N° d'ordre UCBL : 81-2003 Année 2003
N° d'ordre CPPM : CPPM-T-2003-04
DESY-THESIS-2003-023

THESE

présentée

devant l'UNIVERSITE CLAUDE BERNARD - LYON 1

pour l'obtention

du DIPLOME DE DOCTORAT

(arrêté du 25 mars 2002)

présentée et soutenue publiquement le

27 juin 2003

par

Matti PEEZ

Recherche de déviations au Modèle Standard dans les processus de grande énergie transverse sur le collisionneur électron - proton HERA

Search for deviations from the Standard Model in high transverse energy processes at the electron-proton collider HERA

JURY :

M.	E. ASLANIDES	Président
M.	J. DUMARCHEZ	Rapporteur
M.	M. KNECHT	
M.	G. SAJOT	Rapporteur
M.	G. SMADJA	
M.	C. VALLEE	Directeur de Thèse

Table des matières

Introduction

Le Modèle Standard est une théorie impressionante décrivant précisément les résultats de pratiquement toutes les expériences de physique des particules réalisées jusqu'à présent. Il a notamment été vérifié avec succès dans les expériences au LEP, au TEVATRON et à HERA, unique collisionneur électron-proton au monde. Cependant de nombreuses lacunes subsistent dans ce modèle et il est maintenant généralement admis qu'il ne représenterait que la manifestation à basse énergie d'une théorie plus fondamentale. Grâce à l'énergie disponible sur le collisionneur HERA, il est donc possible, en parallèle aux mesures de précisions du Modèle Standard, d'y découvrir des phénomènes nouveaux en physique des particules.

L'objet de cette thèse est la recherche de processus rares et de signaux de nouvelle physique. L'approche habituelle consiste à rechercher les phénomènes de nouvelle physique prédits dans le cadre d'un modèle particulier. La non observation d'un signal permet alors de déduire des limites sur les paramètres libres du modèle testé.

L'analyse présentée ici poursuit l'idée d'une recherche modèle-indépendante, sans à-priori sur la nature ou les propriétés des processus prédits théoriquement. Des particules nouvelles étant susceptibles de se manifester à la limite du domaine cinématique accessible et à grande masse, nous limitons cette recherche aux phénomènes énergétiques, c'est-à-dire de grande énergie transverse. Nous considérons toutes les combinaisons possibles d'électrons, de muons, de jets hadroniques, de neutrinos et de photons dans l'état final. Les événements sont regroupés en différentes classes, selon leur topologie. Les observables pertinentes, comme la masse invariante et la somme des énergies transverses, sont étudiées en détail dans chaque canal et comparées aux prédictions du Modèle Standard.

Afin d'améliorer la sensibilité et la précision statistique, l'ensemble des données H1 de HERA-I (118,3 pb^{-1}) ont été utilisées. L'analyse des données a été effectuée dans un nouvel environnement d'analyse orienté objet, développé avant tout pour la phase haute luminosité de HERA qui démarre actuellement. Le développement et les tests de ces nouveaux outils d'analyse sont partie intégrante de notre travail.

La description du cadre théorique est abordée dans le premier chapitre, avec une attention particulière aux processus du Modèle Standard intervenant dans notre analyse. Le second chapitre décrit l'accélerateur HERA et l'expérience H1. Les chapitres suivants exposent les structures et les algorithmes de reconstruction des particules physiques développés dans le cadre de l'environnement orienté objet. Dans le cinquième chapitre, l'analyse des topologies des événements à grande énergie transverse est présentée. Enfin, le dernier chapitre est consacré à la quantification de l'accord global des observations au Modèle Standard, en utilisant une méthode statistique développée pour l'occasion.

Chapitre 1

Physique à HERA

Audaces fortuna juvat.

Lors d'une collision électron-proton, différentes réactions peuvent se produire. Nous détaillons dans ce chapitre les différents processus du Modèle Standard (abrégé MS par la suite) à HERA qui sont susceptibles de contribuer dans les différents canaux analysés. L'exposé est illustré de visualisations d'événements enregistrés par le détecteur, qui sera décrit dans le chapitre suivant. Nous appelons par la suite $\mathrm{P}_{T,i}$ l'impulsion transverse, θ_i l'angle polaire et ϕ_i l'angle azimutal de la particule i. Le repère utilisé, cartésien, est décrit dans la partie 2.2 : l'axe z est défini positif dans la direction incidente du faisceau de protons et x et y dans le plan transverse à celui-ci.

1.1 Diffusion profondément inélastique

La diffusion profondément inélastique (DIS en anglais) correspond à la diffusion d'un électron sur un quark provenant du proton. Ce processus peut être décrit au premier ordre dans le cadre du modèle des partons. Il permet de sonder de façon précise la structure du proton. Les diagrammes de la figure 1.1 illustrent ce processus. L'interaction entre l'électron incident et le proton se fait soit à travers un photon virtuel γ ou un boson Z^0, soit à travers un boson $\mathrm{W}^\pm$. Dans le cas du photon ou Z^0 on parle de diffusion par Courant Neutre (abrégé NC en anglais) et la topologie finale comprend un électron et une gerbe hadronique collimée appelée jet, issue du quark éjecté. On peut observer l'électron diffusé dans le détecteur. La figure 1.2 montre un événement NC visualisé dans le plan longitudinal et transverse. Dans le cas d'un échange de $\mathrm{W}^\pm$, on parle de diffusion par Courant Chargé (abrégé CC) et on trouve dans l'état final un jet et de l'énergie manquante. Cette topologie est visible sur la figure 1.3. Dans ces deux cas, les restes du proton peuvent en s'hadronisant former un jet supplémentaire très à l'avant. Les particules participant à la réaction peuvent aussi rayonner un photon ou, pour le quark, un gluon. La topologie finale peut donc aussi comporter des photons ou plusieurs jets. En résumé, les courants neutres peuvent donner naissance aux topologies de *e-j*, *e-j-j*, *e-j-j-j*, *e-j-j-j-j* et *e-j-*γ en cas de radiation électromagnétique. Les topologies des courants chargés sont *j-*ν, *j-j-*ν, *j-j-j-*ν et *j-*ν*-*γ. Ces processus ont été étudiés de manière exhaustive dans [1, 2, 3, 4] et récemment dans [5].

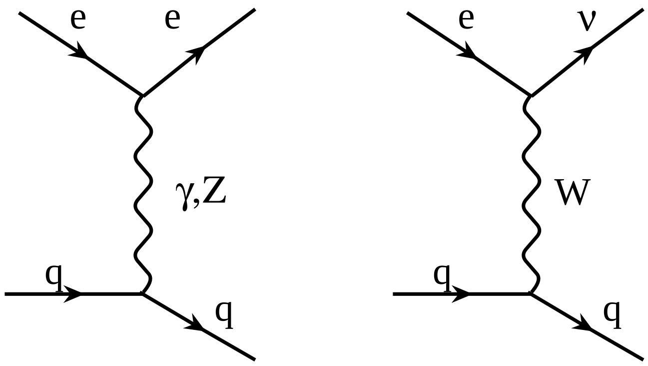


FIG. 1.1 – Diagrammes de diffusion profondément inélastique par Courant Neutre et Courant Chargé.

La cinématique de l'interaction est décrite de la façon suivante : soit k le quadrivecteur de l'électron incident et k' celui du lepton diffusé (électron ou neutrino), l'impulsion transférée au proton est alors q = k - k'. En notant p le quadrivecteur du proton incident, on peut introduire les grandeurs cinématiques suivantes :

- $Q^2 = -\, \mathrm{q}^2$:
  Masse virtuelle du boson échangé
- $x = \frac{-k\cdot q}{p\cdot q} = \frac{Q^2}{2p\cdot q}$:
  La fraction d'énergie d'impulsion longitudinale du proton emportée par le quark dans le référentiel où le proton a un moment infini
- $y = \frac{p\cdot q}{p\cdot k}$:
  La fraction d'énergie perdue par l'électron incident dans le référentiel du proton au repos

Il y a deux variables indépendantes si l'on tient compte de la relation $Q^2 = x \cdot y \cdot s$, (s étant connue, avec $s = (k + p)^2$). On peut donc quantifier ce processus de diffusion si l'on connaît 2 des 3 variables Q^2, x ou y. Celles-ci sont mesurées à partir des produits de la réaction enregistrés dans le détecteur. Nous décrivons une méthode permettant de reconstruire ces variables dans les parties 4.2.2.1 et 4.5.2.1.

Dans le cadre du modèle des partons, la section efficace à courant neutre peut s'écrire sous la forme :

$$\frac{d^2\sigma_{NC}^{e^\pm p}}{dxdQ^2} = \frac{e^4}{8\pi x}\left[\frac{1}{Q^2}\right]^2 \Phi_{NC}^{\pm}(x,Q^2) \tag{1.1}$$

En négligeant la contribution du terme de fonction de structure longitudinal F_L, nous avons :

$$\begin{aligned}\Phi_{NC}^{\pm}(x,Q^2) = Y_+ x \Big[&\frac{4}{9}(u(x)+\overline{u}(x)+c(x)+\overline{c}(x))+\\ &\frac{1}{9}\left(d(x)+\overline{d}(x)+s(x)+\overline{s}(x)+b(x)+\overline{b}(x)\right)\Big]\end{aligned} \tag{1.2}$$

avec $Y_\pm = 1 + (1 \pm \mathrm{y}^2)$, et e la charge électromagnétique. $u(x)$, $\overline{u}(x)$, $c(x)$, $\overline{c}(x)$, $b(x)$, $\overline{b}(x)$, $s(x)$, $\overline{s}(x)$ sont les densités des quarks correspondants. La formule 1.1 nous montre que la section efficace de production dépend essentiellement du propagateur et est donc proportionelle à $1/\mathrm{Q}^4$ pour les courants neutres. A cause du terme d'interférence destructive du boson γ et Z dans les processus e^+p, la section efficace est réduite à grand Q^2 par rapport au processus e^-p, où l'interférence est positive. Pour ce processus ainsi que pour les suivants, nous donnons, dans l'acceptance du détecteur, le comportement visible et la section efficace. Dans le cas de collision e^+p, la section efficace totale, lorsque l'impulsion transverse de l'électron et du jet de plus grand P_T est supérieure à 20 GeV est de 379 *pb*. Elle décroît approximativement en ${P_T^{-}}^4$.

Pour les courants chargés, la section efficace s'écrit :

$$\frac{d^2\sigma_{CC}^{e^\pm p}}{dxdQ^2} = \frac{g^4}{64\pi x}\left[\frac{1}{Q^2+M_W^2}\right]^2 \Phi_{CC}^{\pm}(x,Q^2) \tag{1.3}$$

où g est la constante de couplage électrofaible et M_W la masse du boson $\mathrm{W}^\pm$. $\Phi_{CC}^{\pm}(x,Q^2)$ est relié aux densités de quark par la relation suivante :

$$\Phi_{CC}^{+}(x,Q^2) = x[\overline{u}(x)+\overline{c}(x)] + (1-y)^2 x[d(x)+s(x)] \tag{1.4}$$

$$\Phi_{CC}^{-}(x,Q^2) = x[u(x)+c(x)] + (1-y)^2 x[\overline{d}(x)+\overline{s}(x)] \tag{1.5}$$

Nous remarquons d'après la formule 1.3 que la section efficace à courant chargé dépend essentiellement du propagateur $\frac{1}{Q^2+M_W^2}$. Elle décroît donc moins rapidement que pour les courants neutres, approximativement en ${P_T^{-}}^3$, et vaut environ 19 *pb*, pour une impulsion transverse manquante et un jet d'impulsion transverse au-delà de 20 GeV.

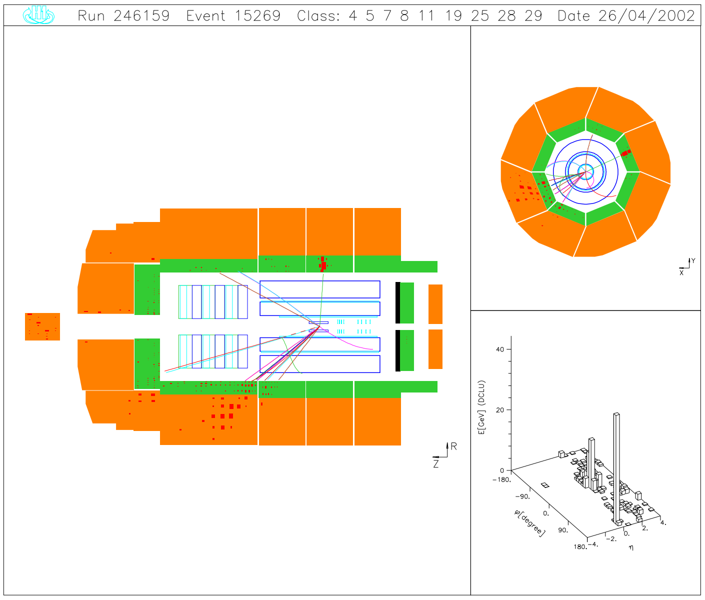


FIG. 1.2 – Un exemple d'événement NC. Sur la vue dans le plan transverse (en haut à droite), les directions des dépôts d'énergie sont opposées, ce qui indique que l'énergie transverse est probablement conservée.

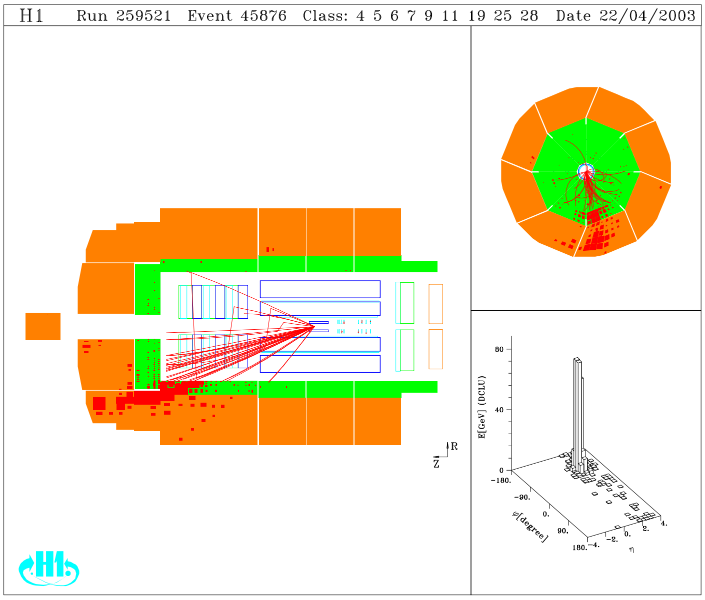


FIG. 1.3 – Un exemple d'événement CC. Sur la vue dans le plan transverse, l'absence de dépôt d'énergie vers le haut de la figure montre un déséquilibre de la quantitée d'énergie transverse déposée. Ce déséquilibre indique qu'une particule (un neutrino) n'a pas été détectée.

1.2 Photoproduction

Le terme "photoproduction" regroupe tous les processus à courant neutre pour lesquels l'impulsion transférée $Q^2 \approx 0$. Ils sont souvent appelés processus γp : le photon échangé est quasi-réel.

A HERA, le processus de photoproduction, dont un exemple est montré dans la figure 1.4, est le processus dominant avec une section efficace totale de $\sigma_{\gamma p} \approx 150\ \mu$b. La plupart de ces processus sont "mous", c'est-à-dire avec une impulsion transverse petite (quelques GeV) de l'état final. Dans l'analyse effectuée ne seront considérés que les processus dits "durs" pour lesquels on retrouve des jets énergiques. Lors de ces événements, l'électron incident est presque toujours perdu dans le tube à vide des faisceaux. Ce processus contribue donc principalement aux canaux à jets, comme j-j, j-j-j et j-j-j-j (lorsqu'il y a radiation de gluons). La section efficace de photoproduction de jets dans les collisions électron-proton, σ_{ep}, peut être calculée à partir de la section efficace de collision photon-proton, $\sigma_{\gamma p}$, en utilisant l'ansatz :

$$\sigma_{ep \to eX} = \int dy f_{\gamma,e}(y) \sigma_{\gamma p}(y, Q^2) \tag{1.6}$$

où $f_{\gamma,e}(y)$ est le flux de photon issu de l'électron. Dans ce cadre, la variable usuelle y de diffusion profondément inélastique est interpretée comme la fraction de l'impulsion longitudinale de l'électron emportée par le photon. Dans l'approximation de Weizsäcker-Williams, $f_{\gamma,e}(y)$ s'écrit :

$$f_{\gamma,e}(y) = \frac{\alpha}{2\pi} \left[\frac{1 + (1-y)^2}{y} log \frac{Q^2_{max}(1-y)}{m_e^2 y^2} + 2 m_e^2 y \left(\frac{1}{Q^2_{max}} - \frac{1-y}{m_e^2 y^2} \right) \right] \tag{1.7}$$

où α est la constante structure fine et m_e la masse de l'électron. La virtualité maximale Q^2_{max} est donnée par les conditions expérimentales et est reliée à travers l'expression $Q^2 = 2E_e E_e^{'}(1+\cos\ \theta_e)$ à l'énergie $E_e^{'}$ et l'angle polaire θ_e de l'électron diffusé (E_e étant l'énergie de l'électron incident). Pour deux jets avec une impulsion transverse au-dessus de 20 GeV, la section efficace totale vaut 321 *pb*. Elle décroît en fonction des impulsions transverses selon P_T^{-8}. On distingue deux sous-processus à l'ordre dominant : la photoproduction directe et résolue. Des analyses dédiées ont été menées pour mesurer les sections efficaces de ces processus [6, 7, 8]. Ce processus sera dominant par rapport au processus de diffusion profondément inélastique dans les canaux à jets et spécialement à grande impulsion transverse et masse invariante.

1.2.1 Photoproduction directe

Le processus direct est décrit par le couplage du photon en tant que particule élémentaire à un parton du proton. Dans l'état final, nous trouvons, au premier ordre, 2 jets. La diffusion photon-gluon, dont le diagramme est montré sur la figure 1.5, domine la section efficace. Aux ordres supérieurs, des radiations de gluons peuvent donner plus de jets dans l'état final. La section efficace de la photoproduction directe décroit légèrement moins vite que celle de la photoproduction résolue en fonction de P_T. Sa contribution dans les canaux à jets sera donc plus importante.

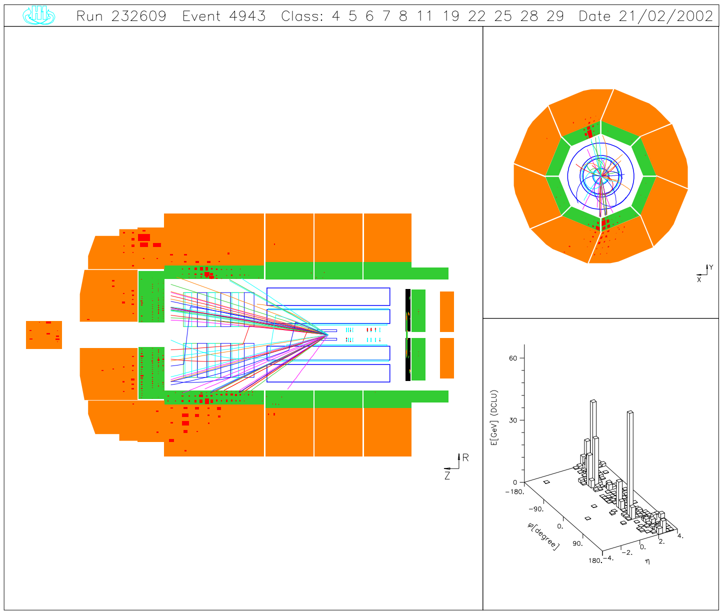


FIG. 1.4 – Un exemple d'événement de photoproduction. Sur la vue transverse, les deux jets de directions opposées correspondent à l'hadronisation de la paire de quarks ou gluons produite.

1.2.2 Photoproduction résolue

Le processus résolu, en revanche, est caractérisé par une fluctuation du photon en paire $q\overline{q}$. Ce processus, résolvant la structure du photon, est décrit dans le cadre du modèle de dominance de mésons vecteurs (VMD en anglais) où le photon fluctue en méson V, comme par exemple un ρ, qui a les mêmes nombres quantiques que lui. Les processus dominants sont $qg \to qg$, qui est illustré sur la figure 1.5, et $gg \to gg$. L'un des quarks provenant de la structure hadronique du photon interagit par interaction forte avec l'un des partons du proton soit directement, soit par échange d'un gluon.

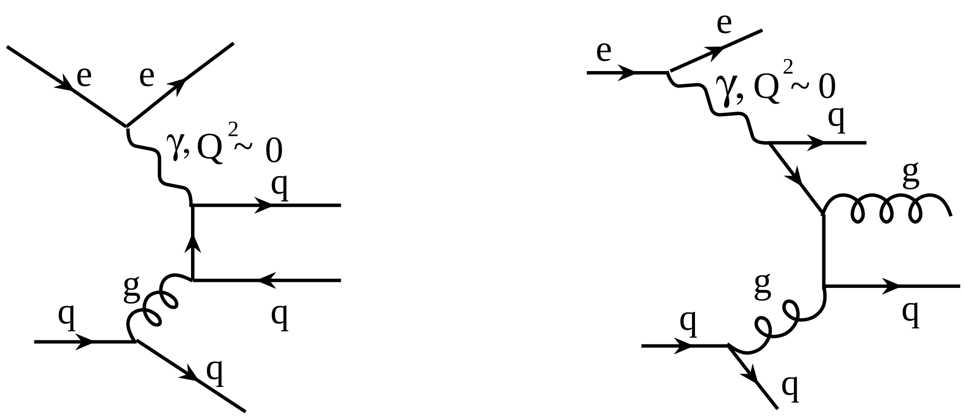


FIG. 1.5 – Exemples de diagrammes de photoproduction directe (à gauche) et résolue (à droite).

1.3 Processus radiatifs électromagnétiques

1.3.1 Processus Bethe-Heitler

Il s'agit d'un processus de radiation par Bremsstrahlung, dans lequel l'électron émet un photon dans le champ du proton. L'électron aussi bien que le photon sont diffusés à très bas angle. Le processus $ep \rightarrow e\gamma p$ domine largement : la structure du proton n'est pas résolue. Comme la section efficace est indépendante de la structure du proton, ce processus est utilisé à HERA pour mesurer la luminosité.

1.3.2 QED Compton

La diffusion Compton est décrite par l'émission par le proton d'un photon quasi-réel qui interagit avec l'électron initial (diagramme 1.6).
Il existe une contribution élastique dans laquelle le proton reste intact, et une contribution inélastique, où le photon initial est émis par un quark du proton incident. Dans l'état final nous reconstruisons un électron et un photon. Un deuxième photon peut être rayonné par l'électron final. Ce processus contribue donc principalement aux canaux e-γ et e-γ-γ.

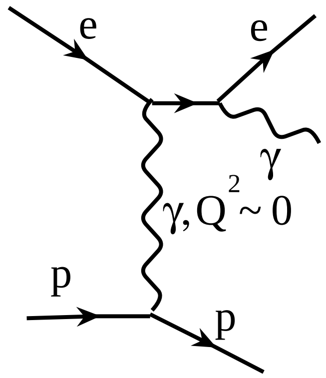


FIG. 1.6 – Diagramme de diffusion Compton QED.

La section efficace de production décroît approximativement en P_T^{-4} et vaut 2,1 *pb* pour des impulsions transverses de l'électron et du photon supérieures à 20 GeV. Nous trouvons de très grandes masses invariantes $\mathrm{M}_{\gamma e}$ lorsque les impulsions du photon et de l'électron sont très importantes ou lorsque l'électron et le photon sont diffusés à de très grands angles. Ce processus est signé par la présence de deux amas électromagnétiques équilibrés en Φ, l'un possédant une trace associée (électron). Un événement typique est montré sur la figure 1.7. Des analyses spécifiques sont présentées dans les références [9, 10].

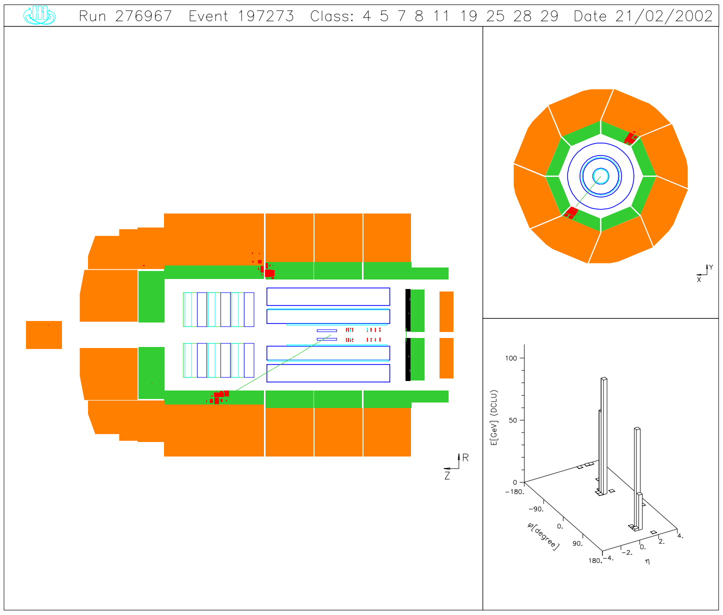


FIG. 1.7 – Un événement QED Compton élastique avec l'électron qui est émis vers le bas du calorimètre et le photon vers le haut.

