RUN II framework", 2000, L. Duflot and R. Hirosky, DØ note 3746.

[52] "A Comparison of Cone and KT Algorithms for the Reconstruction of a Higgs Boson at 110 GeV", 1998, A. Pierce, DØ Note 3486.

[53] http://www-d0.fnal.gov/phys_id/emid/d0_private/variables.html.

[54] "Jets in RunII", 2002, The JETMET group - G. Bernardi et al, DØ Note 3985.

[55] "Jet Energy Scale at D0 , B. Abbott and all., 1997, DØ Note 3287.

[56] http://www-d0.fnal.gov/%7Edaniel/jesgroup/jesgroup.html.

[57] "Corrections for muonic b-jets in P17", 2007, J. Lellouch, DØ Note 5486.

[58] "Towards MissingET certication and unclustered energy studied", 2005, S. Calvet and al. DØ Note 4927.

[59] "Triple Timing Analysis of Calorimeter Read-out", N. Buchanan, B. Calpas, W. Geist, R. Nayyar, D. Schamberger, 2007, DØ note 5494.

[60] "Optimisation of the DØ Online Calorimeter Calibration for RunII", 2001, R. Chiche and al., DØ Note 3914.

[61] "Study of the Non Linearity of the D0 Calorimeter Readout Chain ", Robert Zitoun, 2002, DØ Note 3997.

[62] "Energy Scale studies and calibration of the D0 electromagnetic calorimeter using Z^0 and $J/Psi \rightarrow e^+e^-$ Run II events", Smain Kermiche, Aurélien Mendès, Marie-Claude Cousinou, Alexis Cothenet, 2005, DØ note 4945.

[63] "Triple Sample Studies", B. Calpas, http://www-d0.hef.kun.nl///fullAgenda.php?ida=a071452#2007-08-23.

[64] http://d0dbweb.fnal.gov/sam_data_browsing/ProjectDefinitions.html.

[65] http ://www-d0.hef.kun.nl//askArchive.php ?base=agenda&categ=a071114 &id=a071114s1t1/transparencies/.

[66] "Electron and Photon Identification with p20 data", 2008, O. Atramentov, D. Bandurin, X. Bu, B. Calpas, E. Carrera, D. Duggan, A. Ferapontov, M. Takahashi, T. Uzbyakova, H. Yin, DØ note 5761.

[67] "Single Electron Efficiencies in p17 Data and Monte-Carlo Using p18.05.00 d0correct", J. Hays et al., DØ note 5105.

[68] "Electron Likelihood in p20", 2009, M. Aoki, DØ Note 5675.

[69] "Articial neural network using central preshower detector information for electron and photon selection", 2008, X. Bu and al., DØ Note 5650.

[70] "Study of electron track matching parameters with Run IIb data", O. Atramentov and al., DØ note 5635.

[71] "Utilizing CFT and SMT hits count for photon and electron reconstruction", 2004, O. Atramentov, Y. Maravin, DØ note 4444.

[72] `http://www-d0.fnal.gov/phys_id/tauid/tauid.html`.

[73] "Reconstruction of τ leptons in hadronic final states at DØ Run 2", 2003, The DØ τ-id group, D. Chakraborty and al., DØ Note 4210.

[74] "Single Electron Efficiencies in p17 Data and Monte-Carlo Using p18.05.00 d0correct", J. Hays and al., 2006, DØ Note 5105.

[75] Sohrab Hossain "L3 Tau Trigger", `http://www-d0.hef.kun.nl//fullAgenda.php?ida=a0740\&fid=93`.

[76] "τ identication with neural networks for p17 data", 2006, S. Protopopescu and P. Svoisky, DØ Note 5094.

[77] "Search for $ZH \rightarrow e^+e^-b\bar{b}$ and $ZH \rightarrow \mu^+\mu^-b\bar{b}$ Production in 4.2 fb_{-1} of data with the DØ Detector in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", Conference Note 5876-CONF, 2009, `http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/HIGGS/H71/`.

[78] www-d0.fnal.gov/computing/data_quality.