1.3.3 Processus γ-γ

Les processus γ-γ sont à HERA la source principale de production de paires de leptons. Le diagramme de production est donné sur la figure 1.8. La section efficace de la production de paires de leptons d'impulsion transverse supérieure à 20 GeV est 0,09 *pb*. Elle décroit rapidement en fonction des impulsions transverses des leptons produits, avec une dépendance en P_T^{-8}.
Ces processus γ-γ s'appellent aussi processus 2γ-Bethe-Heitler et appartiennent à la catégorie des processus radiatifs : un photon venant du proton interagit avec un photon émis par l'électron pour donner une paire de leptons. Dans le cas des électrons on peut trouver trois électrons dans l'état final lorsque l'électron incident est diffusé à suffisamment grand angle pour être détecté. Comme dans les processus de diffusion Compton, l'interaction peut être élastique ou inélastique.

La production de paire d'électrons (en paires et en triplets) et de paires de muons a été étudiée dans deux analyses séparées [11, 12] : un bon accord entre les données et les prédictions a été observé ; cependant pour la production de paires d'électrons, des événements atypiques à grande impulsion transverse ont été reportés. L'identification de paires de taus est plus difficile [13] car ils se désintègrent préférentiellement en hadrons.
Ces processus contribuent donc principalement aux canaux e-e, e-e-e, e-μ-μ, μ-μ et e-μ. Un exemple de production d'électrons est donné sur la figure 1.9.

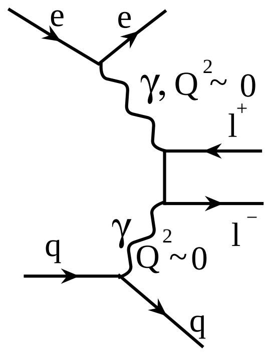


FIG. 1.8 – Diagramme de production de paire de leptons.

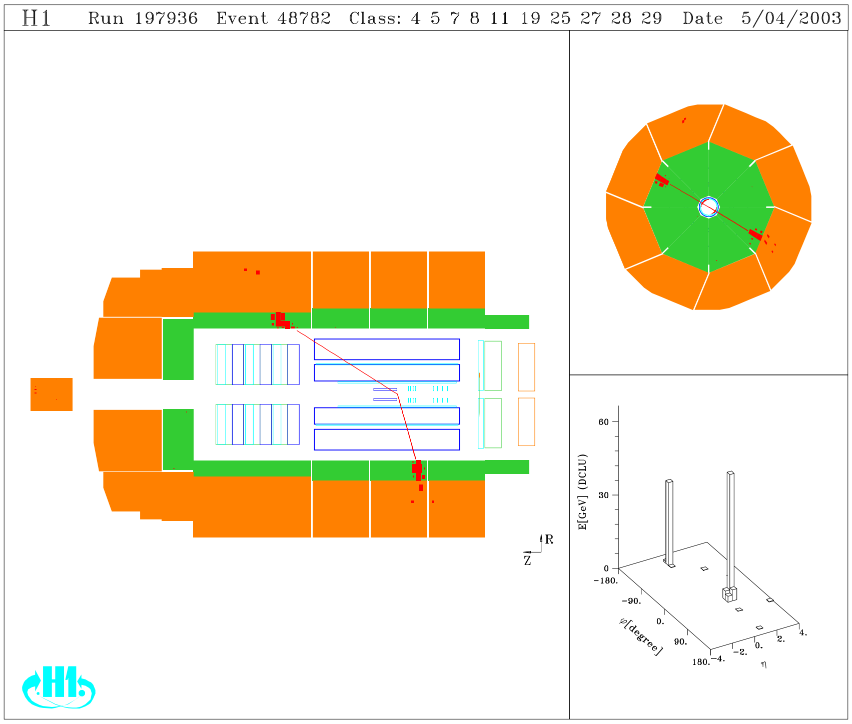


FIG. 1.9 – Un exemple d'événement photon-photon, l'interaction des deux photons a produit deux électrons visibles dans le calorimètre électromagnétique.

1.3.4 Processus Cabibbo-Parisi

Un autre processus pouvant produire des états finaux contenant une paire de leptons est le processus de Bremsstrahlung "Cabibbo-Parisi" schématisé dans la figure 1.10. La paire de leptons créée est issue de l'annihilation de l'électron incident avec un électron provenant de la matérialisation d'un photon émis par le proton. Ce processus prédit préférentiellement des paires d'électrons. La signature de ce processus est la même que celle des processus $\gamma - \gamma$, mais sa contribution est faible dans le domaine cinématique considéré.

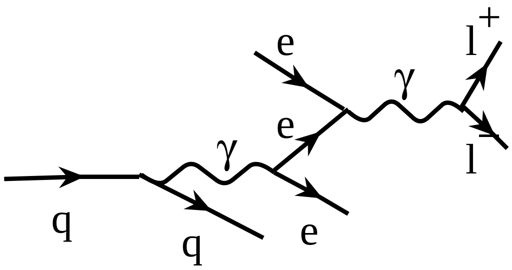


FIG. 1.10 – Diagramme de production de leptons par Cabibbo-Parisi.

1.3.5 Processus Drell-Yan

Les processus Drell-Yan (cf. figure 1.11) mettent en jeu un photon émis par l'électron incident, qui va interagir avec le proton par intermédiaire d'une paire quark-antiquark. Un antiquark issu du photon s'annihile avec un quark du proton pour donner une paire de leptons. La contribution du Drell-Yan à la production de paire de leptons en collision *ep* est négligeable [14].

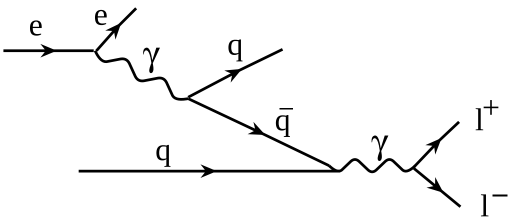


FIG. 1.11 – Diagramme de production de leptons par Drell-Yan.

1.4 Processus radiatifs faibles

L'énergie disponible dans le référentiel du centre de masse ($\sqrt{s} \approx 318$ GeV) permet de produire des bosons $W^{\pm}$ et Z^0 réels. Comme l'énergie initiale est en grande partie absorbée dans l'énergie de masse du boson lourd, ceux-ci sont produits de préférence avec des impulsions transverses petites.

1.4.1 Production de W$^\pm$

Le boson W$^\pm$ peut être produit lors de collision ep à travers les processus (figure 1.12) :

$$e^\pm p \to e^\pm X W^\pm \text{ et } e^\pm p \to \nu X W^\pm$$

Le deuxième processus a une section efficace 20 fois plus petite que le premier, dont la section efficace est déjà très faible : $\sigma_{tot}(e^\pm p \to e^\pm X W^\pm) \approx 1\ pb$. Cette section

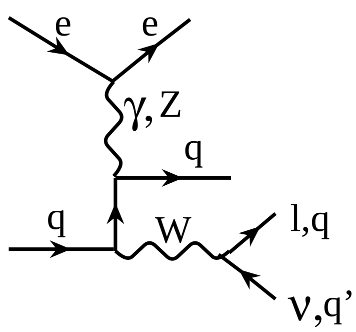


FIG. 1.12 – Diagramme de production de boson vecteur W.

efficace décroît rapidement avec l'impulsion transverse du boson $W^\pm$ (variation en P_T^{-7}), et le lepton et le neutrino issus du $W^\pm$ ont une impulsion transverse de l'ordre de 40 GeV. Le W$^\pm$ se désintègre de préférence en hadrons (2/3 des cas). Le tiers restant est partagé entre les 3 types de leptons. Lorsque le boson $W^\pm$ se désintègre en électron (ou muon) et un neutrino d'impulsion transverse d'au moins 20 GeV, la section efficace totale vaut à peu près 50 fb. La production de boson $W^\pm$ a été le sujet de plusieurs analyses aussi bien dans le canal de désintégration leptonique que dans le canal hadronique. Les résultats de H1 [15, 16, 17] montrent un bon accord global, sauf pour le domaine de P$_T^X > 40$ GeV avec un excès de 6 événements observés pour 1,08 $\pm$ 0,22 prédits. Cet excès d'événement n'est pour l'instant pas confirmé par l'expérience ZEUS qui a effectué une analyse similaire [18, 19].

Dans le détecteur la signature de la désintégration leptonique d'un $W^\pm$ (figure 1.13) est un muon ou électron, de l'impulsion transverse manquante importante (associée au neutrino) et un jet hadronique éventuel. Les taus n'ont pas encore fait l'objet d'étude car il sont très difficiles à identifier.
La contribution principale de ce processus intervient dans les canaux e-j-ν, μ-j-ν, e-ν et μ-ν. Sa contribution aux canaux multi-jets est négligeable par rapport à celle de la photoproduction.

1.4.2 Production de Z^0

De même, les bosons Z^0 peuvent être produits selon les réactions (figure 1.14) :

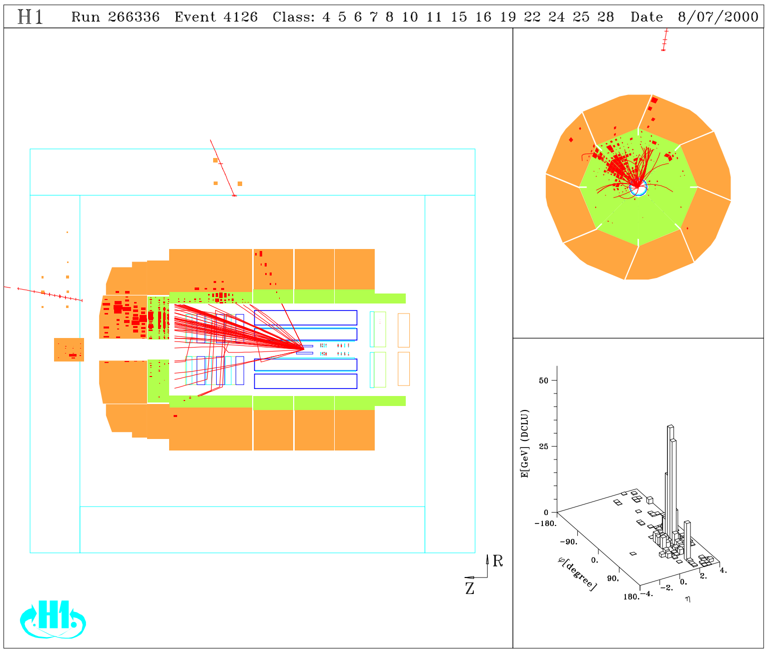


FIG. 1.13 – Un candidat pour le processus de production de $W^{\pm}$ (ici le $W^{\pm}$ se serait désintégré leptoniquement en un muon et son neutrino). L'électron diffusé n'est pas retrouvé car il a été diffusé à trop petit angle. Dans cet événement, le $W^{\pm}$ a une grande impulsion transverse, ce qui se traduit par l'observation du jet hadronique de recul dans le détecteur.

$$e^{\pm}p \rightarrow e^{\pm}XZ^0 \text{ et } e^{\pm}p \rightarrow \nu XZ^0$$

La masse du Z^0 étant de 91,2 GeV, la section efficace de production est encore plus petite que celle des $W^{\pm}$. Le Z^0 se désintégre principalement dans le canal hadronique et leptonique.

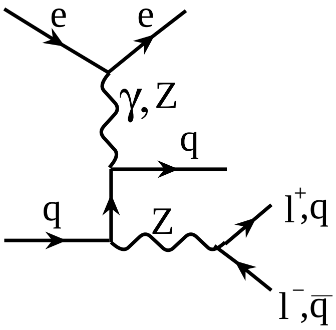


FIG. 1.14 – Diagramme de production de boson vecteur Z^o.

Nous nous attendons donc à trouver ces événements à partir d'une paire de jets très énergiques ou une paire de leptons avec des charges opposées et une masse invariante autour de 90 GeV. Leur contribution à la production totale de paires de leptons est faible par rapport au processus γ-γ.

1.5 Récapitulation

Nous avons énuméré les processus principaux du Modèle Standard intervenant dans l'analyse. Leurs topologies ainsi que leurs sections efficaces totales sont résumées dans le tableau 1.5. La plupart des topologies ont déjà fait l'objet d'études à grande impulsion transverse dans des analyses dédiées. Il s'agit des processus de diffusion profondément inélastique [1, 2, 3, 4, 5], de la production de jets en photoproduction [6, 7, 8], du processus de diffusion QED Compton [9, 10], de la production de paires de leptons [11, 12], et de la recherche de leptons isolés et impulsion transverse manquante [16].

Processus	Topologies principales	σ : deux corps de $P_T > 20$ GeV (pb)	Comportement visible en fonction de P_T^{corps}
DIS NC	e-j, e-j-j, e-j-j-j, e-j-j-j-j, e-j-γ	379	$\sim P_T^{-4}$
DIS CC	j-ν, j-j-ν, j-j-j-ν, j-ν-γ	19	$\sim P_T^{-3}$
γp	j-j, j-j-j, j-j-j-j, j-j-j-j-j	321	$\sim P_T^{-8}$
QED Compton	e-γ-γ, e-γ-γ	2,14	$\sim P_T^{-4}$
$\gamma\gamma \to e^+e^-, \mu^+\mu^-$	e-e, e-e-e, e-μ-μ, μ-μ, e-μ	0,18	$\sim P_T^{-8}$
$e^\pm p \to e^\pm X W^\pm_{\hookrightarrow l\nu}$	e-j-ν, μ-j-ν, e-ν, μ-ν	0,05	$P_T \sim 40$ GeV

TAB. 1.1 – Topologies principales, section efficace totale et comportement visible en fonction du P_T des objets produits par chaque processus.

1.6 Au-delà du Modèle Standard

Malgré son succès, le Modèle Standard n'explique pas tout et certaines questions restent en suspens :

- **Le Modèle Standard comprend un grand nombre de paramètres libres :**
  - Les masses des 12 fermions.
  - La phase et l'angle de la matrice de mélange des quarks (CKM).
  - Les constantes de couplage életromagnétique α et forte α_s.
  - Les masses des bosons $W^\pm$, Z^o et du boson de Higgs.
- **Les 3 générations de fermions** n'ont aucune explication dans le cadre du Modèle Standard. Le nombre de neutrinos légers a été contraint par le LEP à 3, mais rien dans le Modèle Standard ne justifie ce nombre de 3.

- **La gravitation**, l'une des 4 forces fondamentales, n'est pas incluse dans le Modèle Standard. Il ne peut donc pas décrire les phénomènes à très haute énergie où l'influence de la gravité n'est plus négligeable.

Ce nombre important de questions ouvertes laisse supposer l'existence d'une théorie sous-jacente régissant les relations entre les paramètres libres du modèle. Il faut donc rechercher une théorie étendant le Modèle Standard qui, sans remettre en cause ses résultats, vérifiés avec une précision extrême, permette d'aller au-delà. Nous allons donc rappeler brièvement plusieurs pistes théoriques possibles.

1.6.1 Les modèles composites

Pour résoudre les problèmes du Modèle Standard, celui-ci peut être considéré comme une théorie effective de phénomènes plus fondamentaux [20], impliquant qu'une particule au moins du Modèle Standard soit composite. Il existe trois catégories de théories distinctes selon que l'on attribue la structure composite aux bosons faibles, aux fermions ou au boson de Higgs.

1.6.2 La supersymétrie

L'approche de supersymétrie (SUSY) [21] est d'introduire une symétrie supplémentaire entre bosons et fermions en associant aux fermions des partenaires de spin entier et réciproquement, ce qui permettrait de décrire les particules et leurs interactions dans un seul groupe de jauge. Ce modèle très étudié aujourd'hui par les théoriciens permet de résoudre le problème de naturalité et de hiérarchie du Modèle Standard.

En outre, à partir de leurs valeurs mesurées à basse énergie par les expériences LEP, on peut calculer des constantes de couplages à plus haute énergie grâce aux équations du groupe de renormalisation. Dans le Modèle Standard, il n'existe pas d'énergie à laquelle les trois constantes sont égales, et une grande unification est impossible. En revanche, si le Modèle Standard est rendu supersymétrique, l'évolution de ces constantes change à cause de boucles faisant intervenir de nouvelles particules et indique qu'elles convergent vers une même valeur, à une énergie de l'ordre de 10^{16} GeV.

Le modèle le plus simple de supersymétrie est appelé MSSM (**M***inimal* **S***uper* **S***ymmetric* **M***odel*). Il associe à chaque particule du Modèle Standard (fermion ou boson) un partenaire supersymétrique. Les plus légers de ces partenaires auraient une masse inférieure au TeV mais aucun d'entre eux n'a pour l'instant été identifié expérimentalement.

1.6.3 Les dimensions supplémentaires

Nous sommes habitués à vivre dans un espace à 3 dimensions. Cependant l'existence de dimensions supplémentaires [22] n'est pas exclue à condition que celles-ci soient compactifiées (c'est à dire repliées sur elles-mêmes) pour expliquer le fait qu'elles n'aient pas été observées jusqu'à présent. Dans certains modèles de dimensions supplémentaires, où le volume de compactification est relativement grand, la valeur importante de l'échelle de Planck en 4 dimensions résulte d'un effet géométrique. L'échelle fondamentale de Planck dans l'espace complet serait du même ordre que l'échelle électrofaible, ce qui résoud le problème de hiérarchie.

1.6.4 Grande unification et théorie des cordes

Le Modèle Standard est fondé sur un groupe de symétrie qui, pour des masses inférieures à la masse d'unification, c'est à dire la masse à laquelle les couplages des 3 forces (forte, faible et électromagnétique) sont identiques, se décompose en $SU(3) \times SU(2) \times U(1)$. L'unification de trois forces [23] en une seule nécessite un groupe de dimension plus grande. Plusieurs groupes ($SU(5)$, $SO(10)$, etc) ont été étudiés. Certaines théories de grande unification (voir la revue faite dans [24]) prédisent en outre que, aux énergies où ces phénomènes se produisent, les particules ne sont plus ponctuelles mais comparables à des cordes vibrantes dont les différents modes reflèteraient les propriétés que nous observons. Ces théories sont basées sur des groupes plus complexes tels que $SO(32)$ ou $E_8 \times E_8$ et incluent souvent la gravitation. Elles requièrent un espace de 11 dimensions.

Ces modèles ne sont, à l'heure actuelle, pas vérifiables aux énergies atteintes par les accélérateurs de particules.

1.6.5 Approche modèle indépendante

Certains des modèles introduits précédemment, notament les modèles de supersymétrie et de leptons excités, ont été étudiés avec grande précision dans les expériences de collisionneurs à haute énergie. Les analyses ont permis de calculer des limites d'exclusion sur les couplages mis en jeu et sur la masse des particules hypothétiques produites. Dans ces recherches, les techniques de sélection sont adaptées aux caractéristiques des événements attendus pour un modèle particulier : les coupures de sélection sont optimisées pour augmenter l'efficacité de détection du signal. Il y a donc un risque d'ignorer certaines topologies potentiellement intéressantes. Une autre approche consiste à étudier les topologies des événements de manière générique et exhaustive, indépendamment des modèles de nouvelle physique susceptibles de les produire. Cette approche modèle indépendante a été utilisée dans l'expérience D0 [25, 26] et H1 [27, 28, 29] pour étudier différents états finaux exclusifs au sein d'une même analyse.

Dans notre analyse, présentée dans la partie 5, nous recherchons des signaux rares en nous appuyant sur les distributions en impulsion transverse et en masse invariante des particules identifiées dans l'état final. Pour chacun des canaux, la totalité des processus du Modèle Standard présentée dans ce chapitre est prise en compte et comparée aux données. La principale différence par rapport à l'analyse effectuée par l'expérience D0 est que nous étudions au sein de la même analyse de manière systématique toutes les topologies possibles contenant deux, trois et quatres corps (particules de grande énergie transverse). Un gain en sensibilité est atteint par rapport aux analyses précédentes de H1 en considérant toutes les topologies contenant des électrons, photons, muons, jets et neutrinos et en analysant l'ensemble des données de la phase I de HERA.

Chapitre 2

Le dispositif expérimental

Quid dignum tanto feret hic promissor hiatu.

Horace

Les données étudiées ont été prises avec l'expérience H1 sur l'accélérateur HERA à Hambourg.

2.1 L'accélérateur HERA

H1 est une expérience généraliste, située sur l'accélérateur circulaire HERA, sur le site de DESY à Hambourg (figure 2.1). HERA est le premier collisionneur ep jamais réalisé au monde. Il est constitué de deux anneaux indépendants de 6,3 km de périmètre, qui accélèrent les protons à une énergie de 920 GeV (820 GeV jusqu'en 1997) et les électrons ou positrons à une énergie de 27,6 GeV. L'énergie dans le référentiel du centre de masse électron-proton est $\sqrt{s} \approx 318$ GeV (300 GeV jusqu'en 1997).

Une des caractéristiques majeures de HERA est le faible temps séparant deux interactions, égal à 96 ns. A titre d'exemple, le temps de croisement était de 22 μs au LEP et sera de 25 ns au LHC. Les quatre expériences qui se trouvent sur le site de DESY sont HERA-B, HERMES (cible fixe pour les deux), ZEUS et H1. Entre 1992 et 2000 une luminosité d'environ 120 pb^{-1} a été accumulée par les expériences H1 et ZEUS. Cette période de prise de données est appelée “HERA I”. Entre 2000 et 2002, des améliorations importantes ont été apportées à l'accélérateur pour, en premier lieu, améliorer la luminosité instantanée d'un facteur 5, et en second lieu, polariser le faisceau d'électrons longitudinalement. La période de prises de données ayant débuté en 2002 est appelée “HERA II”.

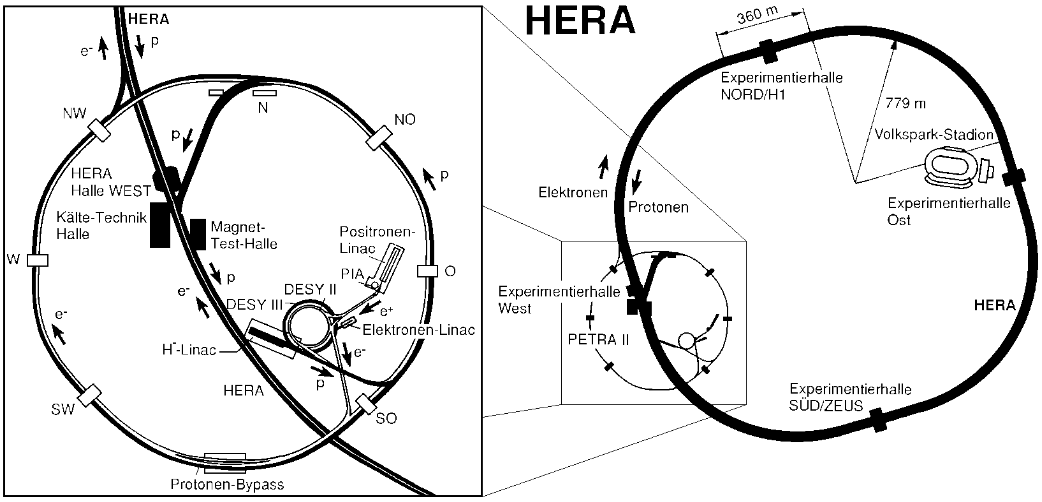


FIG. 2.1 – Vue schématique de HERA

2.2 Le détecteur H1

Le détecteur H1, schématisé dans la figure 2.2, se trouve au nord de HERA et occupe un des 4 points de croisement des faisceaux. Il s'agit d'un détecteur multi-fonctions hermétique et centré sur le point d'interaction, capable de mesurer l'énergie et l'impulsion des particules issues des collisions. Le détecteur pèse environ 2800 tonnes et fait 15 m de long. Comme le faisceau de protons est plus énergétique le détecteur a une forme asymétrique. Le système de référence utilisé par H1 est défini de la façon suivante :

- La direction incidente des protons définit l'axe z positif
- L'axe y est vertical et pointe vers le haut
- L'axe x est défini tel que le repère x, y, z soit direct
- L'angle Φ (azimutal) est défini dans le plan x-y à partir de la direction x
- L'angle Θ (polaire) est défini par rapport à l'axe z positif. Nous utiliserons plus souvent la pseudorapidité η définie par η = -log[tan($\theta/2$)]. Nous distinguerons la partie avant du détecteur correspondant à $\Theta \ll 90^o$ et la partie arrière correspondant à $\Theta \gg 90^o$.