[79] D0 Data Quality Coordination, `http://www-d0.fnal.gov/computing/data_quality/d0_private/forusers.html`.

[80] "PYTHIA 6.2 Physics and Manual", T. Sjostrand et al., arxiv :hep-ph/0108264.

[81] M.L. Mangano et al., ALPGEN, a generator for hard multiparton process in hadronic collisions, JHEP 07, 001 (2003).

[82] "Matching parton showers and matrix elements", S. Hoche et al., arXiv :hepph/0602031 (2006).

[83] "Heavy avour removal and determination of weighting factors for W+ jets Monte-Carlo", D. Gillbert, 2006, DØ Note 5129.

[84] Mika Vesterinen, "EM energy calibration and smearing", EMID meeting during collaboration week, February 04, 2010, https ://plone4.fnal.gov/P1/D0Wiki/object-id/emid/emtool/emresolution_cafe/EMresolution_cafe_v00-07-02.

[85] `http://en.wikipedia.org/wiki/Crystal_Ball_function`.

[86] "ICR Electron Efficiencies for RunIIa", B. Calpas, J.Kraus, T. Yasuda, DØ note 5939.

[87] "p20 ICR Electron Identification", J. Kraus et al., DØ note 5691.

[88] "Generator Level Reweighting of pT of Z Boson", M. Shamim and T. Bolton, DØ Note 5565.

[89] "The Relative Data-Monte Carlo Jet Energy Scale", N. Makovec, J.-F. Grivaz, DØ Note 4807 (2005).

[90] "Shifting, Smearing and Removing Simulated Jets", N. Makovec and J.-F. Grivaz, DØ Note 4914 (2005).

[91] "Studies of Alpgen parametes, corrections and associated uncertainties", W. Fisher, J. Haley, D. Price, DØ Note 5966.

[92] M. Rangel, "Vertex Confirmation SF for Jets", CALGO Meeting, May 20, 2009, `http://www-d0.hef.kun.nl//fullAgenda.php?ida=a09823`.

[93] J-F. Grivaz, Talk in the February 12, 2008 VJets working meeting, http ://www-d0.hef.kun.nl//askArchive.php ?base=agenda&categ=a08266&id=a08266s1t4/ transparencies/K-factors.pdf.

[94] J. Campbell, R. K. Ellis, Phys Rev D, Vol 60, 113006 (1999). `http://mcfm.fnal.gov/`.

[95] M. L. Ciccolini, S. Dittmaier and M. Krmer, Phys. Rev. D. Vol 68, 073003 (2003).

[96] `https://plone4.fnal.gov/P1/D0Wiki/comp/caf/caffaq/LumiReWeight`.

[97] "The longitudinal shape of the luminous region at DØ', H. Schellman, DØ Note 5142.

[98] `https://plone4.fnal.gov/P1/D0Wiki/tdaq/tsg/caftrigger/singleetriggers`.

[99] "Reweighting of Alpgen's Z boson transverse momentum in multijet topologies in p17", S. Calvet and al, DØ Note 5367.

[100] "Taggability in P14 pass2 data", C. Gerber, E. Shabalina, G. Otero y Garzon, 2006, DØ Note 4995.

[101] "Studies of Taggability vs. nPV for p17", I. Anghel, DØ Note 5240.

[102] "A Neural Network b-tagging Tool", T. Scanlon, DØ Note 4889.

[103] "Performance of the Dzero 439 NN b-tagging Tool on p20 Data", T. Gadfort and al., D0 Note 5554, December 19, 2007.

[104] "Update on b-quark jet identication with Secondary Vertex reconstruction using D0reco version p13", 2003, L. Feligioni and al., DØ Note 4182.

[105] "Performance of the SVT b-tagging algorithm on P17 data", 2007, J. Clutter, DØ Note 5385.

[106] "Update of the JLIP b-tagger Performance in p14/pass2 with JES 5.3", D. Bloch, B. Clement, DØ Note 4824.

[107] "Jet Lifetime b-tagging", D.Bloch and al. DØ Note 4069.

[108] Measurement of b-tagging efficiency and mis-tagging rates with CSIP method", R.Demina and al., DØ Note 4432.

[109] "b-tagging with Counting Signed Impact Parameter method", R.Demina and al., DØ Note 4049.

[110] "Search for $ZH \rightarrow l^+l^-b\bar{b}$ in 4.2 fb−1 of Data Using a Random Forest in Dimuon and Dielectron Events.", L. Ancu, J. BackusMayes, B. Calpas, K. Chan, S. Desai, M. Kirby, M. Mulhearn, E. Nagy, M. Naimuddin, E. Strauss, T. Yasuda, 2010, http ://www-d0.fnal.gov/Run2Physics/higgs/zh_llbb/d0_private/drafts/official/ zh_llbb_analysis_v1.4.pdf.

[111] S. Crepe-Renaudin, 2002, DØ Note 4023.

[112] M. Agelou, J.-L. Agram et C. Royon, 2005, DØ Note 4775.

[113] Toolkit for Multivariate Data Analysis with ROOT, `http://tmva.sourceforge.net/`.

[114] "Evidence for production of single toop quarks at DØ' V. M. Abazov et al., DØ Collaboration, Phys. Rev. D 78 (2008) 012005 [arXiv : 0803.0739 [hep-ex]].

[115] http ://www-d0.fnal.gov/Run2Physics/top/public/fall06/singletop/ DT_public_webpage/DecisionTrees.html#dt.

[116] S. Parke and S. Veseli, "Distinguishing WH and Wbb production at the Fermilab Tevatron" Phys Rev D Vol 60 (1999) 093003. (hep-ph/9903231).

[117] K. Herner, 3 Feb. 2009. http ://www-d0.hef.kun.nl //askArchive.php ?base=agenda&categ=a09205&id=a09205s1t0/transparencies.

[118] R.D. Cousins, G.J. Feldman, Phys. Rev. D57, 3873 (1998).

[119] "Systematics and Limit Calculations", 2006, W. Fisher, DØ Note 5309.

[120] "Calculating limits for combined analysis", 2006, W. Fisher, DØ Note 4975.

[121] "COLLIE : a condence level limit evaluator", 2008, W. Fisher, DØ Note 5395.

[122] T. Junk, Nucl. Intrum. Meth. A 434 (1999).

[123] A. Read, J. Phys. G : Nucl. Part. Phys. 28 (2002), 2693.

[124] A. Read, First workshop on confidence limits, 2000, CERN, Geneva.

[125] J. Pumplin et al., JHEP 0207 (2002) 012.

[126] D. Stump et al., JHEP 0310 (2003) 046.

[127] T. Junk, Building a more general χ^2, CDF/DOC/STATISTICS/PUBLIC/7904.

[128] T. Junk, Nucl. Inst. Meth. A434, 435 (1999).

[129] "Search for $ZH \to l^+l^-b\bar{b}$ in 4.1 fb^{-1} of p17 and p20 data using a Boosted Decision Tree in dimuon and dielectron events.", L. Ancu, J. BackusMayes, B. Calpas, K. Chan, S. Desai, M. Kirby, M. Mulhearn, E. Nagy, E. Strauss, T. Yasuda, 2009, DØ note 5862.

[130] "Search for $ZH \to e^+e^-b\bar{b}$ and $ZH \to \mu^+\mu^-b\bar{b}$ Production in 4.1 fb^{-1} of Run II data with the DZero Detector in p-pbar Collisions at $\sqrt{s} = 1.96$ TeV.", L. Ancu, J. BackusMayes, B. Calpas, K. Chan, S. Desai, M. Kirby, M. Mulhearn, E. Nagy, E. Strauss, T. Yasuda, 2009, DØ note 5876.

[131] "Search for ZH Production in the $e^+e^-b\bar{b}$ Final State - Preselection", B.Calpas, K.Chan, S.Desai, M.Mulhearn, E.Nagy, T.Yasuda, 2009, DØ note 5930.