Les principaux constituants du détecteur de H1 utilisés dans notre analyse sont :

- Les détecteurs de traces : ils déterminent l'impulsion des particules chargées et sont utilisés pour reconstruire le point de collision événement par événement. Les trajectographes (notés 2 et 3 sur la figure 2.2) peuvent être divisés en trois catégories :

  1. détecteurs à bandes en silicium qui permettent de déterminer la présence de saveurs lourdes, d'un vertex secondaire ou de valider le passage d'un électron.
  2. chambres proportionnelles à fils rapides (CIP-COP, BPC) qui servent au déclenchement du système d'acquisition de données.

3. chambres à dérive dites lentes (CJC1-CJC2, CIZ, COZ) qui servent à la reconstruction précise des trajectoires des particules chargées.

- Le calorimètre à argon liquide (noté 4 et 5 sur la figure 2.2) abrégé LAr en anglais : il permet de mesurer l'énergie et la direction des particules par rapport au vertex. La plupart des jets sont reconstruits dans le LAr, qui couvre la région angulaire $4° < \theta < 153°$; Le LAr est complété à l'arrière par un calorimètre (SpaCal) pouvant détecter les électrons diffusés à des angles polaires $153° < \theta < 178°$ (noté 12 sur la figure 2.2). A l'avant, il est complété par le bouchon calorimétrique appelé PLUG (noté 13 sur la figure 2.2) pouvant détecter les débris du proton.
- La bobine supraconductrice (notée 6 sur la figure 2.2) : elle crée un champ magnétique de 1,15 Tesla qui courbe les trajectoires des particules chargées, ce qui permet par la suite de déterminer l'impulsion des particules dans les détecteurs de traces.
- La structure en fer (notée 10 dans la figure 2.2) située à l'extérieur du détecteur : elle abrite des chambres à fils permettant de détecter des muons et de compléter la mesure des hadrons non absorbés dans le LAr.

Les calorimètres permettent la mesure de l'énergie des particules chargées et des particules neutres, tandis que les détecteurs de traces ne sont sensibles qu'aux particules chargées. En pratique, les calorimètres sont plus performants à haute énergie, où la résolution des détecteurs à traces devient faible. Le calorimètre LAr est donc la pièce maîtresse de notre analyse. Le détecteur de traces est aussi très important pour mesurer l'impulsion des muons et discriminer entre électrons et photons. La description détaillée des différents sous-détecteurs peut être trouvée dans [30].

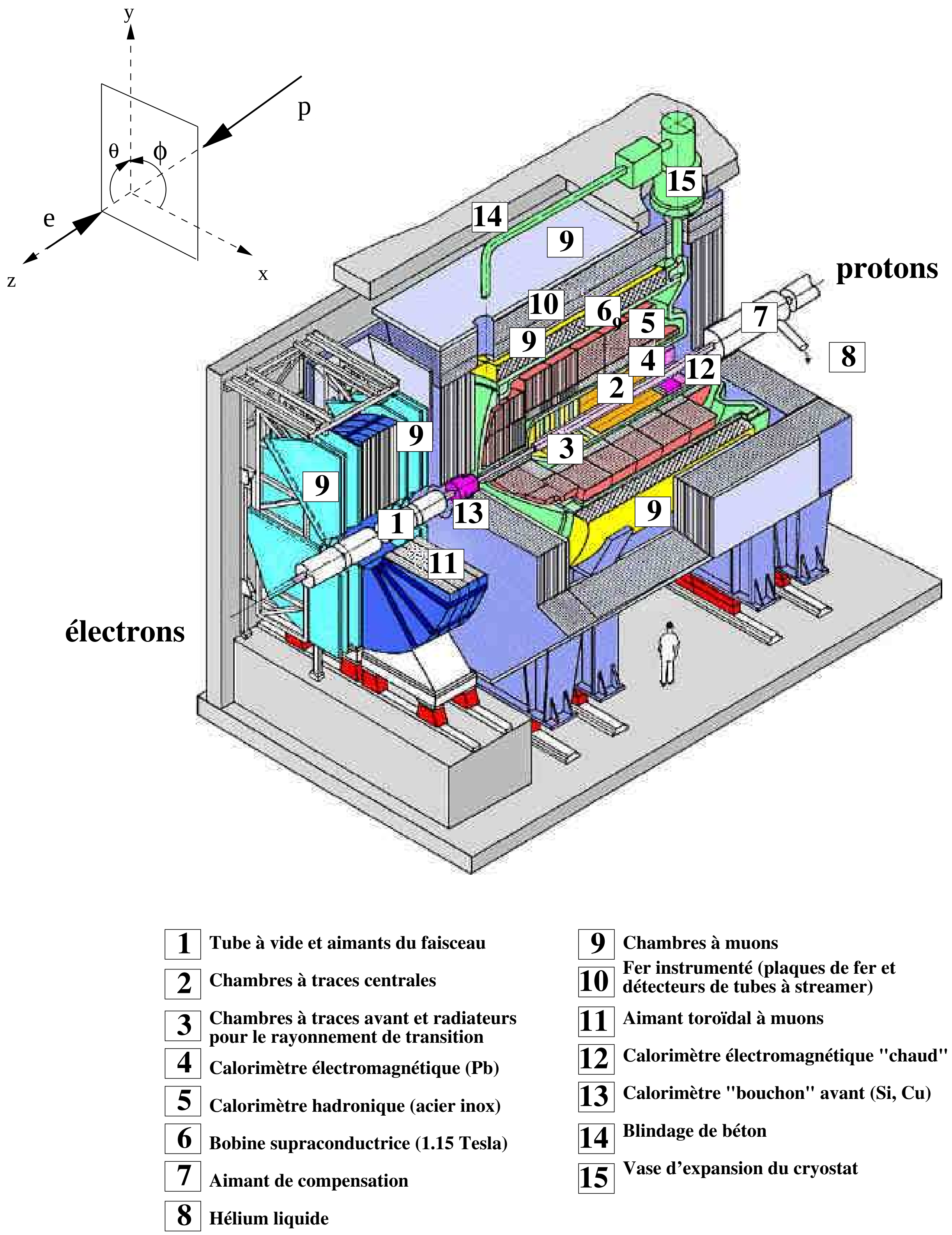


FIG. 2.2 – Vue du détecteur H1.

2.2.1 Les trajectographes

Le système de trajectographie (figure 2.3) a été conçu pour pouvoir reconstruire les trajectoires des particules, ce qui est spécialement important à l'avant où sont enregistrés la plupart des jets, donc un grand nombre de particules chargées. Il couvre la région $7^o < \theta < 176^o$ et permet de mesurer l'impulsion des particules avec une précision de $\sigma_p/p^2 \approx 5\text{x}10^{-3}$ GeV^{-1}. La résolution sur l'angle polaire est de l'ordre de 1 mrad.

2.2.1.1 Les trajectographes en silicium

Deux détecteurs séparés à bandes en silicium constituent le système : les détecteurs BST (**Backward Silicon Tracker**) et CST (**Central Silicon Tracker**) permettent de déterminer le passage d'une particule chargée grâce à l'ionisation dans le silicium. Le CST est utilisé pour déterminer les vertex secondaires et la présence de saveurs lourdes, et le BST pour valider le passage d'un électron diffusé dans la partie arrière du détecteur à petit angle. Pour la phase de haute luminosité un troisième détecteur appelé FST (**Forward Silicon Tracker**) a été installé. Des explications complètes peuvent être trouvées dans la référence [31].

2.2.1.2 Le détecteur central de traces chargées

Les chambres à dérive apportent la majeure partie de l'information trajectographique pour la partie centrale du détecteur. Les deux pièces les plus importantes sont les chambres à dérives CJC1 et CJC2 (**Central Jet Chamber**) emboitées l'une dans l'autre. Elles apportent une précision radiale sur le point de passage du faisceau de l'ordre de 170 μm, mais une précision longitudinale relativement faible de 22 mm. Afin d'améliorer la résolution longitudinale, deux autres chambres, CIZ (**Central Inner Z chamber**) et COZ (**Central Outer Z chamber**), plus petites, apportent une résolution de l'ordre de 300 μm en z. Ces quatre chambres sont complétées par des chambres proportionnelles, CIP (**Central Inner Proportional chamber**) et COP (**Central Outer Proportional chamber**), dont le temps de réponse plus court permet l'utilisation pour le premier niveau de déclenchement. En partant du point d'interaction, une particule rencontre successivement CIP, CIZ, CJC1, COZ, COP et à la fin CJC2.

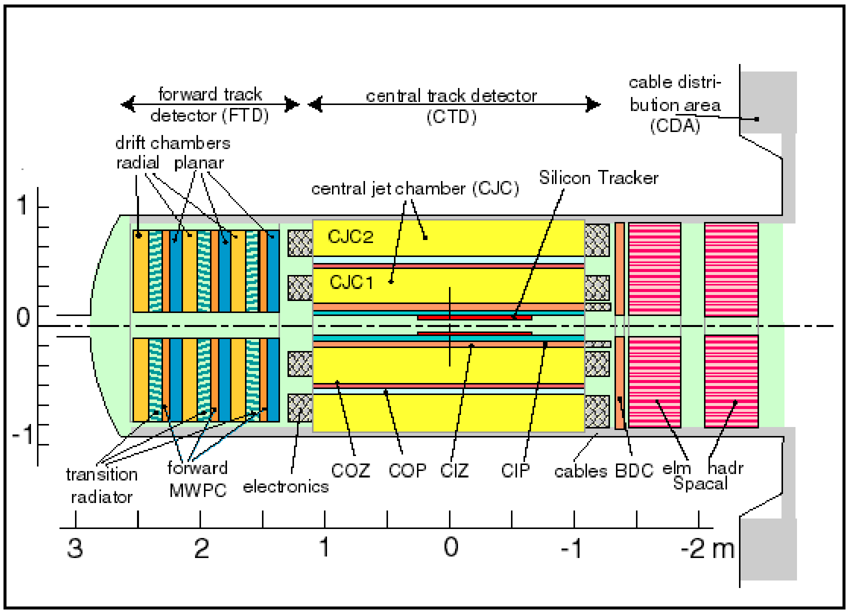


FIG. 2.3 – Coupe longitudinale du système de trajectographie.

2.2.1.3 Le détecteur avant de traces chargées

Situé à l'avant du détecteur central, il est composé de trois modules contenant chacun deux chambres à dérive (planaire et radiale), une chambre proportionnelle et un détecteur de rayonnement de transition.
Il couvre la région $5^o \leq \theta \leq 15^o$ et est utilisable par le système de déclenchement.

2.2.1.4 Le trajectographe arrière

La chambre à dérive BDC (**Backward Drift chamber**) plaçée à l'arrière permet de mesurer les trajectoires des électrons diffusés à petit angle.

2.2.2 Les calorimètres

Quatre calorimètres constituent le système calorimétrique de H1. Leurs caractéristiques sont détaillées dans [32, 33]. Le plus important est le LAr (**Liquid Argon calorimeter**) qui couvre la partie centrale et la partie avant du détecteur ($4^o < \theta < 153^o$). A l'avant, il est complété par le PLUG ($0{,}5^o < \theta < 3,5^o$) et à l'arrière par le SpaCal (**Spaghetti Calorimeter**) qui couvre l'angle polaire $153^o < \theta < 178^o$. Le fer instrumenté permet de mesurer les fins de gerbes hadroniques non contenues dans le LAr.

2.2.2.1 Le calorimètre LAr

Le choix d'un calorimètre à Argon Liquide a été fait en raison de sa bonne réponse électromagnétique, et de sa fine granularité permettant de séparer les électrons des pions et de mesurer l'énergie déposée avec une réponse homogène. Les atouts supplémentaires sont sa stabilité dans le temps et la facilité de calibration. L'ensemble du calorimètre (figure 2.4) se trouve dans un seul cryostat pour limiter les zones non instrumentées. A l'intérieur de ce cryostat, il est segmenté en huit roues : on différencie, en commençant par l'avant, successivement IF2/OF2, IF/OF1, FB2, FB1, CB3, CB2, CB1 et BBE. Ces roues sont segmentées transversalement en huits octants. La disposition des roues en octants est conçue pour que les zones non instrumentées qui les séparent ne pointent pas toutes vers le point d'interaction. Les interstices sont appelés "crack" en anglais. La partie hadronique du calorimètre est constituée d'un empilement de plaques d'acier séparées par des cellules de lecture. Cette partie sert de support à la partie électromagnétique du calorimètre, équipée de plaques en couches successives de plomb, de G10 (alliage d'époxy et de fibre de verre) et de plans de lecture. La résolution en énergie est de l'ordre de $\frac{\sigma_E}{E} \simeq \frac{12\%}{\sqrt{E}} \oplus 1\%$ pour la partie électromagnétique et $\frac{\sigma_E}{E} \simeq \frac{50\%}{\sqrt{E}} \oplus 2\%$ pour la partie hadronique.

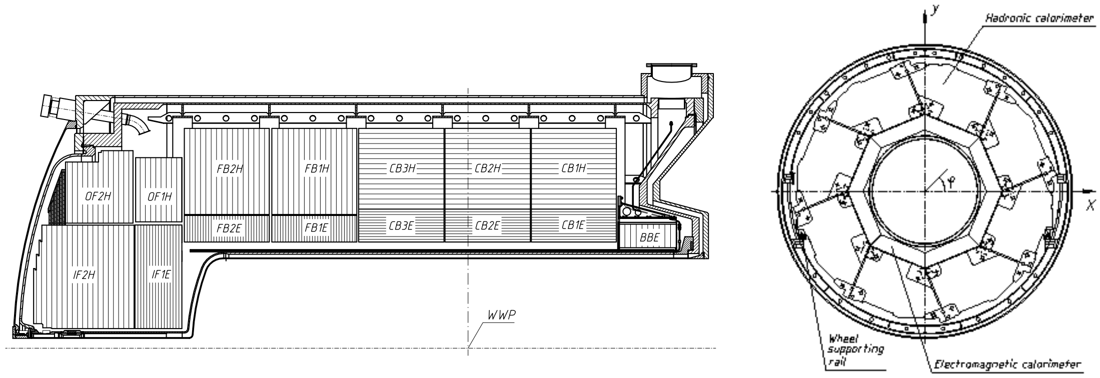


FIG. 2.4 – Coupes longitudinale (à gauche) et transverse (à droite) du calorimètre à Argon liquide.

Nous allons utiliser les dépôts d'énergie dans le LAr pour identifier les jets, les électrons et les photons. Les procédures d'identification sont détaillées dans la partie 4.

2.2.2.2 Le SpaCal

Le SpaCal (**Spa***ghetti* **Cal***orimeter*) est constitué de plomb et de fibres scintillantes. Il contient une partie électromagnétique et une partie hadronique. Sa résolution est $\frac{\sigma_E}{E} \simeq \frac{7{,}1\%}{\sqrt{E}}$ pour la partie électromagnétique et $\frac{\sigma_E}{E} \simeq \frac{56\%}{\sqrt{E}}$ pour la partie hadronique. Il sert principalement à mesurer l'énergie des électrons faiblement déviés vers la partie arrière du détecteur.

Le SpaCal a été installé en 1995 en remplacement d'un autre calorimètre (BEMC, **B***ackward* **E***lectro***M***agnetic* **C***alorimeter*) fait d'une alternance de plaques de plomb et de scintillateurs.

2.2.2.3 Le PLUG

Le PLUG permet de mesurer l'énergie des objets émis très en avant (principalement des jets). Formé de couches de plaques de cuivre avec une électronique de lecture en silicium entre les plaques, il a une résolution de $\frac{\sigma_E}{E} \simeq \frac{150\%}{\sqrt{E}}$.

2.2.3 Système à muons

Le système à muons (noté 9 sur la figure 2.2) est composé de chambres à muons, du fer instrumenté et d'un spectromètre à muons.
Au-delà du calorimètre se trouve un solénoïde supraconducteur (noté 6 sur la figure 2.2) produisant un champ magnétique de 1,15 Tesla et permettant donc la mesure de la charge et de l'impulsion des particules chargées avec les trajectographes. Le retour des lignes de champ est assuré par le fer instrumenté (noté 10 dans la figure 2.2). Son électronique de lecture permet de mesurer l'énergie des gerbes hadroniques qui ne sont plus contenues dans le LAr ainsi que la trajectoire des muons. Ce détecteur de structure octogonale, est constitué de 10 plaques d'acier de 7,5 cm équipées de 16 plans de "tubes à streamer". Dans la partie centrale, les tubes sont orientés selon l'axe des faisceaux pour permettre une bonne mesure de l'angle azimutal et de l'impulsion des muons. A l'avant et l'arrière, ils sont orientés perpendiculairement à l'axe z. Le fer instrumenté, dont un octant est représenté sur la figure 2.5, permet de mesurer des muons dont l'énergie est comprise entre 1,5 et 170 GeV.

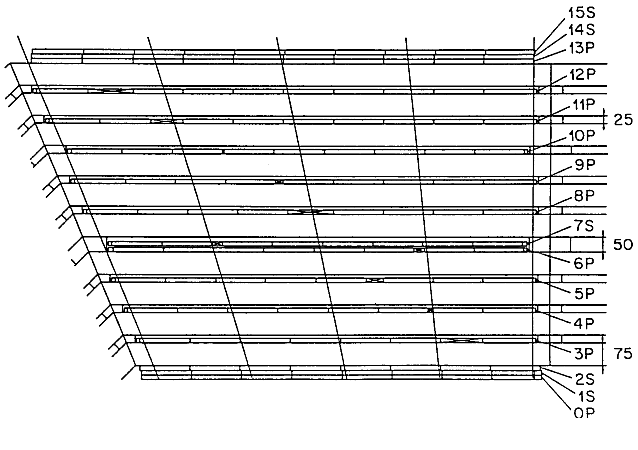


FIG. 2.5 – Vue schématique d'un demi-octant du fer instrumenté.

A l'avant, le système à muons central est completé par le spectromètre à muons avant (appelé **FMD** en anglais). Il couvre un angle polaire de $4^o < \theta < 16^o$. En raison des pertes d'énergie dans les détecteurs centraux et dans le toroïde, les muons doivent avoir une impulsion d'au moins 5 GeV pour être détectés. Au-delà de 100 GeV, la courbure n'est plus assez importante pour permettre une mesure précise de leur impulsion. Le spectromètre à muons est constitué d'un aimant toroïdal, fournissant un champ de 1,6 Tesla, entouré de 6 couches doubles de chambres à dérive (figure 2.6). Deux parmi ces six sont orientées dans le plan radial pour mesurer l'angle azimutal, et les quatre autres dans le plan longitudinal.

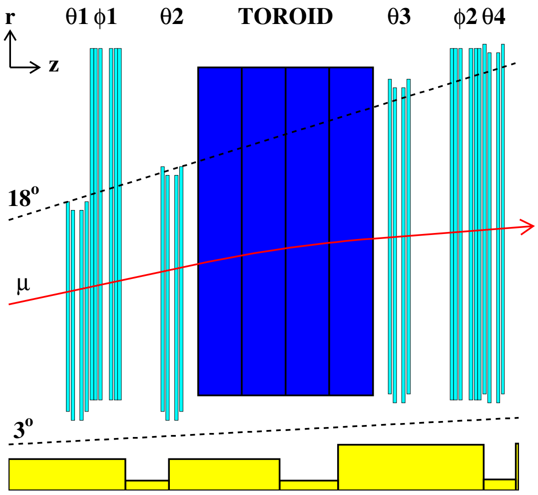


FIG. 2.6 – Vue schématique du spectromètre à muons avant. Les chambres à dérive sont nommées $\theta 1$-$\theta 4$ et $\phi 1$-$\phi 2$.

2.2.4 Système de luminosité

A HERA, la luminosité est déterminée à partir de la mesure du taux de diffusion Bethe-Heitler, processus de radiation de photons par les électrons incidents dans le champ électromagnétique des protons incidents. Ce processus étant purement électromagnétique, sa section efficace est connue avec une grande précision. Le luminomètre est situé en aval du point d'interaction dans la direction des positrons incidents. Dans les processus Bethe-Heitler, le positron perd une fraction de son énergie et modifie très peu sa trajectoire. Celle-ci est défléchie par les quadripôles situés en aval du détecteur (-23.8 m $< z <$ -5.8 m) permettant ainsi sa détection par un calorimètre à cristaux (appelés ET pour *Electron Tagger*) situé à 33,4 m du point d'interaction. Un autre calorimètre (appelé PD pour *Photon Detector*) situé à 102,9 m du point d'interaction, permet de mesurer le photon émis dans la direction du positron initial. La résolution en énergie des deux calorimètres ET et PD est σ (E)/E $= 25\%/\sqrt{E}$. Ces calorimètres seront utilisés dans la partie 5.2.4.1 pour déterminer l'efficacité de déclenchement de jets.

2.2.5 Acquisition de données et système de déclenchement

Le système de déclenchement est chargé d'une analyse rapide afin de sélectionner les événements venant des interactions électron-proton. Cette analyse qui va déclencher ou non l'acquisition de données est divisée en 4 niveaux différents.

- **Niveau L1 :** Le premier niveau, appelé L1, doit prendre une décision avec une fréquence de $\approx$ 10 MHz, correspondant au taux de croisement de faisceau. Ce niveau rassemble les décisions partielles des différents détecteurs qui sont gardées dans une mémoire pipeline pour la prise de décision qui a lieu 2,5 μs après le croisement de faisceaux : chaque sous-détecteur envoie des informations codées sur 8 bits, appelés éléments de déclenchement, au déclencheur central. Celui-ci combine les 200 éléments de déclenchement en 128 sous-déclencheurs. Chaque sous-déclencheur décide de la lecture ou du rejet d'un événement.

– **Niveau L2 :** Le deuxième niveau de déclenchement est basé sur une logique programmable qui doit prendre une décision en moins de 20 μs. Deux systèmes fonctionnent en parallèle : le "L2 topologique" et le "L2 réseau de neurones". Ils utilisent les mêmes informations que le L1, mais de façon plus raffinée, pour classifier les événements. Si le L2 décide que l'événement présente un intérêt, l'acquisition complète des données commence, mais peut être interrompue en cas de décision négative du L3.
– **Niveau L3 :** Le L3 dispose de 800 μs pour prendre sa décision de rejet de l'événement. Ce troisième niveau de déclenchement n'a pas été utilisé par H1 sur la phase I de HERA.
– **Niveau L4 :** Le L4 utilise une ferme de processeurs. Il est fondé sur une reconstruction rapide des données complètes de l'événement et permet une analyse plus approfondie.
  Les événements acceptés par le L4 sont écrits sur bande à un taux de quelques Hz.
– **Niveau L5 :** Afin de réduire davantage le lot, ces données sont ré-analysées dans un délai de quelques heures et classifiées selon des critères de physique. Cette classification est appelée L5. Elle repose sur une reconstruction complète de l'événement. Ces événements sont écrits sur bande avec tous les détails possibles, format POT (pour Production Output Tape), et sur disque en format réduit DST (pour Data Summary Tape).

2.2.5.1 Sous-déclencheurs utilisés dans l'analyse

L'analyse présentée dans cette thèse prend en compte les efficacités des sous-déclencheurs associés aux interactions à courant chargé [1, 2], à la production de paires de muons [12] et à la production de jets dans les processus de photoproduction [6]. L'efficacité des sous-déclencheurs pertinents pour les topologies contenant au moins un dépôt électromagnétique (électron ou photon) avec une impulsion transverse au-dessus de 20 GeV [1], comme par exemple les courants neutres, est proche de 100% [2] : une correction des simulations Monte Carlo pour tenir compte des efficacités de déclenchement n'est pas nécessaire dans ce cas.
Les sous-déclencheurs de niveau L1, utilisés pour la détection d'impulsion manquante dans le calorimètre LAr sont les suivants :
– 2 sous-déclencheurs (S66 et S77) basés sur la présence d'énergie transverse manquante dans le calorimètre.
– 1 sous-déclencheur (S71) basé sur la présence d'au moins une trace reconstruite en coïncidence en temps avec le croisement de faisceau.
– 2 sous-déclencheurs (S67 et S75) basé sur la présence d'énergie électromagnétique ou hadronique dépassant un certain seuil dans le calorimètre LAr.

Pour la détection de jets, les sous-déclencheurs utilisés sont le S67, le S75, le S77 décrits auparavant et le S64, basé sur la présence d'énergie transverse dans le calo-

[1] Dans l'analyse décrite dans la partie 5, nous recherchons des objets physiques d'impulsion transverse d'au moins 20 GeV.

rimétre LAr.

Pour la détection de muons, les sous-déclencheurs utilisés sont :
- 1 sous-déclencheur (S19) fondé sur un signal dans le tonneau central du fer instrumenté et une forte multiplicité de traces.
- 1 sous-déclencheur (S22) basé sur un signal dans la région avant ou arrière du fer instrumenté en combinaison avec une forte multiplicité de traces.
- 1 sous-déclencheur (S34) fondé sur un signal dans le tonneau central du fer instrumenté et une faible multiplicité de traces.
- 1 sous-déclencheur (S56) basé sur un signal dans le SpaCal en coïncidence avec un signal dans l'un des détecteurs à muons.

Dans tous les cas, les sous-déclencheurs ci-dessus sont mis en coïncidence avec des contraintes temporelles issues des chambres proportionnelles ou du calorimètre.
Dans l'analyse des canaux contenant des muons, un neutrino (impulsion transverse manquante) ou deux jets, un événement est retenu lorsque l'un des sous-déclencheurs appropriés décrits ci-dessus est activé. Les efficacités de déclenchement correspondantes seront étudiées au chapitre 5.

2.3 Simulations des processus physiques

Afin de comparer les données au prédictions théoriques, il est nécessaire de simuler les processus des collisions *ep* intervenant dans l'analyse. La simulation est effectuée en trois étapes distinctes :
- Génération des particules primaires à l'interaction
- Simulation des signaux déposés dans le détecteur
- Reconstruction des particules à partir des signaux des détecteurs

2.3.1 Génération des processus

La génération elle-même se divise en trois étapes principales :
- **Calcul des éléments de matrice** : Pour ces calculs, on tient compte des densités partoniques du photon et du proton pour déterminer les états initiaux des quarks et gluons intervenant dans certains processus. Pour le processus de production de paire de leptons nous avons utilisé GRAPE [34], pour le processus de photoproduction PYTHIA [35], pour le processus de diffusion profondément inélastique LEPTO [36], pour le processus de production et désintégration de bosons $W^{\pm}$ EPVEC [37] et pour le processus de QED Compton WABGEN [38].

- **Calcul des corrections radiatives :** Pour tenir compte des corrections radiatives électromagnétiques, on utilise le module HERACLES [39]. Il est intégré dans tous les générateurs utilisés dans l'analyse sauf PYTHIA, qui calcule lui-même ces corrections, et WABGEN, dans lequel les effets radiatifs sont négligés. Il existe principalement deux modèles différents, décrits dans la thèse

[40], permettant de tenir compte des corrections radiatives QCD : on utilise d'une part le modèle dit de "cascades partoniques" (abrégé MEPS pour "Matrix Element + Parton shower"), et d'autre part le modèle dit de dipôles de couleur (abrégé CDM pour "Color dipole model"). LEPTO tient déjà compte des corrections radiatives QCD dans le modèle MEPS, mais peut être relié au module ARIADNE qui les traite dans le modèle CDM. Ainsi, les modules complets de DJANGO [41] et RAPGAP [42] (regroupant les modules LEPTO, HERACLES et ARIADNE), ont été utilisés pour simuler des processus de diffusion profondément inélastique dans les deux approches.

– **Fragmentation :** Le modèle de cordes de Lund permet de décrire la fragmentation de l'état hadronique, c'est-à-dire l'hadronisation de paires $q - \overline{q}$. Il est incorporé dans JETSET [43]. Ce module est utilisé dans les générateurs précédents pour décrire l'état hadronique final.

Le tableau 2.1 résume les caractéristiques des générateurs utilisés pour les différents processus physiques étudiés.

2.3.2 Simulation et reconstruction des événements

Les particules stables résultantes sont alors introduites dans la simulation du détecteur H1 qui est basée sur le programme GEANT. La simulation inclut d'une part la géométrie du détecteur, et d'autre part la description des matériaux. Pour pouvoir comparer les prédictions théoriques aux données, on simule les processus en tenant compte des conditions de run correspondant aux données, comme par exemple les sections mortes de détecteur ou le bruit d'empilement.

Processus physique	Générateur	Corrections radiatives QED	QCD	Fragmentation
DIS NC	DJANGO 6.2	HERACLES	MEPS	JETSET
DIS CC	DJANGO 6.2	HERACLES	MEPS	JETSET
DIS NC	RAPGAP 2.08	HERACLES	CDM	JETSET
γp	PYTHIA 5.7, 6.1	PYTHIA	MEPS	JETSET
Paires e^+e^-	GRAPE	HERACLES	MEPS	JETSET
Paires $\mu^+\mu^-$	GRAPE	HERACLES	MEPS	JETSET
QED Compton	WABGEN	-	-	JETSET
$e^{\pm}p \to e^{\pm}XW^{\pm}$	EPVEC	HERACLES	MEPS	JETSET

TAB. 2.1 – Caractéristiques des générateurs utilisés dans l'analyse. MEPS est un modèle de cascades partoniques et CDM le modèle de dipôles de couleur.

Finalement, les événements simulés sont reconstruits de la même manière que les données.

Chapitre 3

Outils d'analyse orientés objet

Hoc opus, hic labor est.

Virgile

En 2001, l'arrêt de la machine pour la préparation de la nouvelle phase haute luminosité de HERA a permis l'installation de plusieurs nouveaux détecteurs au sein de l'expérience et d'initier le passage des programmes d'analyse vers un environnement orienté objet. Les buts étaient de réorganiser les outils d'analyse de physique et de les intégrer dans un environnement unique.

Jusqu'à la fin de l'année 2000, tous les outils d'analyse de physique étaient développés en fortran. Afin de faciliter les analyses et de regrouper les différents outils, un projet a été mis en place. Le nouveau logiciel issu de ce projet devait :

- constituer un environnement unique et extensible,
- permettre l'analyse de tous les processus physiques de H1,
- standardiser les algorithmes de physique, les critères de sélection et l'identification de particules,
- et présenter un accès plus rapide et simple aux données.

Afin de répondre à ces différents critères, de faciliter les échanges d'information entre les différents groupes de physique et l'usage pour les non experts, l'utilisation d'un langage orienté objet C++ dans l'environnement ROOT [44, 45, 46] a été retenue. L'un des principaux avantages de la conception orientée objet consiste en la capacité à définir des objets ayant des propriétés proches des objets physiques, comme par exemple la définition de particules par leur quadri-vecteur. Durant la première partie de ma thèse, j'ai pris une part importante de responsabilités au sein du projet en contribuant aux développements de la structure de base orientée objet et de plusieurs algorithmes C++ d'identification de particules, décrits dans le chapitre 4.

Ce chapitre présente une description générale du projet et de ses résultats avec une attention particulière à mes propres contributions. Après une brève introduction concernant la structure générale des données, constituée de quatre couches parallèles, je détaillerai le contenu des objets calorimétriques, la structure de l'algorithme de reconstruction hadronique et les définitions et relations des particules reconstruites. Le contenu est détaillé dans le document interne [47].

3.1 Structure générale des données

Un gros effort a été effectué pour développer les structures C++ des données car l'organisation et la hiérarchie des informations stockées a un impact primordial sur les performances finales et la facilité d'utilisation.
La structure développée est basée sur la reconstruction des différents détecteurs, écrite en fortran et qui n'a pas été modifiée (partie gauche de la figure 3.1). Les données reconstruites sont stockées sous forme de banques dans les fichiers appelés POT ($Production\ Output\ Tape$) à partir desquels les informations DST ($Data - Summary\ Tape$) suffisantes pour les analyses de physique sont extraites. Le format de banques utilisé est BOS [48], et elles sont accédées par FPACK [49]. L'environnement d'analyse développé, indiqué en rouge sur la figure 3.1, repose sur l'information contenue dans ces banques. Nous avons développé une structure permettant de convertir l'information de DST en information orientée objet. Le choix du format de données s'est dirigé vers $ROOT$, déjà utilisé par d'autres expériences de physique à haute énergie (astroparticule, physique des accélérateurs). Le nouveau code est

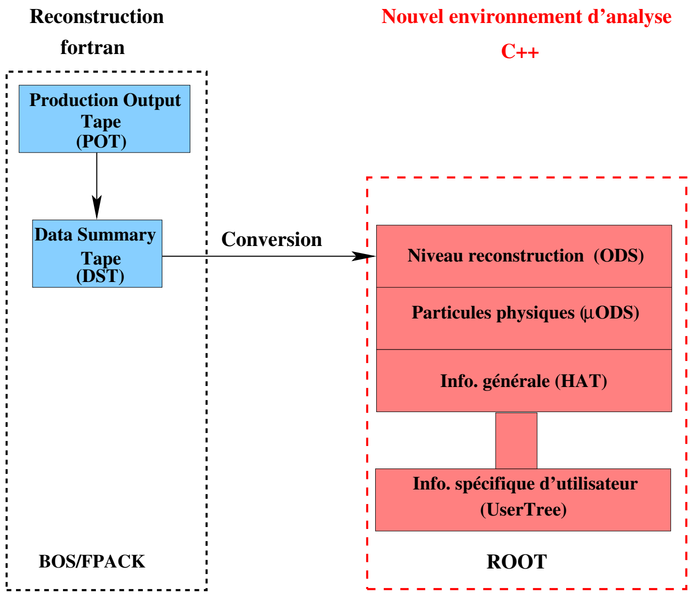


FIG. 3.1 – Structure des données : μODS, HAT et $UserTree$ constituent les principales couches du modèle de données.

organisé sous forme de modules dont les différentes versions sont gérées sous CVS

[50].

Afin de déconnecter le plus possible les informations du détecteur des objets utiles aux analyses de physique, les données sont stockées dans une structure composée de trois couches prédéfinies plus une couche qui peut être remplie par l'utilisateur (partie droite de la figure 3.1) :

- La couche la plus fondamentale, *ODS* (pour *Object Data Store*), est produite à partir des *DST* ou *POT*. Elle contient les données reconstruites des sous-détecteurs. Les fichiers *ODS* correspondent un à un aux informations contenues dans les banques *BOS* de *DST*. Cette exacte correspondance est maintenue afin de permettre une reconversion des fichiers *ODS* en *DST*. Pour éviter la duplication de l'information, la possibilité de transcription virtuelle (sans écriture) événement par événement des *ODS* a été codée. La taille d'un événement contenu dans les *ODS* est de 13 kO environ.
- La seconde couche dite μODS (pour *Micro Object Data Store*), contient les listes de particules reconstruites et certaines informations propres aux particules. Elle n'occupe que 3 kO par événement et est remplie en utilisant les informations *ODS*.
- La derniere couche appelée *HAT* (pour $H1$ *Analysis Tag*) contient les informations générales de l'événement, comme par exemple l'état du système de déclenchement et les variables cinématiques globales.
  Deux cent variables différentes sont actuellement stockées, ce qui correspond à 0,4 kO par événement.
- Une couche supplémentaire appelée $UserTree$ peut être produite en parallèle pour contenir des informations propres à l'utilisateur.

Pour la majorité des analyses de physique, les couches μODS, *HAT* et UserTree de taille réduite sont suffisantes. Elles sont intégrées dans une structure en arborescence appelée *skeleton*. L'accès et l'écriture à l'intérieur du *skeleton* sont gérés par la classe $H1Tree$. Celle-ci représente l'interface entre l'utilisateur et les données : pour ouvrir et fermer les fichiers de données, seules les fonctionnalités du $H1Tree$ sont à connaître. Pour faciliter la navigation entre les différentes couches, des pointeurs appelés $H1Pointer$ ont été implémentés. Grâce à une nomenclature standardisée, ils permettent par exemple de trouver les caractéristiques de la trace associée (stockée sur *ODS*) à une certaine particule reconstruite (stockée sur μODS).

La production de fichiers se fait dans l'ordre suivant :

- Les fichiers *ODS* sont une transcription de leur analogue, les *DST*, mais ne sont pas stockés sur disque (*ODS* virtuel).
- La production de μODS et *HAT* aura lieu plus souvent afin d'intégrer les améliorations des algorithmes de physique et des calibrations. Elle sera effectuée à partir des *ODS* virtuels et stockée de manière permanente.

En résumé, les données sont stockées sous forme de fichiers *DST* contenant les informations reconstruites des détecteurs, de fichiers μODS contenant les informations particulaires et des fichiers *HAT* contenant des informations plus générales de l'événement. La structure *ODS* est produite “on the fly” à partir des *DST* pendant le traitement d'un événement.

3.2 Informations reconstruites du détecteur (*ODS*)

La couche appelée *ODS* est une transcription une à une en structure C++ des informations contenues sur *DST*, c'est-à-dire des informations reconstruites venant des différents détecteurs. C'est pour cela que la plupart des fonctions d'accès ont gardé la même forme et nomenclature.
Seules les informations correspondant aux dépôts calorimétriques (cluster) et les informations des trajectographes (trace) ont été réorganisées. Il s'agit des objets fondamentaux utilisés ensuite par les identificateurs de particules. Nous avons donc transcrit l'information correspondante dans des classes exploitant les avantages de l'environnement orienté objet, comme par exemple l'héritage. Nous distinguons ainsi trois types d'objets sur *ODS* : les dépôts calorimétriques (appelés *H1Cluster*), l'information des trajectographes (appelée *H1Track*) et les informations des autres banques de données (appelée *H1Bank*). Les informations stockées sur *ODS* ne sont pas accédées directement à travers ces classes, mais à travers une classe d'interface appelée *H1OdsEvent*. Celle-ci gère l'accès aux informations requises en passant par les classes *H1TrackEvent*, *H1ClusterEvent* et *H1BankEvent* correspondant aux traces, clusters et aux informations restantes stockées sous forme de banques. J'ai plus spécialement contribué à l'organisation, l'implémentation et la validation des clusters calorimétriques décrits dans la partie 3.2.2.

3.2.1 Les traces

L'information contenue initialement dans plus de dix banques *DST* a été réorganisée et centralisée dans le module *H1Tracks*. Celui-ci contient une classe principale, appelée *H1Track*, définissant les fonctions principales communes à tous types de traces. Dans les classes qui en héritent les calculs et les variables importantes, comme par exemple le calcul du quadri-vecteur impulsion-énergie, ont été implémentés. Nous distinguons trois types de traces dérivant de *H1Track* :

- Les traces non ajustées au vertex central (*H1NonVertexFittedTrack*).
- Les traces ajustées au vertex central (*H1VertexFittedTrack*).
- Les traces de muons cosmiques (*H1CosmicTrack*).

Au moment de la conversion de l'information contenue sur *DST* en structures C++ sur *ODS*, ces classes sont remplies, comme nous l'avons déjà mentionné, à travers la classe *H1TrackEvent* qui contient entre autres ces trois listes de traces.

3.2.2 Clusters et cellules calorimétriques

Pour notre analyse décrite dans la partie 5, l'implémentation et l'intégration de l'information calorimétrique dans le code est importante car à grande impulsion transverse, domaine cinématique de notre analyse, la résolution calorimétrique est

meilleure que celle des trajectographes.
Le calorimètre de H1 n'étant pas intrinsèquement compensateur, des méthodes de pondération des énergies mesurées sont appliquées au niveau de la reconstruction, décrite dans la thèse [51]. Les cellules d'énergie supérieure à un certain seuil sont regroupées en clusters (conglomérats ou amas de cellules). Ces informations sont stockées sous forme de banques (de nom $DCLU$ pour les clusters et $DCEL$ pour les cellules) sur DST.

Le but était alors de transcrire ces informations dans l'environnement orienté objet avec les contraintes suivantes :
- Minimisation de la taille occupée
- Accès rapide
- Intégration de toutes les informations contenues sur DST
- Définition d'objets fondée sur les conventions de $ROOT$

3.2.2.1 Structure

Afin d'utiliser les informations disponibles, des classes $H1Cluster$ et $H1Cell$ ont été implémentées, proches des caractéristiques des banques $DCLU$ et $DCEL$. Les deux classes dérivent, au sens orienté objet, de l'objet *TObject* pour permettre le stockage dans les fichiers $ROOT$.
Les cellules contiennent deux types d'informations principales :
- Une information déterminée événement par événement, comme par exemple l'énergie au niveau électromagnétique
- Une information déterminée année par année, comme par exemple la position de la cellule du calorimètre
  (information lue à travers la base de données H1 appelée NDB)

Le premier type d'information est stocké de façon permanente pour chaque événement, alors que le second est relu lors de la lecture des informations de cellule ou de cluster.

De la même manière que sur DST, les objets $H1Cell$ sont créés à travers les clusters $H1Cluster$ et ne peuvent donc exister sans cluster associé (voir figure 3.2). La classe $H1Cluster$ permet l'accès aux quantités d'un cluster calorimétrique. $H1Cluster$ contenant essentiellement la liste des cellules associées, toutes les propriétés d'un cluster sont recalculées à partir des cellules qui le composent. Afin d'identifier les clusters associés à une particule, pour chaque cluster, un lien vers la particule candidate correspondante a été développé. Le concept de particule candidate est détaillé dans le paragraphe 3.3. Afin d'en faciliter l'utilisation, plusieurs mécanismes pour ordonner les cellules et clusters ont été implémentés.

L'information de géométrie, c'est-à-dire la position de la cellule dans le calorimètre est stockée dans l'objet global $H1CaloGeometry$ du module de géométrie appelé $H1Geom$. Cette information est remplie pendant la production standard des fichiers *ODS* et provient de la base de données NDB. La géométrie est accédée lorsqu'une fonction de géométrie des cellules est appelée. Toutes les informations, comme par exemple le numéro de roue de calorimètre correspondant à la cellule sont

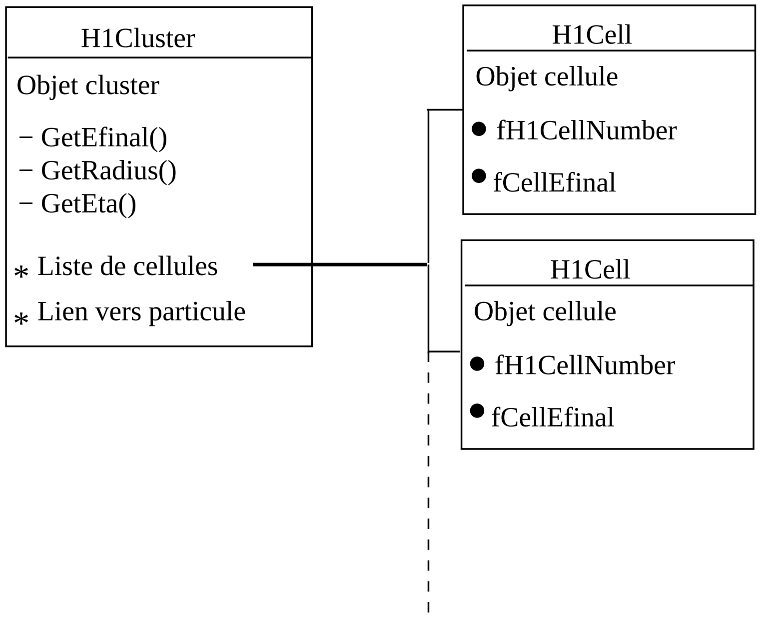


FIG. 3.2 – Structure des clusters ($H1Cluster$) et cellules ($H1Cell$) calorimétriques : Une liste de cellule et un lien vers la particule associée sont intégrés dans chaque cluster.

disponibles à travers ce mécanisme.

3.2.2.2 Reconstruction

La classe $H1ClusterEvent$ centralise, comme cela est visible sur la figure 3.3, l'écriture et la lecture des objets calorimétriques. Elle contient d'une part une liste de tous les clusters calorimétriques, et d'autre part une liste de toutes les cellules calorimétriques. La distinction est nécessaire pour permettre l'écriture des objets dans les fichiers $ROOT$.
La routine de remplissage des clusters, et donc des cellules s'appelle $H1ClusterFiller$. Son rôle est de construire les objets clusters $H1Cluster$ et cellules $H1Cell$ à partir des informations stockées dans les banques DST. Celles-ci sont lues à travers des commandes en format BOS (correspondant au format des banques DST). La procédure (figure 3.3) de remplissage peut être divisée en deux étapes :

- Lecture de chaque cluster calorimétrique sur DST :
  - Pour chaque cluster trouvé, création d'un objet $H1Cluster$.
  - Pour chaque cellule associée, création d'un objet $H1Cell$.
- Pour chaque objet créé, transcription de toutes les quantités et association des objets.

La routine $H1ClusterFiller$ est appelée au même niveau que les autres routines de remplissage $H1BankFiller$ et $H1TrackFiller$, par la routine principale de remplissage des ODS, appelée $H1OdsFiller$.

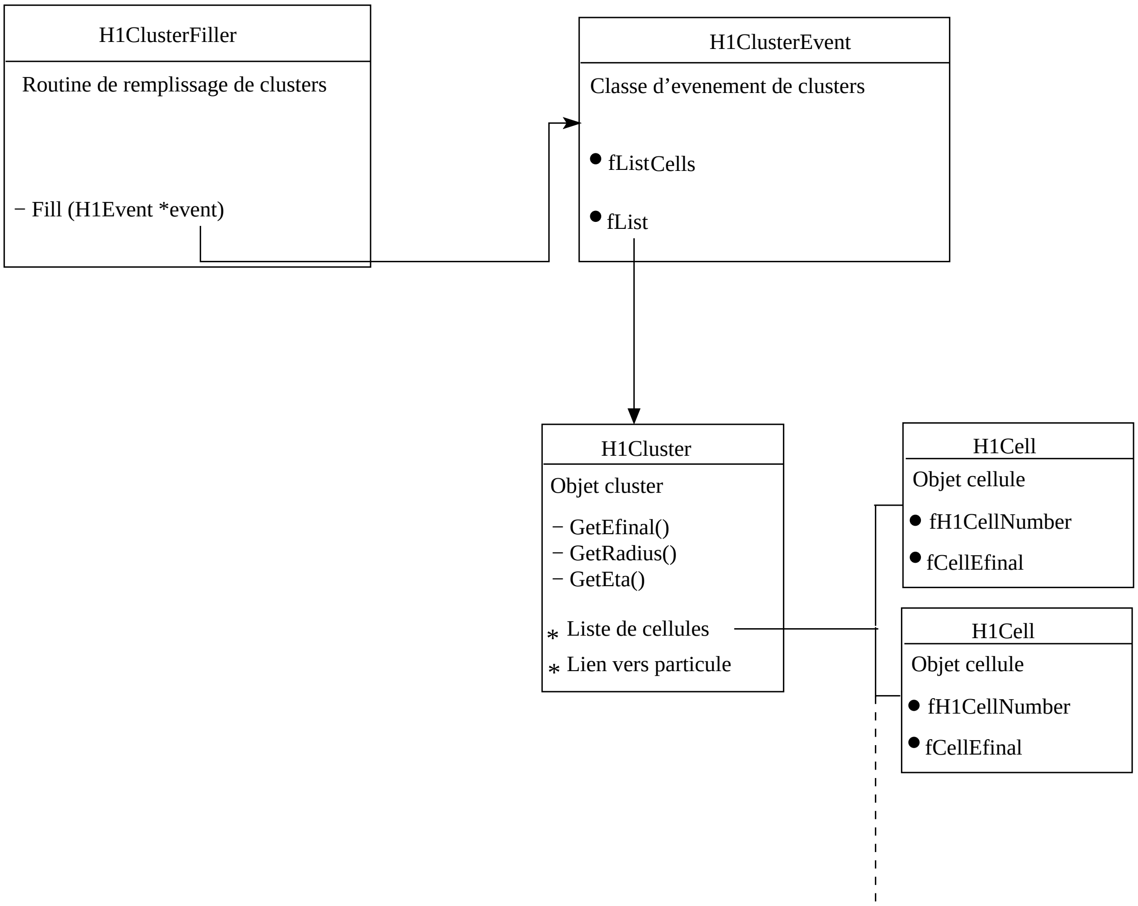


FIG. 3.3 – Structure décrivant la procédure de reconstruction des objets clusters $H1Cluster$ et cellules $H1Cell$: $H1ClusterFiller$ remplit les listes de clusters $H1Cluster$ et de cellules $H1Cell$ contenues dans $H1ClusterEvent$.

3.2.3 Banques

En plus des traces et des clusters calorimétriques décrits ci-dessus, les fichiers *ODS* contiennent l'équivalent des banques *DST* non associées aux deux catégories ci-dessus, sous forme d'objets C++. Ces objets contiennent aussi bien des informations de détecteurs comme par exemple l'état du système de déclenchement que des informations plus générales de l'événement comme par exemple le vertex central. Son contenu est décrit plus en détails dans [47].

3.3 Information particulaire (μODS)

Les fichiers μODS sont construits à partir des structures *ODS* et contiennent toutes les données des particules reconstruites d'un événement. Les spécifications définies lors de la conception des structures étaient les suivantes :

– Accès rapide et facile aux informations particulaires, comme par exemple le

quadri-vecteur (au niveau de l'analyse).
- Définition d'une liste de particules sans double comptage en énergie entre les différents sous-détecteurs, contenant toute l'expertise en reconstruction de H1.
- Définition de listes de particules spécifiques à chaque type de particules, comme par exemple une liste de muons.

- Intégration aisée de nouvelles listes de particules, comme par exemple une liste de taus.

A partir de ces contraintes, nous avons développé la structure décrite dans la suite. Nous avons défini d'une part, les informations nécessaires pour décrire l'objet physique "particule", et d'autre part, les différentes listes de particules nécessaires pour permettre la plus grande variété d'analyses de physique possible. Ainsi, quatre types de listes de particules ont été implémentés au niveau μODS :
- La liste des particules candidates
- La liste des particules identifiées
- La liste des particules composées
- La liste des particules Monte-Carlo, c'est-à-dire générées

Ces quatre types se divisent en plusieurs listes de particules reconstruites (cf. tableau 3.1). Par la suite, nous décrirons les différents types de listes de particules, qui seront également utilisées dans l'analyse.

Nom des Listes	*Type de particules*	*Contenu*
Particles	H1PartCand	Particules candidates
Muons	H1PartMuon	Muons
EmParticles	H1PartEm	Particules électromagnétiques
HFS	H1PartHFS	Hadrons
SelectedTracks	H1PartSelTracks	Particules chargées
FNC particles	H1PartFNC	Neutrons
Jets	H1PartJets	Jets reconstruits avec Kt
Dstar	H1PartDstar	D^* identifié selon $D^* \to K\pi\pi$
DiffractiveJPSI	H1PartJPsiDiff	J/Psi
MCParticles	H1PartMC	Particules générées (MC)

TAB. 3.1 – *Listes de particules disponibles sur μODS.*

Nous définissons les listes suivantes :

- **Liste des particules identifiées :**

  Chaque identificateur remplit, lorsqu'il est appelé, sa propre liste de particules dites "simples", que l'on appelle de façon générique liste de particules identifiées. Ainsi, pour l'instant, les quatre identificateurs implémentés créent quatre listes de particules identifiées distinctes qui peuvent présenter du double-comptage en énergie entre-elles. Donc plusieurs hypothèses de particules identifiées peuvent correspondre à la même particule physique :
  - Les particules dites, par tradition, traces "Lee West" [52] ($H1PartSelTrack$) correspondant à une sélection de particules chargées identifiées à partir de traces de bonne qualité
  - Les particules électromagnétiques ($H1PartEm$) qui correspondent à la reconstruction d'un dépôt électromagnétique dans les calorimètres
  - Les muons ($H1PartMuon$)
  - Les hadrons ($H1PartHFS$)
- **Liste des particules candidates :**

  Elle correspond à la couverture d'un angle solide de 4Π autour du vertex de toutes les particules reconstruites sans double comptage en énergie. Lorsque plusieurs particules identifiées correspondent à la même particule physique, seule la meilleur hypothèse est retenue. Les caractéristiques des particules candidates sont :
  - Un quadri-vecteur énergie-impulsion de la particule
  - Un bit par type de particule identifié correspondante
  - Un lien en forme de pointeur vers la liste de clusters ou traces associées sur *ODS*
  - Un lien vers les particules identifiées correspondantes
- **Liste des particules composées :**

  Après la reconstruction des particules dites "simples" (particules candidates et identifiées), la reconstruction des particules dites composées est lançée. Ainsi, pour l'instant, les particules suivantes sont reconstruites :
  - Jet : reconstruit à partir des particules hadroniques
  - D^* : reconstruit à partir des particules chargées $H1PartSelTrack$
  - J/Ψ : reconstruit à partir des particules chargées $H1PartSelTrack$
- **Liste des particules générées :**

  Cette liste n'est remplie que pour les simulations Monte-Carlo des processus. Elle contient toutes les caractéristiques des particules générées, comme par exemple leur quadri-vecteur énergie-impulsion et leur type.

Par la suite, nous détaillons les principes de reconstruction des différents identificateurs.

3.3.1 Principes des identificateurs de particules

Les identificateurs de particules remplissent à tour de rôle les différentes listes de particules. Afin d'optimiser la reconstruction d'un événement les principes suivants

doivent être respectés :
- L'identificateur doit verrouiller les traces et clusters calorimétriques utilisés pour éviter le double comptage.
- Les performances des algorithmes en terme de reconstruction d'objets physiques doivent être au moins aussi bonnes que les identificateurs écrits en fortran.

A partir de ces critères, nous avons établi un ordre dans lequel les différents identificateurs sont appelés. Ceci permet de verrouiller en premier les traces et amas calorimétriques des particules les plus faciles à identifier. Ainsi, l'identificateur de particules chargées utilisant les traces uniquement est appelé en premier. Puis, l'identificateur des particules électromagnétiques, c'est-à-dire des électrons et photons, est appelé à son tour car leur signature dans le calorimètre permet facilement de les identifier. Ensuite, un algorithme de reconstruction de muons est lançé. Le lepton tau étant très difficile à reconstruire, un algorithme d'identification est en cours de développement. A la fin, toute l'énergie restante est inclue dans ce qu'on appelle l'état hadronique final.

Chaque identificateur remplit deux listes : la liste des particules candidates ne contenant pas de double comptage en énergie et la liste des particules identifiées correspondant à son propre type de particules. Au fur et à mesure de l'association des clusters et des traces aux particules candidates, ces informations sont verrouillées et ne sont donc plus visibles pour les identificateurs suivants. Néanmoins une distinction doit être faite : seuls les clusters et traces associés aux particules candidates dites "isolées" sont ignorés par les identificateurs suivants. Le critère d'isolation est propre à chaque type de particule et a été optimisé en fonction des caractéristiques du détecteur et des analyses précédentes. Les principes de reconstruction, ainsi que les critères d'isolation sont détaillés dans les parties correspondant aux identificateurs (chapitre 4).

Dans le cadre de cette thèse, un algorithme de reconstruction hadronique et un algorithme de reconstruction de jets en K_T ont été développés. Leur structure est décrite dans la suite, et leur contenu physique sera donné dans le chapitre suivant.

3.3.2 Structure de l'identificateur hadronique

Le module a été divisé en trois sous-modules différents pour répondre aux besoins de l'algorithme. Les informations d'entrée de l'algorithme sont d'une part les clusters calorimétriques et d'autre part les traces associées aux particules chargées. Ces dernières ont été déjà reconstruites lorsque l'identificateur hadronique est lançé. Contrairement aux particules chargées, les clusters calorimétriques stockés sur ODS ne répondent pas aux critères de qualité requis et nous devons corriger leurs propriétés. Nous avons donc développé trois classes distinctes :
- Une classe appelée $H1HfsCluster$ redéfinissant les caractéristiques d'un amas calorimétrique.
- Une seconde classe appelée $H1HfsClusterCalib$ recalculant les caractéristiques des clusters, corrigées des effets du détecteur comme par exemple la calibration en énergie et l'alignement.

Variable	Contenu
fBarycenterAlign	Le barycentre après toutes les corrections
f4VectorCalib[4]	Le quadri-vecteur après toutes les corrections
fKilledEFraction	"1 - Fraction d'énergie enlevée" (1,0 = rien enlevé)
fEMFraction ("transient")	La fraction électromagnétique du cluster
fLArFraction ("transient")	La fraction du cluster se trouvant dans le LAr
fODSCluster	Lien vers le cluster *ODS* initial
fAssoc	Flag interne pour signaler l'association à une particule

TAB. 3.2 – Information disponible sur les clusters calorimétriques de type "H1HfsCluster"

- Une troisième classe appelée $H1HfsCreatePart$ qui représente le cœur de l'algorithme et fournit, à partir des traces des particules chargées et clusters, et à travers l'algorithme décrit dans le section 4.4.3, les hadrons.

Les clusters de bonne qualité ("H1HfsCluster")

Nous avons doté les clusters $H1HfsCluster$ de propriétés supplémentaires par rapport aux clusters $H1Cluster$ stockés sur ODS. Pour chaque cluster $H1Cluster$, nous définissons un nouveau cluster $H1HfsCluster$ avec des propriétés recalculées par $H1HfsClusterCalib$. Les propriétés sont résumées dans le tableau 3.2 :

- **Fraction électromagnétique** $fEMFraction$ **:**
  Il s'agit de la fraction d'énergie du cluster déposée dans la section électromagnétique du calorimètre. Elle est typiquement supérieure à 0,9 pour un électron.
- **Fraction LAr** $fLArFraction$ **:**
  La fraction LAr correspond à la fraction d'énergie contenue dans le calorimètre LAr.
- **Barycentre** $fBarycenterAlign$ **:**
  Le vecteur de position après les corrections d'alignement et d'inclinaison des faisceaux.
- **Quadri-vecteur** $f4VectorCalib$ **:**
  Le quadri-vecteur énergie-impulsion du cluster après toutes les corrections d'alignement et de calibration.
- **Pointeur** $fODSCluster$ **:**
  Il s'agit du pointeur vers le cluster d'origine $H1Cluster$.
- **Variables internes** $fAssoc$ **et** $fKilledEFraction$ **:**
  $fAssoc$ permet à l'algorithme de savoir si un cluster a été déjà associé à une particule. $fKilledFraction$ est une variable interne permettant de savoir avec quelle fraction d'énergie initiale le cluster doit être considéré.

Nous allons décrire dans la suite quelles sont les corrections appliquées aux clusters dans $H1HfsClusterCalib$.

Application des corrections ("H1HfsClusterCalib")

Les effets expérimentaux, dont $H1HfsClusterCalib$ tient compte dans un premier temps pour corriger le quadri-vecteur énergie-impulsion des clusters, sont les suivants :

- **Calibration en énergie :**
  Nous appliquons à tous les clusters des facteurs correctifs pour obtenir des hadrons bien mesurés en énergie. Les facteurs ont été obtenus dans l'analyse d'événements à courant neutre [2].
- **Alignement :**
  D'une période de prises de données à l'autre, la position des trajectographes par rapport au calorimètre LAr peut varier. Pour en tenir compte, des constantes correctives dépendant de la période de prise de données sont appliquées.
- **Inclinaison des faisceaux :**
  Les faisceaux se croisent au point d'interaction avec un petit angle par rapport au détecteur H1. Nous en tenons compte en appliquant des facteurs correctifs dépendant de la période de prise de données sur le vecteur de position.
- **Position du point d'interaction :**
  L'algorithme de reconstruction hadronique associe de l'énergie calorimétrique à l'énergie de traces. Pour ce faire, il faut tenir compte de la position du vertex primaire.

Dans un second temps, nous appliquons des critères de qualité et de réjection du bruit analogique possible :

- **Critères de qualité généraux :**
  Nous gardons uniquement, pour les hadrons ayant déposé une partie de leur énergie dans le fer instrumenté ou le PLUG, l'énergie déposée dans les calorimètres LAr ou SpaCal. Nous demandons ensuite que l'énergie du cluster $E_{cluster}^{finale}$ soit supérieure à 0. Enfin, les cellules d'énergie négative seront rejetées.
- **Suppression de bruit analogique :**
  Du bruit analogique peut simuler un dépôt d'énergie dans certaines cellules du calorimètre. Les clusters dans le LAr doivent remplir les critères suivants pour supprimer les cellules bruyantes isolées :
  - Region avant ($\theta_{cluster} < 15^o$) : $E_{cluster} > 0.4$ GeV OU $E_{cluster}^{sphere} > 0.4$ GeV
  - Region centrale ($\theta_{cluster} \geq 15^o$) : $E_{cluster} > 0.8$ GeV OU $E_{cluster}^{sphere} > 0.8$ GeV

  où $E_{cluster}$ est l'énergie du cluster et $E_{cluster}^{sphere}$ l'énergie contenue dans une sphère autour du cluster d'un rayon de 20 cm dans la région centrale et de 40 cm à l'avant. Ces coupures de suppression de bruit, élaborée dans [53], rejettent les dépôts de faible énergie concentrés en une seule cellule. Dans un dernier temps, les clusters contenus dans SpaCal doivent remplir les critères suivants :
  - $R_{0,cluster} = \sqrt{X_0^2 + Y_0^2} \geq 9.6$ cm

  où $R_{0,cluster}$ est la distance de l'amas dans le plan transverse à l'axe z du tube à vide. Cette coupure permet de rejeter les dépôts calorimétriques trop proches du faisceau incident de proton.

Structure de l'algorithme ("H1HfsCreatePart")

Lors de la première partie de ma thèse, j'ai implémenté trois algorithmes différents (décrits dans la partie 4.4) pour reconstruire les hadrons :

- L'algorithme "Cluster" reconstruisant les hadrons à partir des clusters calorimétriques uniquement. Chaque cluster est identifié comme hadron. Cet algorithme a lontemps été utilisé pour décrire les jets à grande impulsion transverse.
- L'algorithme "FSCOMB" [54, 55], qui a été développé pour améliorer la résolution sur les objets hadroniques : il tient compte des traces à basse impulsion transverse.
- L'algorithme "HADROO", développé dans le cadre de ma thèse qui tient compte des traces en fonction de leur erreur de mesure.

Les objets d'entrée des algorithmes développés dans $H1HfsCreatePart$ sont les classes $H1ModsEvent$ gérant l'accès à toutes les listes de particules stockées sur μODS et $H1ClusterEvent$. A partir de $H1ModsEvent$, l'algorithme trouve en premier les traces des particules chargées déjà reconstruites. Dans un second temps, les clusters $H1Cluster$ sont lus à travers $H1ClusterEvent$ et une liste de clusters corrigés $H1HfsCluster$ est produite. Nous avons donc une liste contenant les traces des particules chargées et une seconde liste contenant les clusters de bonne qualité corrigés de calibration. En itérant sur les traces des particules chargées, l'algorithme produit alors au fur et à mesure de la reconstruction, des hadrons qui sont stockés dans les listes finales de particules candidates et de particules identifiées. Celles-ci sont remplies à travers $H1ModsEvent$.

3.4 Sélection rapide (HAT)

Les fichiers HAT contiennent des informations générales de l'événement avec stockage de type simple (réel, entier). La HAT permet d'une part une sélection rapide des événements, mais aussi la création d'histogrammes en ligne. La structure contient environ deux-cents variables que l'on peut accéder facilement grâce aux pointeurs. Ces variables se divisent en plusieurs catégories :

- Des informations particulaires comme par exemple le nombre d'électrons reconstruits.
- Les grandeurs cinématiques comme par exemple Q^2 ou y_{had}.
- Les quantités décrivant l'état des sous-détecteurs comme par exemple l'état du système de déclenchement.
- Les informations supplémentaires comme par exemple le résultats des identificateurs de bruit de fond non physique [56, 57].

La liste détaillée peut être trouvée dans [58]. Lors de notre analyse, ce sont surtout les quantités décrivant l'état des sous-détecteurs qui ont été utilisées pour une sélection rapide des événements. La structure est remplie à partir des informations disponibles sur ODS ou μODS.

Chapitre 4

Identificateurs de particules

Si tu ne peux pas être une étoile au firmament,
sois une lampe dans ta maison

Proverbe arabe

Lors de la première phase de ma thèse j'ai pu non seulement contribuer à la mise en place de la structure des données décrites dans le chapitre 3, mais aussi au développement et à l'implémentation des différents identificateurs de particules. Dans la suite, je vais détailler les identificateurs de particules chargées, de particules électromagnétiques (électron et photon), de muons et de hadrons. Une attention particulière sera portée à la reconstruction des hadrons et des jets hadroniques. Pour la reconstruction des hadrons, un nouvel algorithme a été développé afin de mieux exploiter les performances du détecteur. Les jets sont reconstruits à partir de ces hadrons avec un algorithme classique de clustering. Le développement et l'implémentation a nécessité une compréhension précise des algorithmes déjà écrits en fortran. A partir de ces algorithmes existants, des améliorations ont été proposées et mises en place dans la version C++.

Afin de valider les nouveaux identificateurs, des études sur des lots physiques ont été effectuées. Nous avons comparé les résultats obtenus avec les identificateurs en C++ aux résultats obtenus avec du code fortran. Puis, pour valider les nouvelles performances, des comparaisons entre les particules reconstruites et les particules générées pour des simulations Monte-Carlo ont été effectuées. Enfin, dans certains cas, nous avons aussi comparé les données aux simulations Monte-Carlo pour valider par exemple les calibrations en énergie appliquées. Ces études ont été faites dans le domaine cinématique pertinent de notre analyse : nous nous sommes donc restreints au domaine dit de grande impulsion transverse avec des particules de plus 10 GeV d'impulsion transverse. Ces études ont permis de valider les algorithmes et de jouer le rôle important de débogage des nouveaux outils.

Pour les études de validation, ainsi que dans notre analyse, nous utilisons les variables suivantes :

- La position en z du point d'interaction appelée z_{vtx} est mesurée à partir des ajustements de traces contraintes. Cette variable permet de rejeter le bruit de fond ne venant pas des interactions électron-proton.
- La variable $E - P_z$ est calculée à partir des grandeurs cinématiques de toutes les particules d'un événement et vaut, dans le cas où toute l'énergie a été mesurée, $2E_{ele}^{ini}$ =55,2 GeV.
- Les variables indiquées par gen sont calculées à partir des particules générées du processus pour une simulation Monte-Carlo. N_{jet}^{gen} désigne par exemple le nombre de jets reconstruits au niveau hadronique.
- Lorsqu'il y a lieu, les variables indiquées par rec désignent les variables reconstruites dans le détecteur. $P_{T,ele}^{rec}$ est par exemple l'impulsion transverse de l'électron reconstruit. Lorsqu'il n'y a aucune indication, il s'agit des variables reconstruites dans le détecteur.
- Les indices ele, jet, μ et γ se rapportent respectivement à l'électron, au jet, au muon et au photon.

4.1 Reconstruction des particules chargées

Au niveau de la reconstruction fortran des événements, nous distinguons deux types de traces selon leur qualité de mesure :

- **Les traces DTNV**
  Il s'agit de segments de traces élémentaires. Ces traces $DTNV$ ne sont pas obligatoirement contraintes au vertex primaire. De plus, la même trace physique peut être présente plusieurs fois : en effet, une particule chargée peut être reconstruite lorsqu'elle traverse les trajectographes centraux (CJC) et avant (FTD), avec des paramètres différents. De plus, elle peut être contrainte au vertex primaire et/ou à un vertex secondaire reconstruit. Toutes les différentes hypothèses sont gardées.
- **Les traces contraintes au vertex DTRA**
  Les traces $DTRA$ sont définies à partir de toutes les traces $DTNV$ ajustées au point d'interaction primaire de l'événement. Les différentes hypothèses sont aussi gardées.

La sélection dite de Lee West [52] permet de résoudre le problème des traces à plusieurs hypothèses en imposant un ordre de préférence. De plus, des critères de qualité spécifiés ci-dessous permettent d'affiner la sélection finale. Dans le module $H1TrkFinder$, la sélection finale de Lee West a été programmée. Les particules chargées identifiées à partir de ces traces sont alors stockées. Nous allons juste évoquer ici les principes fondamentaux et les critères utilisés.

4.1.1 Critères de sélection

L'algorithme implémenté est basé sur les traces dites *DTRA* contraintes au vertex et stockées au niveau des *ODS*. Il permet de lever le problème des hypothèses multiples et d'affiner la sélection. L'algorithme procède de la manière suivante :

- Dans un premier temps, les coupures issues de la sélection d'électrons et de muons décrites dans [52] sont appliquées à toutes les traces contraintes au vertex.
- Puis, dans un deuxième temps, la meilleure hypothèse de trace est retenue.

Une trace *DTRA* quelconque peut être reconstruite avec les hypothèses suivantes :

- Elle est uniquement mesurée par le trajectographe central CJC (dénommée C),
- Elle est mesurée par le trajectographe avant FTD (dénommée F),
- Elle est mesurée par la combinaison des deux (dénommée K).

De plus, elle peut être associée au vertex primaire (dénommé "1") ou vertex secondaire (dénommé "2"). L'ordre de préférence "KCF" et "12" par exemple, va contraindre l'algorithme à reconstruire préférentiellement les traces qui ont été mesurées en même temps par les trajectographes central et avant, et contraintes au vertex primaire. S'il ne trouve pas de trace, il sélectionnera les traces purement centrales et ainsi de suite. Dans la reconstruction standard, l'ordre "KCF" et "12" a été choisi afin d'améliorer l'efficacité d'identification des traces à grande impulsion transverse dans la partie avant du détecteur.

4.1.2 Etudes de validation

L'algorithme a été validé une première fois juste après son implémentation sur un lot de données 1999 de 34551 événements aléatoires. Cette validation consistait à comparer les résultats de la sélection écrite en C++ aux résultats obtenus avec le code fortran.

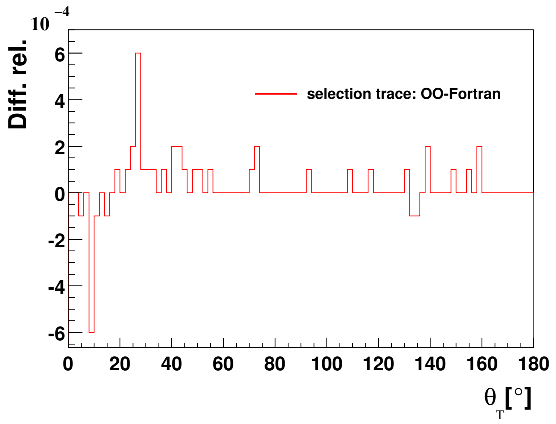


FIG. 4.1 – Distribution de la différence relative du nombre de traces reconstruites en C++ et fortran en fonction de l'angle polaire pour un lot de données 1999.

La figure 4.1 montre la différence relative du nombre de traces entre les deux algorithmes en fonction de l'angle polaire θ des traces. Nous remarquons un très bon accord avec des différences maximales entre le nombre de traces sélectionnées en C++ et en fortran inférieures à 10^{-3}.

Les différences observées sont dues à une définition légèrement différente de la longueur d'une trace R_{long} :
- OO : $R_{long}(K) = R_{end}(F)-R_{start}(C)$ (K : Trace combinée F : Trace avant, C : Trace centrale)
- Fortran : $R_{long}(K) = R_{end}(C)$

où R_{start} et R_{end} sont les distances radiales respectivement du premier et dernier dépôt d'énergie associé au passage d'une particule chargée dans les trajectographes. J'ai aussi verifié la validité de la sélection sur un lot de simulation Monte-Carlo de courants chargés à grand transfert d'impulsion $Q^2 > 100$ GeV^2 (4000 événements).

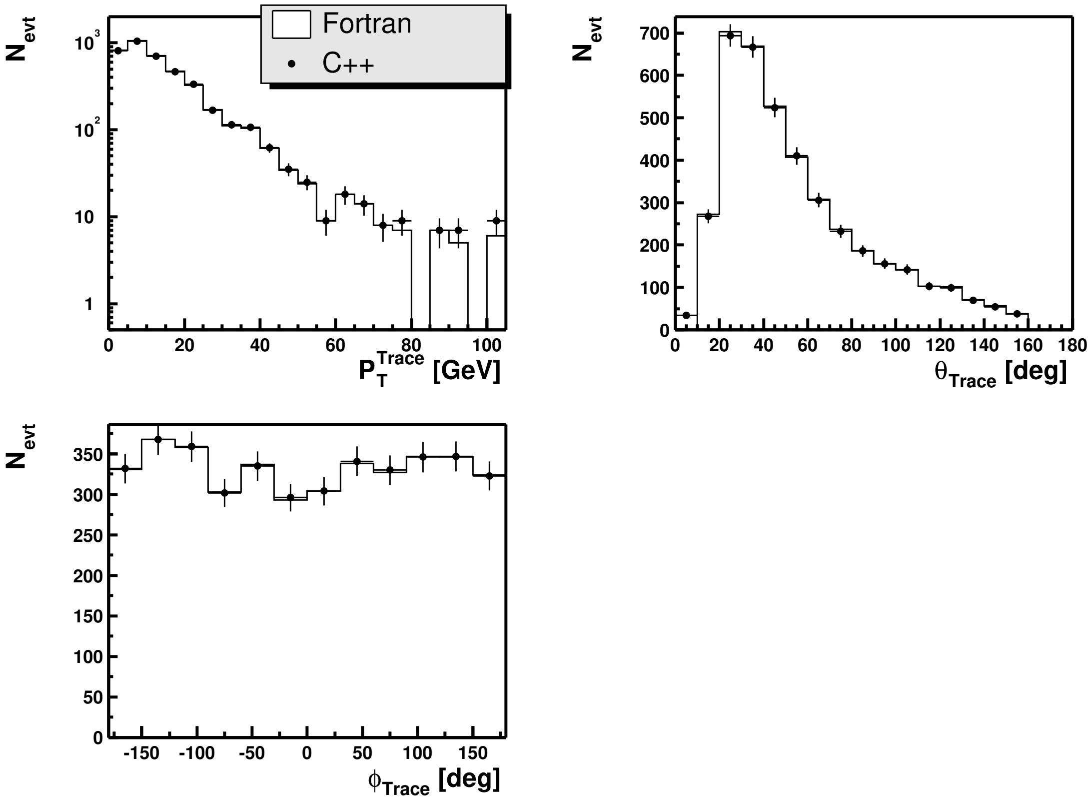


Fig. 4.2 – Distribution de l'impulsion transverse, de l'angle polaire et azimutal de la trace de plus grande impulsion transverse pour un lot de simulation Monte-Carlo à courant chargé.

D'après les distributions de la figure 4.2 montrant l'impulsion transverse, l'angle polaire et l'angle azimutal de la trace de plus grande impulsion transverse, nous remarquons un bon accord global entre les versions fortran et C++.

4.2 Reconstruction des électrons et photons

Nous allons discuter les principales propriétés du dépôt électromagnétique dans le calorimètre LAr utilisées pour identifier les électrons et les photons. Nous faisons l'hypothèse que les propriétés de leur gerbe sont identiques. Le codage de l'algorithme de reconstruction d'un photon ou électron est basée sur un algorithme existant déjà en fortran [59].

4.2.1 Propriétés des dépôts électromagnétiques

Nous présentons uniquement les caractéristiques des dépôts électromagnétiques dans le calorimètre LAr car c'est le détecteur où sont reconstruits les électrons et photons à grande impulsion transverse que nous analysons. Par la suite, nous utiliserons le terme électron pour désigner de façon générale un électron ou photon.
Nous reconstruisons l'électron à partir de clusters dans le calorimètre. L'identification inclut des coupures générales, et des estimateurs de description d'un dépôt typique d'électron en tenant compte des caractéristiques du détecteur. Les détails sont données dans la thèse [60].
Les clusters candidats doivent remplir les critères suivants :

- **L'impulsion transverse du dépôt** : P_t^{ele} doit être supérieure à 3 GeV.
- **L'énergie du dépôt** : E^{ele} doit dépasser 5 GeV.
- **Le nombre de cellules associées doit être supérieure à 3**

Nous définissons des candidats électrons à partir des caractéristiques suivantes du dépôt :

- **La fraction électromagnétique** : f_{em}
  Afin de rejeter les pions qui n'interagissent presque pas dans la section électromagnétique (dont la longueur est d'environ 1λ), on définit f_{em} = $\mathrm{E}_{em}/\mathrm{E}_{tot}$ où seules les cellules ayant une énergie positive sont considérées. Cette fraction est proche de un pour les électrons, excepté dans les insterstices en ϕ et en z, et autour de $\theta_{em} \approx 90^o$ où l'électron dépose de l'énergie dans la première couche hadronique.
- **Le rayon transverse** : σ_r
  La taille transverse d'une gerbe électromagnétique est de l'odre du rayon de Molière R_M, qui est bien inférieur à la longueur d'interaction λ_l. On calcule pour cela l'écart type transverse $\sigma_r = \sqrt{< r^2 > - < r >^2}$ où r^n sont les moments d'ordre n des distances transverses (r_i) des cellules par rapport à l'axe de la gerbe. Ces moments s'écrivent : $< r^n > = \mathrm{w}^{-1}\ \Sigma w_i r_i^n$ avec $\mathrm{w}_i = E_i/V_i$ et $\mathrm{w} = \Sigma\ w_i$. V_i et E_i sont respectivement les volumes et les énergies des cellules. Pour les électrons, ce rayon est faible.
- **La fraction contenue dans le cœur chaud** : $\mathrm{f}_{Hot\ N}$
  Le profil longitudinal de dépôt d'énergie étant caractérisé par une courbe en cloche et puisque la taille transverse est faible, la gerbe comporte un cœur chaud. Une grande fraction de l'énergie d'un électron est donc déposée dans un nombre restreint de cellules voisines. Cette fraction est définie par $\mathrm{f}_{Hot\ N} = \mathrm{E}_{Hot\ N}/E_{em}$ où $\mathrm{E}_{Hot\ N}$ correspond à l'énergie contenue dans le groupe de N

cellules voisines le plus énergétique. Ici encore on ne considère que les cellules ayant une énergie positive. La taille des cellules variant avec θ, il faut faire varier N pour que cet estimateur soit à peu près indépendant de l'angle polaire.

– **Le critère d'isolation du candidat** : f_{iso}
  Pour estimer l'isolation du candidat on utilise l'estimateur $f_{iso} = E_{tot}/E_{iso}$ où E_{tot} est l'énergie de l'amas et E_{iso} est l'énergie comprise dans le cône d'isolation dont l'axe passe par le vertex et le barycentre de l'amas, dont le sommet est le vertex et l'ouverture en pseudorapidité η et l'angle azimutal ϕ de $\sqrt{\eta^2 + \phi^2} <$ 0.25 radian. Lorsqu'un électron n'est pas exactement isolé, cela est dû soit à du bruit électronique dans les cellules qui se trouvent à l'arrière, soit à la présence d'un photon rayonné dans l'état final. Dans le premier cas, nous nous attendons à avoir au plus quelques centaines MeV dans la section hadronique. Dans le second cas, l'énergie déposée en dehors du cône d'isolation se trouve dans la section électromagnétique. Ainsi, pour qu'un amas soit considéré comme un candidat électron il faut qu'il vérifie : soit f_{iso} >0.98 ou f_{iso} >0.95 et $E_{iso}^{had} <$ 300 MeV où E_{had} est l'énergie hadronique dans le cône d'isolation entourant l'électron.

Tous les candidats ainsi trouvés sont ordonnés par P_T décroissant et stockés sur μODS. A ce stade, nous ne faisons aucune distinction entre un électron et un photon. Néanmoins, pour toute particule dite électromagnétique, des informations supplémentaires sont stockées. Nous détaillons deux informations supplémentaires utilisées pour sélectionner des candidats électron et photon dans notre analyse :

– **Critère d'isolation calorimétrique :**
  Afin de sélectionner des électrons et photons isolés par rapport à des hadrons, un critère d'isolation supplémentaire est introduit au niveau particulaire. Nous parlons d'électron ou photon isolé lorsque l'énergie calorimétrique supplémentaire contenue dans un cône de rayon 1,0 dans le plan (η-ϕ) autour de l'électron (ou photon) est inférieure à 3% de l'énergie de l'électron (ou photon).
– **Critère de trace associée :**
  Pour distinguer l'électron du photon, nous calculons la distance de la trace la plus proche au point d'impact dans le calorimètre par rapport au dépôt calorimétrique. Si celle-ci est inférieure à 12 cm, la trace est associée au dépôt et la particule est identifiée comme électron.

Ces critères sont nécessaires, mais pas suffisants pour distinguer l'électron du photon. Nous allons détailler des critères supplémentaires dans le chapitre 5.

4.2.2 Etudes de validation

Afin de valider l'algorithme, nous avons procédé à trois vérifications différentes. Dans un premier temps, nous avons testé la validité de l'implémentation en comparant le quadri-vecteur impulsion-énergie de l'électron reconstruit par rapport au quadri-vecteur vrai pour un lot de simulation Monte-Carlo à courant neutre à grand transfert d'impulsion $Q^2 > 100$ GeV^2 avec la sélection suivante :

– $|z_{vtx}| < 40$ cm

– $N^{rec}_{ele,jet} \geq 1$, $N^{gen}_{ele,jet} \geq 1$
– $P_T^{jet,ele\ rec} > 20$ GeV, $P_T^{jet,ele\ gen} > 20$ GeV
– $10^o < \theta^{rec}_{jet,ele} < 150^o$, $10^o < \theta^{gen}_{jet,ele} < 150^o$

Les critères requis pour l'électron et le jet sont les critères de sélection serrés de l'analyse, détaillés dans les parties respectives 5.1.1 et 5.1.4. Ce lot nous permet d'estimer les résolutions angulaires et en énergie sur l'électron reconstruit.

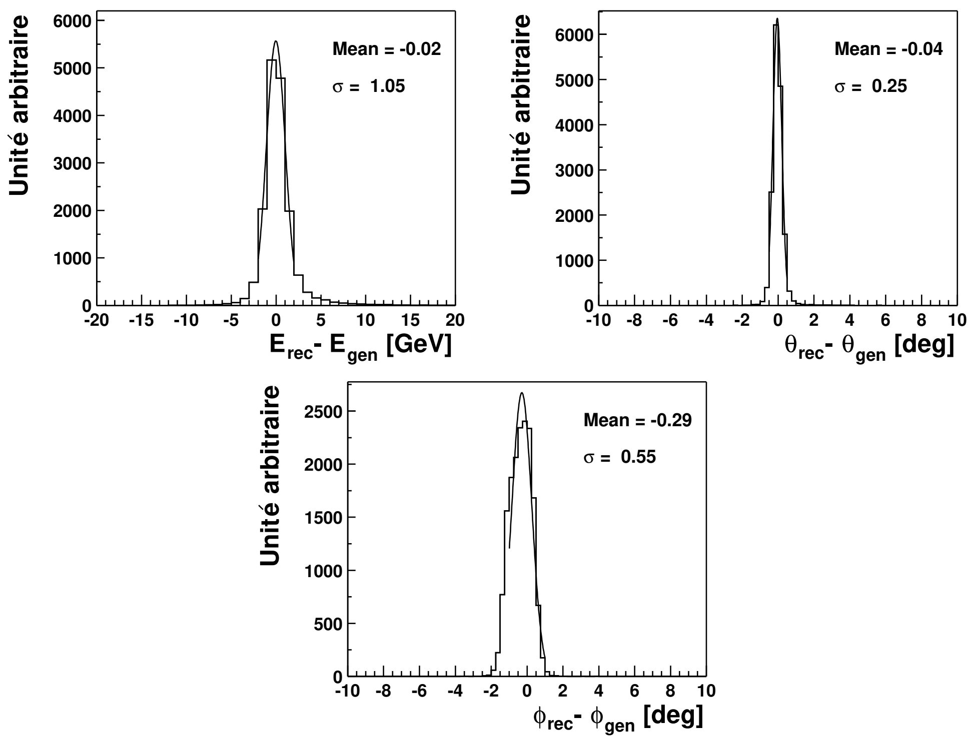


FIG. 4.3 – Comparaison de l'énergie et des angles polaires et azimuthaux reconstruits de l'électron par rapport aux quantités générées pour des simulations Monte-Carlo à courant neutre.

D'après les distributions de la figure 4.3, nous obtenons une résolution en énergie d'environ 1 GeV, une résolution angulaire de 0,25^o (4 mrad) pour l'angle polaire et 0,55^o (10 mrad) pour l'angle azimutal. Nous remarquons l'impact du champ magnétique dans le plan transverse car la distribution ϕ_{rec}-ϕ_{gen} est décalée de 0,29^o.

Dans un second temps, nous avons vérifié la validité de l'implémentation en comparant les résultats obtenus avec l'algorithme écrit en C++ aux résultats obtenus avec le code fortran existant. Pour cela, un lot de simulation Monte-Carlo à courant neutre à grand transfert d'impulsion $Q^2 > 100$ GeV^2 a été utilisé. La seule coupure appliquée consiste à demander la présence d'un vertex primaire reconstruit à moins de 40 cm du point d'interaction nominal. Nous obtenons un lot avec 15563 événements en C++ et 15520 événements en fortran.

Les distributions de la figure 4.4 montrent un bon accord entre l'implémentation de l'algorithme en C++ et du code fortran existant. Les distributions de P_T^{ele} et

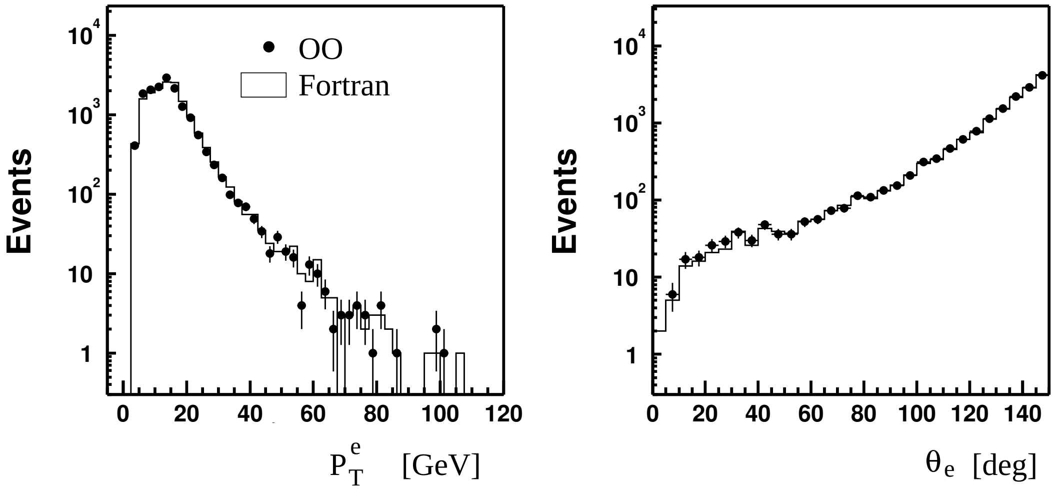


FIG. 4.4 – Distributions de P_T^{ele} et de θ_{ele} de l'électron diffusé pour une simulation Monte-Carlo de processus courant neutre.

de θ_{ele} sont celles de l'électron diffusé. Les différences résiduelles sont dues à une calibration différente de l'energie calorimétrique entre les deux implémentations.

4.2.2.1 Calibration de l'énergie des électrons

Nous allons étudier la calibration des particules qui déposent leur énergie dans la section électromagnétique du calorimètre. La réponse du calorimètre doit être calibrée afin d'obtenir une précision maximale sur l'énergie mesurée. Comme tous les objets physiques que nous recherchons sont reconstruits dans le LAr, nous nous concentrons sur la calibration de celui-ci. Cette calibration peut se faire en utilisant des événements dans lesquels l'énergie d'une particule est contrainte par d'autres variables indépendantes de la réponse en énergie du calorimètre. C'est le cas des événements profondément inélastiques avec échange de courant neutre : ces événements étant sur-contraints cinématiquement (voir partie 1.1), les énergies de l'électron et du jet hadronique peuvent être prédites à partir de leurs angles polaires $\theta_e^{'}$ et γ. Il est donc possible de comparer, pour un événement donné, la mesure directe de l'énergie d'un électron avec celle fournie par la méthode dite "double angle" qui donne l'énergie E_{DA} de l'électron :

$$E_{DA} = \frac{2E_e \sin\gamma}{\sin\gamma + \sin\theta_e' - \sin(\gamma + \theta_e')} \tag{4.1}$$

$$Q^2_{DA} = \frac{4E_e^2 \sin\gamma(1 + \cos\theta_e')}{\sin\gamma + \sin\theta_e' - \sin(\gamma + \theta_e')} \tag{4.2}$$

$$x_{DA} = \frac{E_e}{E_p}\frac{\sin\gamma + \sin\theta_e' + \sin(\gamma + \theta_e')}{\sin\gamma + \sin\theta_e' - \sin(\gamma + \theta_e')} \tag{4.3}$$

où E_e et E_p sont respectivement l'énergie de l'électron et du proton incident. Nous obtenons aussi $P_{T,DA}$ à partir de la relation suivante : $P_{T,DA} = Q^2_{DA}/(1 - y_{DA})$, avec

$y_{DA} = Q^2_{DA}/(x_{DA}s)$. Nous allons utiliser cette méthode pour vérifier la calibration en énergie des électrons, définie dans [2], sur un lot de donnés e^+p 99-00 de courant neutre à grand transfert d'impulsion Q^2. Ce lot est sélectionné de la manière suivante :

- $|z_{vtx}| < 35$ cm
- Un dépôt électromagnétique isolé
- Un seul jet reconstruit
- $43 < E - P_z < 67$ GeV
- $P_T^{ele,jet} > 10$ GeV

Aucun critère supplémentaire de qualité de l'électron, comme par exemple l'association d'une trace n'est requis.

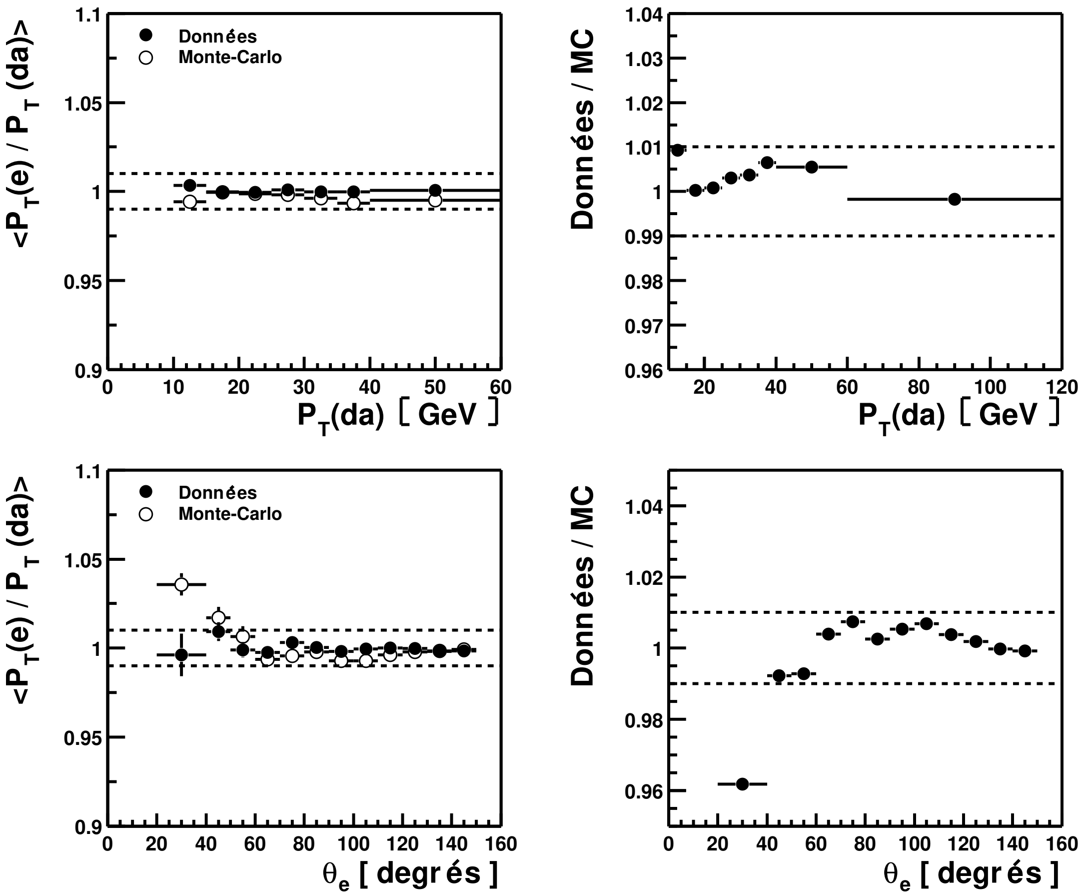


FIG. 4.5 – Distributions de gauche : Rapport P_T^{Bal} en fonction de P_T^{DA} (en-haut) et θ_{ele} (en-bas) pour les données et les simulations Monte-Carlo. Distributions de droite : Double rapport données sur simulation Monte-Carlo de P_T^{Bal} (les bandes indiquent un écart de $\pm$ 1%).

Les distributions de P_T^{Bal}=P_T^{ele}/P_T^{DA} en fonction de P_T^{DA} et θ_{ele} (figures de gauche de 4.5) montrent que la calibration absolue est obtenue avec une erreur plus petite que 1%, sauf à l'avant, région où la statistique restre trop faible pour calibrer avec une meilleure précision. Les figures de droite de 4.5 montrent la calibration relative : l'accord entre les données et les simulations est meilleur que 1% dans la partie centrale.

4.3 Reconstruction des muons

La reconstruction des muons est basée sur la reconnaissance de signatures typiques de muons dans le détecteur. Leur masse en font des particules très pénétrantes, mesurables dans le fer instrumenté qui entoure le calorimètre.

4.3.1 Principes

Les muons peuvent être identifiés dans le fer instrumenté soit à partir des segments de traces, soit à partir des dépôts d'énergie. L'algorithme actuel utilise pour l'identification cinq types de muons selon des critères de qualité. On associe à chaque type un critère qualité ("grade") allant de 1 à 5 : Le grade "1" désigne les muons de meilleure qualité, c'est-à-dire reconstruits avec la plus grande certitude, et "5" ceux de plus mauvaise qualité, identifiés avec la moins grande précision. Les critères de qualité sont les suivants :

- **Les muons de grade 1 :**
  La reconstruction est faite à partir d'une trace centrale et un segment de trace dans le fer instrumenté, en utilisant le calcul de Bethe-Bloch pour estimer l'énergie perdue par le muon : on extrapole alors la trace centrale jusqu'au segment et on calcule une valeur de χ^2 entre la trace centrale et la trace dans le fer. A partir du χ^2, on déduit un niveau de confiance. Ce niveau de confiance permet d'estimer la probabilité de concordance entre les deux traces. Le grade "1" est alors associé à un candidat muon lorsque la probabilité de concordance dépasse un certain seuil.
- **Les muons de grade 2 :**
  Ce type de muon est reconstruit de la même manière que le précédent sauf qu'aucune probabilité minimale n'est requise. L'association d'une trace centrale à une trace reconstruite dans le fer instrumenté est calculée à partir de leur distance $\Delta_{\eta-\phi}$ dans le plan $(\eta - \phi)$. Elle doit être inférieure à 0,5 radian.
- **Les muons de grade 3 :**
  Le détecteur de fin de gerbe (abrégé TC en anglais pour $Tail\ Catcher$) destiné à mesurer les hadrons déposant encore au-delà de la section hadronique du calorimètre peut être utilisé pour augmenter l'efficacité de reconstruction des muons. Souvent le nombre de dépôts permettant de reconstruire une trace dans le système à muons est plus faible lorsque le muon traverse des régions à l'intersection de deux modules du fer instrumenté. En utilisant les critères décrits dans la documentation [47], on peut augmenter l'efficacité de reconstruction, tout en supprimant la misidentification due aux hadrons. La trace centrale, ainsi que les dépôts calorimétriques dans le détecteur de fin de gerbes doivent être alignés dans le plan $(\eta - \phi)$: leur distance doit être inférieure à

0,5 radian.
- **Les muons de grade 4 :**
Les muons de basse énergie sont souvent arrêtés dans le calorimètre LAr avant d'atteindre le fer instrumenté et peuvent être identifiés à partir de leur signature dans le calorimètre uniquement. Un jeu de variables discriminantes décrit dans [47] est utilisé pour comparer la signature d'une particule candidate à celle d'une particule à ionisation minimale comme le muon.
- **Les muons de grade 5 :**
On ne peut pas toujours associer une trace centrale de bonne qualité aux muons produits à très bas angle polaire, à l'avant. Cependant le toroïde du détecteur avant de muon permet de mesurer l'impulsion avec une bonne précision. Des muons de grade "5" peuvent alors être identifiés à partir de l'information du toroïde avant seul.

L'algorithme final tient compte, pour chaque trace, de toutes les hypothèses de muons possible et choisit celle de niveau de confiance le plus élevé, c'est-à-dire avec de grade le plus faible. L'impulsion transverse est mesurée à partir des trajectographes centraux (CJC), avant (FTD) ou le toroïde suivant l'angle polaire.

Enfin, un critère d'isolation calorimétrique supplémentaire est introduit au niveau particulaire afin de distinguer les muons isolés : les amas calorimétriques mesurés par le calorimètre LAr et situés dans un cylindre d'un rayon de 25 cm dans la section électromagnétique et 50 cm dans la section hadronique autour de la trace centrale du muon sont associés au muon. Un muon est dit isolé dans le calorimétre lorsque l'énergie de ces amas est inférieure à 8 GeV.

4.3.2 Etudes de validation

Nous avons fait deux types d'études. Dans un premier temps, nous avons comparé le quadri-vecteur reconstruit au quadri-vecteur généré du muon pour une simulation Monte-Carlo de production de paires de muons. Pour cela, nous avons sélectionné un échantillon avec les critères suivants :
- $|z_{vtx}| < 40$ cm
- $N_\mu^{rec} \geq 2$, $N_\mu^{gen} \geq 2$
- $P_T^{\mu\ rec} > 20$ GeV, $P_T^{\mu\ gen} > 20$ GeV
- $10^o < \theta_\mu^{rec} < 150^o$, $10^o < \theta_\mu^{gen} < 150^o$

Nous obtenons un échantillon contenant au moins deux muons reconstruits d'impulsion transverse supérieure à 20 GeV. Les distributions de la figure 4.6 montrent que la résolution en énergie est d'environ 1,4 GeV et la résolution angulaire de 0,13^o (2 mrad) pour l'angle polaire.

Dans un second temps, nous avons vérifié la validité de l'algorithme codé en comparant les distributions de muons reconstruits avec le code fortran au distributions obtenues avec l'algorithme C++ sur un lot de simulation Monte-Carlo de production de boson W$^\pm$. Nous recherchons alors les événements contenant le muon issu de la désintégration du boson W$^\pm$:
- $|z_{vtx}| < 40$ cm

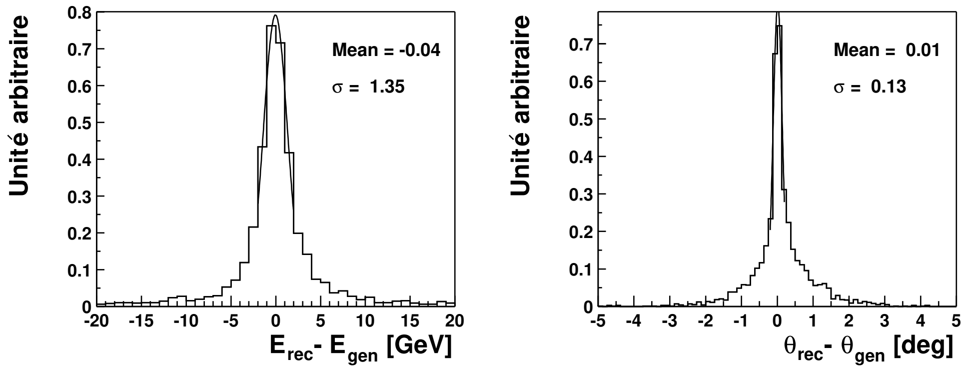


FIG. 4.6 – Comparaison de l'énergie et de l'angle polaire reconstruit du muon par rapport aux quantités générées pour des simulations Monte-Carlo de production de paires de muons.

- $N^{\mu}_{rec} >= 1$
- $P^{\mu}_{T} > 2.0$ GeV

Après ces coupures, subsistent 2326 événements dans le lot C++ et 2356 événements dans le lot fortran.

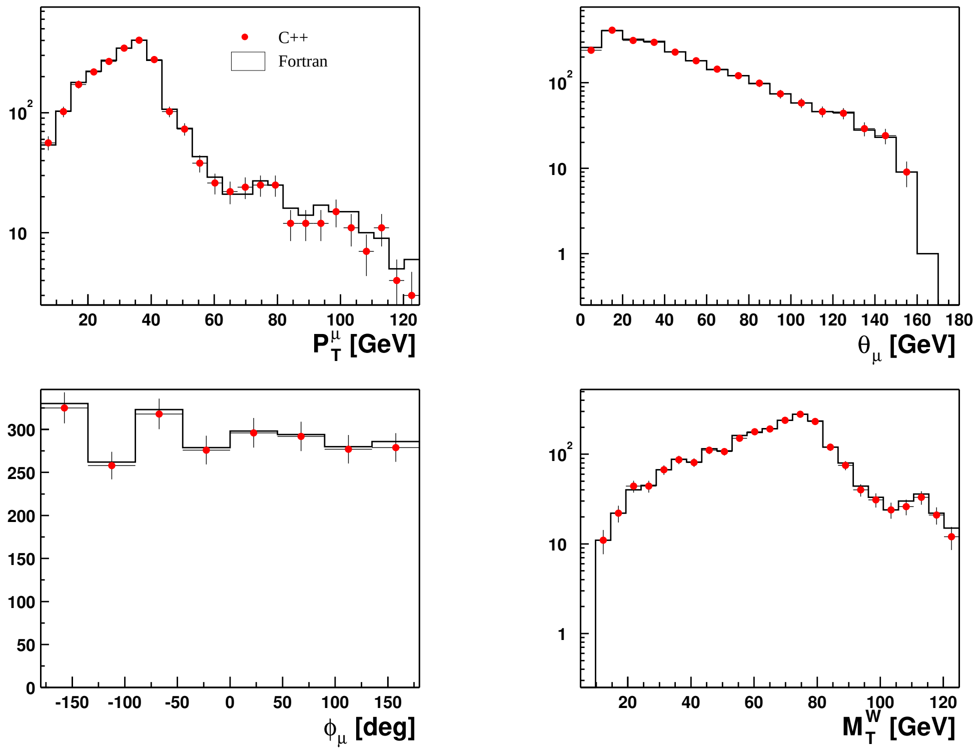


FIG. 4.7 – Distributions de P^{μ}_{T}, θ_{μ}, ϕ_{μ} et de la masse dans le plan transverse M^{W}_{T} du muon pour un échantillon de simulation Monte-Carlo de production de $W \to \mu\nu$.

Les distributions de P_T^μ, θ_μ, ϕ_μ et M_T^W de la figure 4.7 montrent un accord remarquable de l'ordre de 1% entre les deux algorithmes. Seule, une différence pour la région très à l'avant ($\theta_\mu < 20^o$) subsiste : dans l'algorithme C++ moins de muons sont reconstruits. Cela est dû à un critère d'identification défini différemment entre les deux algorithmes : le muon doit déposer de l'énergie dans au moins six couches du fer instrumenté dans la partie avant pour le code C++, alors que dans le code fortran le minimum requis est deux.

4.4 Reconstruction des objets hadroniques

Nous allons décrire dans cette partie les trois algorithmes implémentés pour reconstruire les hadrons, chargés ou non et montrer leurs performances. Il s'agit de deux algorithmes transcrits de programmes existants en fortran, et, d'un troisième, développé en s'inspirant des algorithmes existants, afin d'améliorer, entre autre, la détermination de l'énergie transverse des jets. Un jet hadronique contient principalement des pions chargés et des pions neutres. La gerbe qu'il forme dans le calorimètre est la superposition d'une gerbe électromagnétique ($\pi_0 \rightarrow \gamma\gamma$) et d'une gerbe hadronique.
Nous pouvons diviser les algorithmes de reconstruction hadronique en deux catégories : d'une part, il existe l'algorithme utilisant l'information calorimétrique uniquement (appelé $Cluster$) et d'autre part les deux autres, combinés (appelés $FSCOMB$ et $HADROO$) utilisant l'information calorimétrique et l'information des trajectographes. Pour ces deux derniers, nous parlons alors de reconstruction du flux d'énergie. L'idée à la base d'un algorithme de reconstruction du flux d'énergie ou "energy flow" est de mettre en commun les informations provenant de plusieurs sous-détecteurs. En effet chaque sous-détecteur ne possède pas la même couverture en acceptance, mesure mieux les basses énergies ou les plus hautes, et a des zones mortes à différents endroits. En sélectionnant chaque sous-détecteur dans le domaine cinématique où sa résolution est optimale on obtient une mesure plus précise que celle fournie par un détecteur unique.
Dans la structure du code orienté objet, décrit dans la partie 3.1, l'algorithme de reconstruction hadronique est lancé à la fin de la chaîne de reconstruction. Ainsi, les dépôts et traces associés aux leptons bien identifiés et isolés ont été déjà verrouillés, ce qui facilitera la tâche. Ces algorithmes sont intégrés dans le module $H1HfsFinder$ du programme officiel de reconstruction.

4.4.1 Algorithme calorimétrique $Cluster$

Nous savons pertinemment que la mesure d'énergie du calorimètre à grande impulsion transverse est dotée d'une erreur moins importante que la mesure de l'impulsion, et donc de l'énergie, à partir des traces. En se plaçant dans cette hypothèse, la reconstruction des hadrons peut être basée sur les objets calorimétriques uniquement. Sachant que les identificateurs de leptons ont déjà verrouillé les leptons bien identifiés, les traces et clusters subsistants peuvent être associés aux hadrons. Ainsi,

l'algorithme appelée *Cluster* identifie alors pour chaque cluster calorimétrique, un hadron qu'il soit chargé ou non. Les clusters calorimétriques utilisés proviennent d'une sélection, décrite dans la partie 3.2.2, des clusters stockés sur *ODS*. Cette sélection corrige les effets expérimentaux comme par exemple les bruits analogiques simulant un dépôt d'énergie concentré dans une cellule.
Finalement, nous obtenons donc une liste de hadrons stockée sur μODS et pouvant être utilisée par l'identificateur de jets. Les performances de cet algorithme seront décrites dans la partie 4.4.3, en comparant ses résultats à ceux des autres algorithmes.

4.4.2 Algorithme combiné *FSCOMB*

L'algorithme combiné*FSCOMB* a été développé [54, 55] pour améliorer la résolution de la mesure en énergie de l'état hadronique final. Les objets d'entrée sont d'une part les objets calorimétriques et d'autre part les traces centrales associées aux particules chargées. L'idée de base est d'utiliser l'information des traces à basse énergie, où elles sont les plus précises, et celles du calorimètre à haute énergie, où sa résolution est la meilleure.
L'algorithme associe alors des traces à des clusters pour reconstruire les hadrons chargés et les hadrons neutres. L'impulsion de chaque trace de moins de 2 GeV est comparée à l'énergie déposée dans un cylindre de 25 cm de rayon autour de l'impact de la trace dans le calorimètre électromagnétique LAr et de 50 cm de rayon dans le calorimètre hadronique LAr. Dans tous les cas, la trace est gardée en tant qu'objet hadronique. Si l'impulsion de la trace est plus faible que l'énergie contenue dans le cylindre associé, la mesure de la trace est gardée et, successivement, les dépôts calorimétriques contenus dans le cylindre sont verrouillés jusqu'à atteindre l'énergie de la trace [1]. Si l'impulsion de la trace est plus importante, elle est aussi gardée en tant qu'objet hadronique et tous les clusters contenus dans le cylindre sont verrouillés. Les clusters restants à la fin, c'est-à-dire ceux qui n'ont pas été associés à une trace de moins de 2 GeV sont ajoutés au flot hadronique.

Cet algorithme ne considère que les traces centrales d'impulsion transverse en-dessous de 2 GeV pour les raisons suivantes :
- Les paramètres et quadri-vecteurs des traces avant, c'est-à-dire reconstruites uniquement à partir du trajectographe avant ($\theta_{trace} < 25^o$), sont, dans des événements à grande impulsion transverse, parfois mal mesurées et mal décrites par les simulations Monte-Carlo.
- Les traces d'impulsion transverse supérieure à 2 GeV étaient, au moment de la conception de l'algorithme, peu utilisées et peu étudiées dans les analyses à grande impulsion transverse.

Les performances de *FSCOMB* seront détaillées dans la partie 4.4.3 en comparant ses résultats aux autres algorithmes.

[1] L'énergie de la trace est donnée par son impulsion avec l'hypothèse de masse de pions.

4.4.3 Algorithme combiné $HADROO$

Afin d'exploiter au mieux les performances du détecteur, l'algorithme combiné $HADROO$ (*Hadronic Reconstruction in OO*) a été développé en s'inspirant de l'algorithme $FSCOMB$ décrit précédemment. Nous avons alors redéfini les contraintes d'un tel algorithme :

- L'identification de hadrons doit être faite à l'aide de toute l'information disponible des sous-détecteurs.
- A grande impulsion transverse, l'énergie totale mesurée à partir des hadrons doit correspondre approximativement à l'énergie totale déposée dans le calorimètre.
- Il doit garder le plus grand nombre de traces de bonne qualité possible en tant que particules hadroniques.
- Il doit éviter le double comptage en énergie, c'est-à-dire la sommation des informations redondantes de plusieurs sous-détecteurs correspondant à la même particule.
- A bas P_T, la résolution en énergie doit être améliorée par rapport à $FSCOMB$.

L'algorithme est divisé en deux étapes distinctes, pour lesquelles deux paramètres sont introduits.
Il associe les clusters aux traces et définit les quadri-vecteurs des objets hadroniques en fonction des erreurs de mesure de l'impulsion de la trace et de l'énergie du cluster :

- A bas P_T, toutes les traces sont gardées et l'énergie calorimétrique associée à la trace, sous forme de clusters, est diminuée de celle des traces.
- A grand P_T, des estimateurs pour calculer la compatibilité en énergie entre la trace et le cluster sont utilisés pour décider quelle information est plus précise.

La partie nouvelle de l'algorithme par rapport à $FSCOMB$ réside dans le traitement de la transition entre les deux domaines, qui dépend à la fois de la trace et de la résolution attendue du calorimètre. Dans l'algorithme, décrit en détails dans [47] nous introduisons un paramètre, appelé α, qui doit être ajusté afin de permettre une transition régulière entre ces domaines en P_T. Celle-ci est définie par l'inéquation suivante :

$$\left(\frac{dP_T^{trace}}{P_T^{trace}}\right) < \alpha\,\frac{50\%}{\sqrt{E}_{trace}} \tag{4.4}$$

où P_T^{trace} est l'impulsion transverse de la trace et E_{trace} son énergie. La partie de gauche décrit l'erreur relative sur l'impulsion transverse de la trace donnée par les trajectographes. La partie de droite décrit l'erreur de mesure attendue sur l'énergie déposée dans le calorimètre [32] multipliée par la constante α, à ajuster. Cette inéquation décrit donc à partir de quelle énergie mesurée, il nous faut passer de la mesure de trace à la mesure calorimétrique. L'inéquation 4.4 est basée sur la mesure de P_T^{trace} et E^{trace} uniquement. Elle constitue une prescription permettant de définir séparément les deux domaines à partir de l'erreur de l'impulsion et non plus à partir de l'impulsion même des traces.

A grand P_T, afin de déterminer la compatibilité en énergie entre la trace et le cluster, nous introduisons un deuxième paramètre, appelé N_σ. Il permet de définir, à partir de quelle valeur la mesure de l'énergie d'une particule chargée n'est plus considérée comme compatible avec l'énergie calorimétrique associée.

Le traitement spécifique de l'algorithme dépend donc du domaine en P_T et est décrit en détails dans l'appendice A.

Paramètres

Nous avons donc introduit deux paramètres dans l'algorithme :
- α : Paramètre séparant les régimes à grand et à bas P_T. Si les résolutions du détecteur sont bien comprises, on s'attend à une valeur optimale de α de l'ordre de 1.
- N_σ : Paramètre définissant la fenêtre de compatibilité entre l'énergie du dépôt calorimétrique et celle de la trace :
  $N_\sigma = 2.0 \rightarrow E_{trace} \in [E_{calo} - 2.0\sigma_{calo}, E_{calo} + 2.0\sigma_{calo}]$

où E_{trace} est l'énergie de la trace et E_{calo} celle des dépôts calorimétriques associés. Leur ajustement est décrit dans la partie suivante.

4.4.4 Résultats et validation

Dans cette partie, nous allons détailler les résultats obtenus avec l'algorithme combiné $HADROO$. Nous allons montrer dans une première étape comment le comportement de l'algorithme varie en fonction des paramètres α et N_σ. Ensuite, les résultats obtenus avec $HADROO$ seront comparés aux résultats obtenus avec les deux algorithmes $Cluster$ et $FSCOMB$.
Pour étudier le comportement en fonction du paramètre α, nous avons fixé le second paramètre N_σ dans un premier temps à 2.0. Afin de trouver la valeur de α la plus adéquate, nous avons étudié la reconstruction sur un lot de simulation Monte-Carlo à courant chargé avec uniquement une coupure sur la présence d'un vertex reconstruit dans l'événement : $|z_{vtx}| < 35$ cm. En utilisant les courants chargés, nous pouvons nous assurer que seules des particules hadroniques forment l'état final.

Ensuite, nous avons considéré la quantité appelée P_T^{Bal} définie comme suit :

$$P_T^{Bal} = \frac{P_T^{Had}}{P_T^{\nu}} \tag{4.5}$$

où P_T^{Had} est l'impulsion transverse reconstruite du système hadronique et P_T^{ν} l'impulsion transverse du neutrino généré du processus.

Pour un calorimètre parfait et une reconstruction parfaite des hadrons, la quantité P_T^{Bal} vaut 1.0 car le système hadronique équilibre le neutrino dans le plan transverse. La distribution (a) de la figure 4.8 montre l'évolution de la valeur moyenne de P_T^{Bal} en fonction de P_T^{ν}, obtenue par ajustement gaussien. Nous observons d'une part que P_T^{Bal} est proche de la valeur optimale de 1.0 à partir de $P_T^{\nu} = 10$ GeV. De plus, nous remarquons que P_T^{Bal} augmente avec α. Ce comportement est dû au

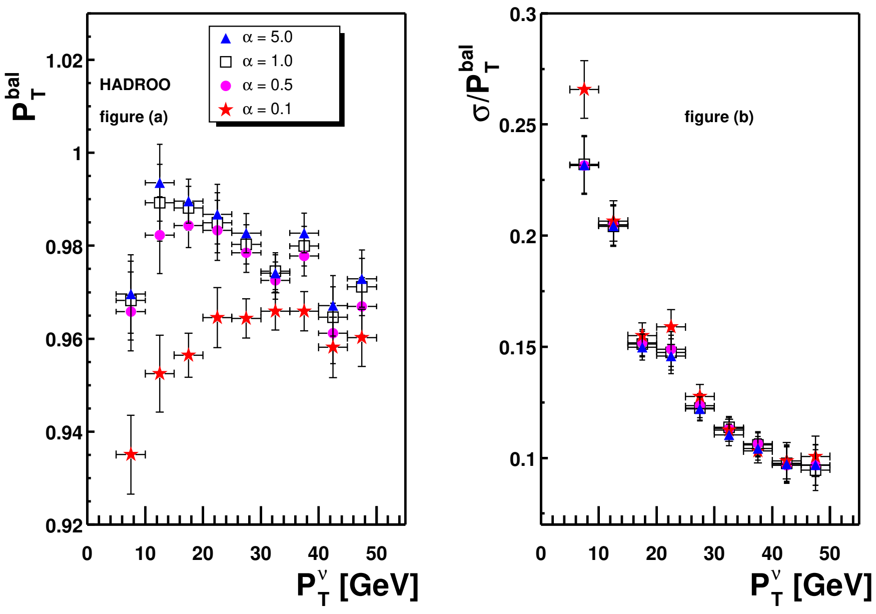


FIG. 4.8 – Distributions de P_T^{Bal} (figure (a)) et de la résolution relative (figure (b)) en fonction de P_T^{ν} pour un lot de simulations Monte-Carlo à courant chargé à $Q^2 > 4\ GeV^2$ avec l'algorithme *HADROO* en fonction des différentes valeurs du paramètre α. Le paramètre N_σ est fixé à 2,0.

fait que lorsqu'on augmente α, on augmente le nombre de traces considérées comme bien mesurées. Ce phénomène est atténué pour des valeurs de $\alpha > 1.0$. La figure (b) de 4.8 permet de choisir la valeur optimale pour α. C'est celle qui va donner la meilleure résolution relative, définie par :

$$\sigma_{Res} = \frac{\sigma_{P_T^{Bal}}}{P_T^{Bal}} \tag{4.6}$$

Nous remarquons immédiatement que les différences en résolution entre $\alpha =$ 0,5 et $\alpha =$ 5,0 sont comprises à l'intérieur des barres d'erreur de l'ajustement. L'algorithme avec $\alpha =$ 0,1 donne des résolutions relatives moins bonnes, à bas P_T, car moins de traces sont considérées. En conséquence, une valeur de α entre 0,5 et 5,0 semble optimale. Naturellement, nous avons alors choisi la valeur de 1,0 pour les études comparatives suivantes.

Le second paramètre, N_σ, décrit la largeur de la fenêtre centrée sur l'énergie calorimétrique pour laquelle la mesure de l'énergie de la trace est considéré comme compatible avec celle du calorimètre. Contrairement au paramètre α, qui modifie drastiquement le comportement de l'algorithme, l'impact de N_σ est plus petit. La distribution (a) de la figure 4.9, obtenue sur un second lot de courant chargé indépendant montre l'évolution de P_T^{Bal} en fonction de P_T^{ν} obtenue par ajustement gaussien. Les valeurs P_T^{Bal} sont identiques. Nous choisirons par la suite la valeur de N_σ à 2,0.

Afin de s'assurer que l'algorithme améliore bien les résolutions relatives nous avons

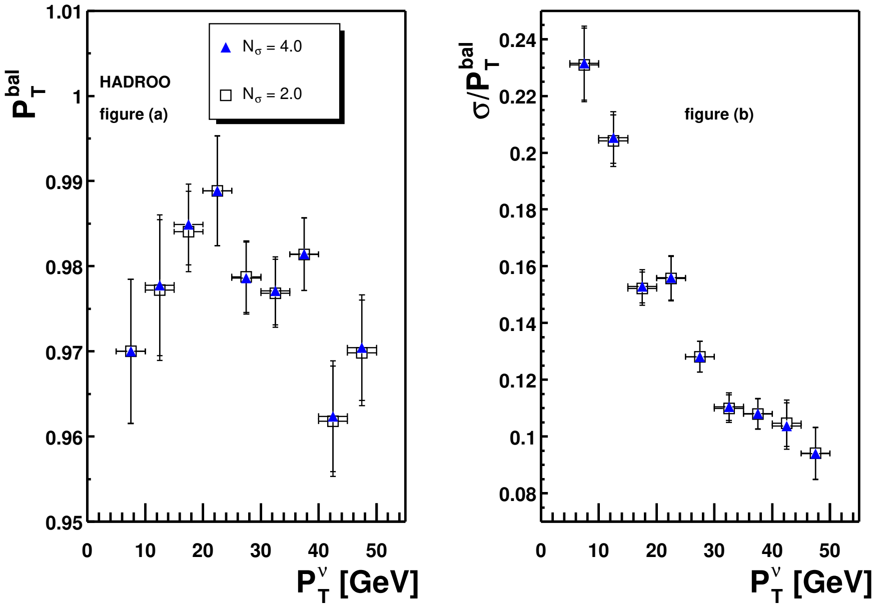


FIG. 4.9 – Distributions de P_T^{Bal} (figure (a)) et de la résolution relative (figure (b)) en fonction de P_T^ν pour un lot de simulations Monte-Carlo à courant chargé à $Q^2 > 4\ GeV^2$ avec l'algorithme $HADROO$ en fonction de deux valeurs du paramètre N_σ. Le paramètre α est fixé à 1.

comparé $HADROO$ à l'algorithme combiné $FSCOMB$ et à l'algorithme purement calorimétrique.

La distribution (a) de la figure 4.10 montre clairement les différences de comportement des algorithmes. Le comportement de l'algorithme $CLUSTER$ se détériore notablement à bas P_T, où le calorimètre n'a pas une réponse optimale. Le comportement est plus stable pour l'algorithme combiné $FSCOMB$ pour lequel les valeurs de P_T^{Bal} ne varient qu'entre 0,91 et 0,97. De plus, elles sont plus proches de la valeur optimale de 1,0. Finalement, nous remarquons des valeurs comprises entre 0,97 et 0,99 pour $HADROO$. Les résolutions relatives de la figure (b) montrent que les algorithmes combinés n'améliorent vraiment la résolution relative que pour des valeurs de P_T^ν jusqu'à 20 GeV. Au-delà, les différences sont comprises dans les barres d'erreur. Nous observons néanmoins que pour des P_T^ν jusqu'à 20 GeV, $HADROO$ améliore légèrement la résolution par rapport à $FSCOMB$. Au-delà, c'est-à-dire à grand P_T, nous retrouvons pour les algorithmes combinés des résolutions identiques à l'algorithme purement calorimétrique.

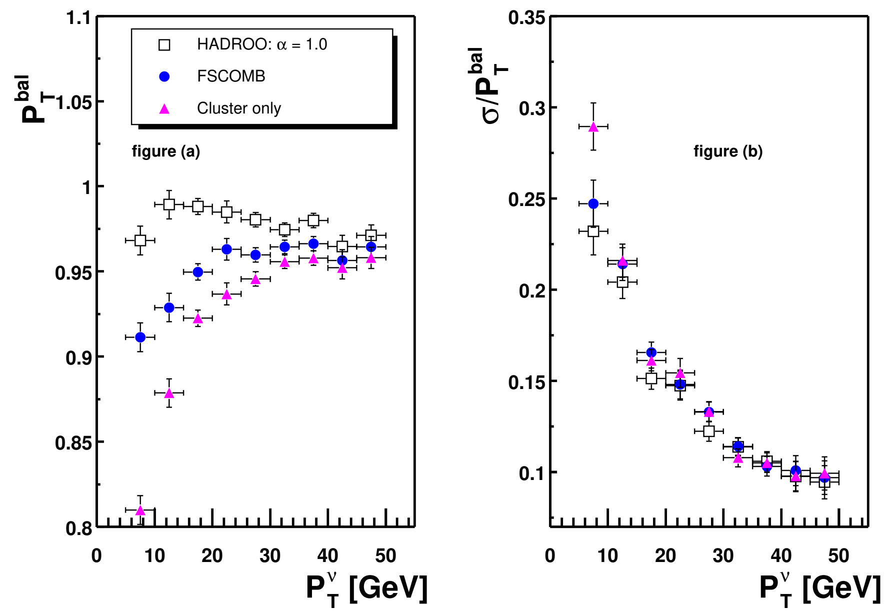


FIG. 4.10 – Distributions de P_T^{Bal} (figure (a)) et de la résolution relative (figure (b)) en fonction de P_T^{ν} pour un lot de simulations Monte-Carlo à courant chargé à $Q^2 > 4\ GeV^2$. Les paramètres α et N_σ sont fixés à 1 et 2,0 respectivement.

La figure 4.11 en fonction de P_T^{ν} montre la décomposition en contributions des traces et clusters. Les deux distributions du haut montrent que le comportement de *HADROO* pour les valeurs de 1,0 et de 0,5 de α est similaire : à bas P_T, 45% du P_T^{Had} provient des traces et 55% des clusters. Pour des P_T^{ν} atteignant 50 GeV, plus de 65% proviennent de clusters et moins de 35% de traces. Dans *FSCOMB*, par contre, la contribution des traces pour P_T^{ν} compris entre 5 et 10 GeV n'est que de 25%. A grand P_T^{ν}, la contribution des traces diminue encore, car seules les traces d'impulsion transverse en-dessous de 2 GeV sont gardées. Pour l'algorithme *Cluster*, seul les clusters contribuent.

Conclusions et perspectives

Nous avons vu que l'algorithme *HADROO*, avec des valeurs de α de l'ordre de 1, améliore sensiblement la mesure de l'énergie totale hadronique et que la résolution est améliorée pour des P_T jusqu'à 20 GeV par rapport aux algorithmes *Cluster* et *FSCOMB* utilisés jusqu'à présent. Au-delà, le comportement de *HADROO* est similaire à l'algorithme purement calorimétrique. Nous avons vu aussi, que *HADROO* garde plus de traces en tant qu'objets hadroniques.

Afin de déterminer la valeur de α la plus optimale aussi bien pour les analyses de physique à grand P_T que pour les analyses à plus bas P_T, des études supplémentaires sont effectuées en ce moment. Pour sonder le comportement de l'algorithme à bas

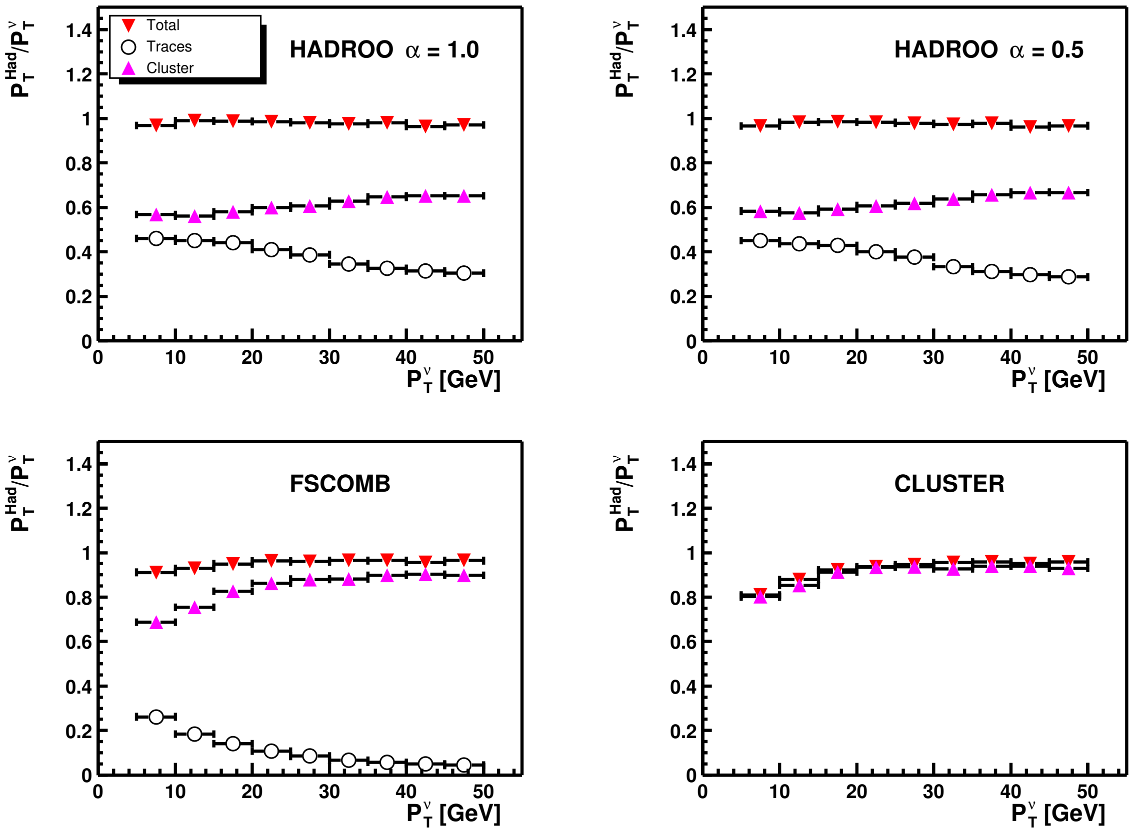


FIG. 4.11 – Distributions de P_T^{Bal} et des contributions individuelles des types d'objets hadroniques en fonction de P_T^{ν} pour un lot de simulations Monte-Carlo à courant chargé à $Q^2 > 4\ GeV^2$.

P_T, des lots contenant des saveurs lourdes dans l'état final sont utilisés. Une valeur de $\alpha = 0{,}5$ semble être une valeur optimale. En même temps, un certain nombre d'améliorations sont envisageables. Dans l'algorithme $HADROO$, aucune hypothèse sur le type de particule n'est faite. L'énergie associée à chaque dépôt calorimétrique est celle obtenue après correction à l'échelle hadronique. Donc, l'énergie de pions neutres reconstruits en tant que hadrons par exemple est un peu surestimée. Il est donc envisageable d'utiliser l'échelle d'énergie correspondant au type de particule suivant la fraction électromagnétique du dépôt calorimétrique. En ce moment, un gros effort est fourni pour améliorer la mesure de l'échelle d'énergie surtout à bas P_T. Ces modifications sont en cours d'études, et peuvent à terme améliorer encore sensiblement la mesure de l'énergie des hadrons à bas P_T.

4.5 Reconstruction des jets

L'identificateur de jets fait partie de la catégorie des identificateurs de particules composées décrits dans la partie 3.3. Ces identificateurs reconstituent des particules composées à partir de particules simples déjà reconstruites auparavant. L'identificateur de jets utilise les hadrons reconstruits pour les associer en jets. Lors de la première partie de ma thèse, j'ai implémenté la structure de reconstruction des jets en C++, ainsi que l'algorithme de K_T décrit ci-dessous. Nous avons choisi l'algorithme K_T pour plusieurs raisons :

- Il a été montré dans [61] que l'algorithme de K_T est l'algorithme le plus adéquat pour minimiser des effets d'hadronisation à longue distance entre les partons.
- De plus, contrairement aux algorithmes de cone, les jets ne se chevauchent pas.

4.5.1 L'algorithme K_T

L'agrégation des particules en jet est faite en fonction d'un paramètre d'approche minimale pour toutes les paires de particules qui associe sucessivement les particules les plus proches jusqu'à atteindre une condition limite.
Pour l'algorithme de K_T, le paramètre d'approche est l'impulsion transverse relative entre deux particules.

Schéma de reconstruction

L'algorithme K_T a été codé avec un schéma de reconstruction pondéré par l'impulsion transverse. Ce schéma définit le quadri-vecteur de particules filles associées ensemble en un jet selon :

$$\begin{aligned} E_T^{jet} &= \Sigma_i \; E_{T,i} \\ \eta^{jet} &= \frac{\Sigma_i \; E_{T,i} \; \eta_i}{\Sigma_i \; E_{T,i}} \\ \phi^{jet} &= \frac{\Sigma_i \; E_{T,i} \; \phi_i}{\Sigma_i \; E_{T,i}} \end{aligned} \tag{4.7}$$

avec i = 1..N particules.
L'impulsion transverse P_T^{jet} et l'énergie transverse E_T^{jet} sont équivalentes car les jets sont reconstruits avec une masse nulle.

L'algorithme

La reconstruction suit les pas suivants :

- La reconstruction commence avec une liste de particules et une liste de jets vides.
- La distance d_{ij} est calculée pour chaque paire de particules et d_i pour chaque particule individuelle, avec

$$d_i = E_{t,i}^2 \qquad d_{ij} = \min\,(E_{T,i}^2, E_{T,j}^2) R_{ij}^2 / R_0^2 \qquad R_{ij}^2 = \Delta\eta_{ij}^2 + \Delta\phi_{ij}^2$$

et R_0 constante

- d_{min} est définie comme la plus petite valeur entre d_i et d_{ij}
- Si d_{min} appartient à l'ensemble des $\{d_{ij}\}$, les particules i et j sont associées en une nouvelle pseudoparticle qui est ajoutée à la liste des particules (i et j sont effacés de la liste des particules)
- Si d_{min} appartient à l'ensemble des $\{d_i\}$, la particule est effacée de la liste des particules et ajoutée à la liste des jets
- La procédure prend fin lorsque la liste des particules ne contient plus de particules. A ce stade, toutes les particules ont été attribuées à un jet.

Cette définition des jets implique que des particules avec $R_{ij} < R_0$ sont associées, et donc que tous les jets finaux sont éloignés les uns des autres d'au moins R_0. Plus de détails peuvent être trouvés dans [7]. Lors de l'implémentation on a défini la variable $R_0 = 1.0$. Afin de limiter la taille des fichiers *ODS*, nous avons choisi de stocker uniquement des jets ayant une impulsion transverse supérieure à $P_{T,min}^{jet} = 2.5$ GeV, selon un ordre décroissant en P_T. $P_T^{jet\ 1}$ désigne alors le jet de plus grande impulsion transverse.

Dans l'implémentation, une attention particulière a été portée à la rapidité de l'algorithme. Il identifie les jets à partir de matrices contenant la distance entre toutes les particules i et j. Le rang de celle-ci est donc N_{part} pour un événement à N_{part} particules. L'algorithme initial était très lent pour les deux raisons suivantes :

- L'utilisation massive d'objets C++ complexes qui doivent être alloués et effacés pour chaque événement ralentit la reconstruction en général.
- Pour chaque pas de l'algorithme, les distances entre les N_{part}^2 particules étaient recalculées.

Les deux solutions implémentées sont les suivantes :

- Seuls les quadri-vecteurs des particules sont utilisés et stockés sous forme de tableau. Cette information est extraite des particules au début de l'algorithme.
- Seules les distances modifiées entre particules sont recalculées. Au lieu de recalculer de l'ordre de N_{part}^2 éléments de matrice, seuls $\sim N_{part}$ sont recalculés à chaque étape de l'algorithme.

4.5.2 Etudes de validation et de comparaisons

Afin de valider l'algorithme, nous avons choisi de mener deux études différentes :

- Un premier test montrant l'accord entre le niveau reconstruit et le niveau généré des jets.
- Un second test vérifiant l'accord entre les algorithmes C++ et fortran.

Dans un premier temps, nous avons comparé les jets reconstruits aux jets générés. Nous avons utilisé un lot de simulations Monte-Carlo de courant neutre de $Q^2 > 100\ GeV^2$ en sélectionnant au moins un jet et un électron au niveau généré et reconstruit d'impulsion transverse supérieure à 20 GeV :

- $|z_{vtx}| < 40$ cm
- $N_{ele,jet}^{rec} \geq 1$, $N_{ele,jet}^{gen} \geq 1$

– $P_T^{jet,ele\ rec} > 20$ GeV, $P_T^{jet,ele\ gen} > 20$ GeV
– $10^o < \theta^{rec}_{jet,ele} < 150^o$, $10^o < \theta^{gen}_{jet,ele} < 150^o$

Les critères requis pour l'électron et le jet sont les critères de sélection serrés de l'analyse, détaillés dans les parties respectives 5.1.1 et 5.1.4. Ce lot nous permet d'estimer les résolutions angulaires et en énergie sur le jet reconstruit. La figure 4.12 montre que la résolution est à peu près de 5 GeV pour l'énergie (figure (a)), de 0,5^o pour l'angle polaire (figure (b)) et de 1,2^o pour l'angle azimutal (figure (c)) du jet.

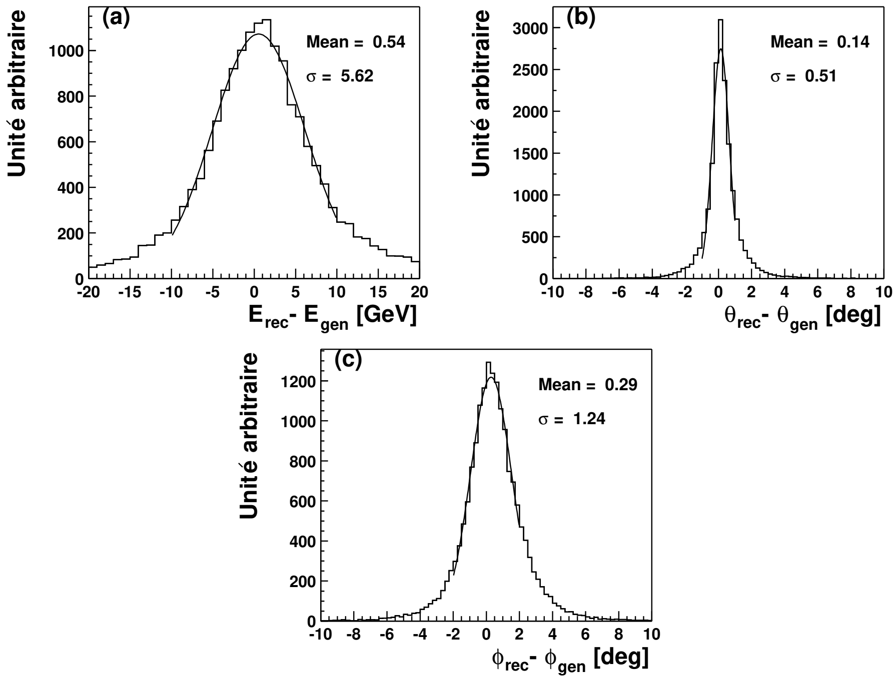


FIG. 4.12 – Résolutions en énergie (a), angle polaire (b) et azimutal (c) du jet pour un lot de courants neutres de $Q^2 > 100\ GeV^2$.

Pour la comparaison avec l'algorithme fortran, nous avons choisi de travailler sur un lot de courants chargés à $Q^2 >\ 4\ GeV^2$. Afin de pouvoir étudier les erreurs possibles de reconstruction de jets, nous avons généré un lot avec une reconstruction de hadrons avec $FSCOMB$ en fortran et en C++, suivie d'une reconstruction des jets en fortran ou C++. Nous avons alors sélectionné des événements contenant au moins un jet dans l'acceptance du calorimètre LAr suivant :

– $N_{jet,rec} \geq 1$
– $P_{T,jet1} \geq 10$ GeV
– $\theta_{jet1} \geq 10^o$

où $jet1$ désigne le jet de plus grande impulsion transverse. En utilisant ces coupures sur un lot de de simulations Monte-Carlo courant chargé, nous obtenons 2397 événements sélectionnés pour le lot C++ et 2412 en fortran.

Les distributions de P_T, des angles polaires et azimutaux de la figure 4.13 montrent un très bon accord entre les codes fortran et C++. Néanmoins, nous remarquons un léger déficit de 10% de jets reconstruits en C++ pour un P_T compris entre 7,5 et 15 GeV et au-delà de 90 GeV. Un traitement différent de cellules bruyantes entre les deux algorithmes pourrait expliquer ces différences. Des études comparatives sont en cours.

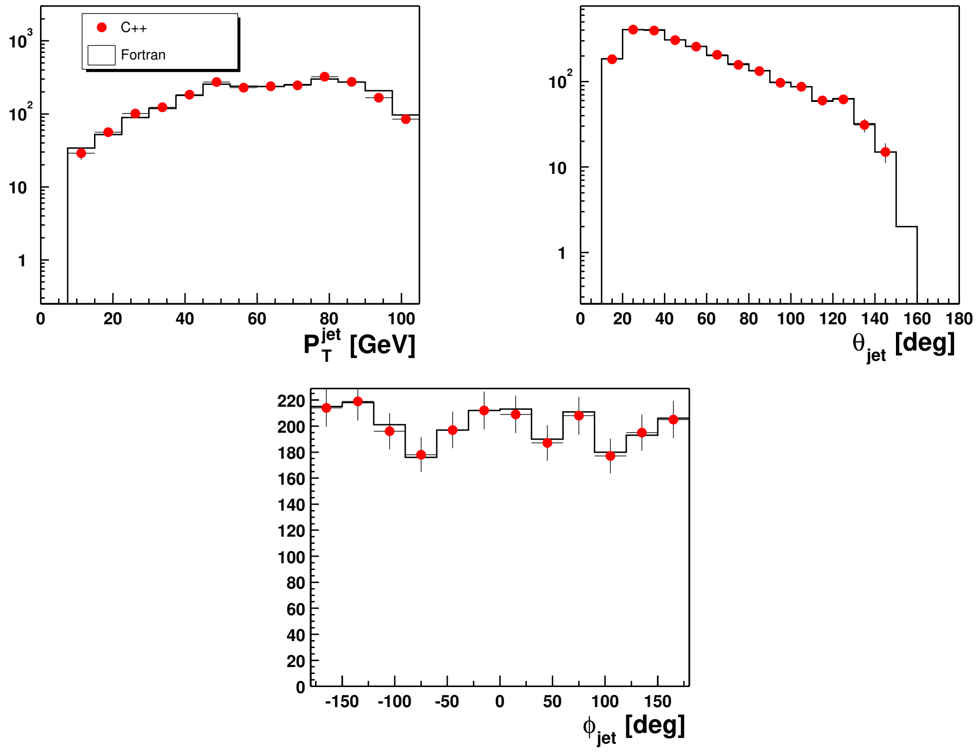


FIG. 4.13 – Distributions en impulsion transverse, angle polaire et azimutal du jet de plus grand P_T pour un lot de simulations Monte-Carlo courant chargé à $Q^2 > 4$ GeV^2.

Dans notre analyse finale, décrite dans le chapitre 5, nous allons sélectionner des jets hadroniques dans la liste des jets reconstruits sur μODS, en utilisant en plus les estimateurs de fraction électromagnétique et de taille dans le plan transverse.

D'après la figure 4.14, l'allure de ces estimateurs de l'algorithme fortran est bien reproduite par l'algorithme en C++.

Nous avons choisi, dans un premier temps, d'utiliser pour l'analyse finale du chapitre 5 les jets reconstruits avec l'algorithme *FSCOMB*, car les constantes de calibration hadronique et de calibration des jets ont été précisément mesurées dans des analyses antérieures.

4.5.2.1 Calibration de l'énergie des jets

De la même manière que pour l'électron, nous pouvons vérifier la calibration des jets et donc des objets hadroniques qui les constituent. La calibration implémentée repose sur une calibration des objets hadroniques calorimétriques roue par roue

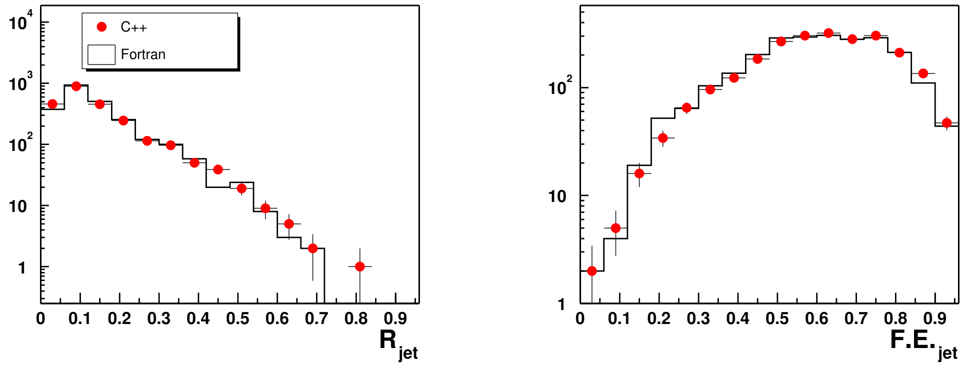


FIG. 4.14 – Distributions de taille transverse R_{jet} et de fraction électromagnétique $F.E._{jet}$ du jet de plus grand P_T pour un lot de simulations Monte-Carlo courant chargé à $Q^2 > 4$ GeV^2.

complétée par des facteurs correctifs dépendant de l'énergie et de l'angle polaire du jet appliqués aux jets mêmes [2]. Pour vérifier l'accord final, nous avons constitué un lot de contrôle de courants neutres sélectionnés de la façon suivante :

- Coupures de qualité :
  - Utilisation des identificateurs de bruit de fond décrits dans [56, 57]
  - $|z_{vtx}| < 35$ cm
- $43 < E - P_z < 67$ GeV
- Un électron bien identifié
- $P_T^{had} > 15$ GeV et un seul jet reconstruit
- $10^o < \theta_{ele,jet} < 150^o$
- $P_T^{ele} > 15$ GeV

où l'électron a été sélectionné avec les critères décrits dans la partie 5.1.1. Nous avons sélectionné des échantillons correspondant aux données e^+p 94-97, e^-p 98-99 et e^+p 99-00 et étudié le rapport P_T^{Jet}/P_T^{DA}, qui devrait être centré sur 1.0 pour des jets parfaitement calibrés. Les distributions pour l'échantillon e^+p 99-00 de la figure 4.15 nous montrent que pour des P_T^{Jet} compris entre 20 et 80 GeV et pour l'acceptance du calorimètre LAr ($10^o < \theta < 150^o$), les écarts des énergies mesurés par rapport à la valeur optimale sont toujours inférieurs à 4% (intervalle délimité par les bandes horizontales).

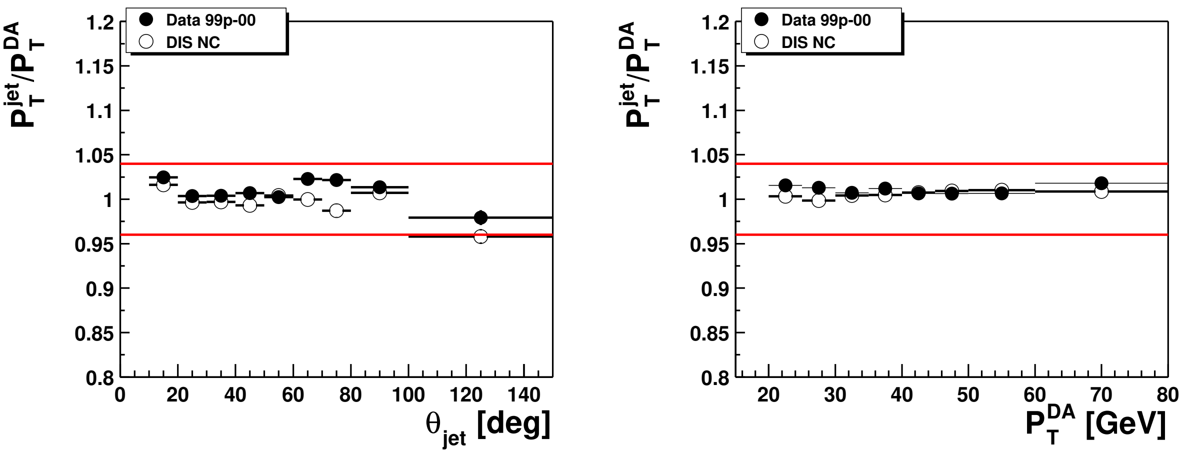


FIG. 4.15 – Distribution de P_T^{Jet}/P_T^{DA} en fonction de θ_{jet} et de P_T^{DA}.

Afin d'estimer l'accord entre les données et le Monte-Carlo, nous avons étudié le double rapport P_T^{Jet}/P_T^{DA} (données/MC). La valeur moyenne de ce double rapport est présentée sur les figures 4.16 en fonction de θ_{jet} et P_T^{DA}, pour les trois périodes de prises de données analysées. Nous voyons que, à part à grand θ_{jet} pour les données e^-p 98-99, les données et les simulations Monte-Carlo sont en accord à mieux de 4% près. Cette valeur sera retenue comme incertitude systématique globale sur la calibration des jets.

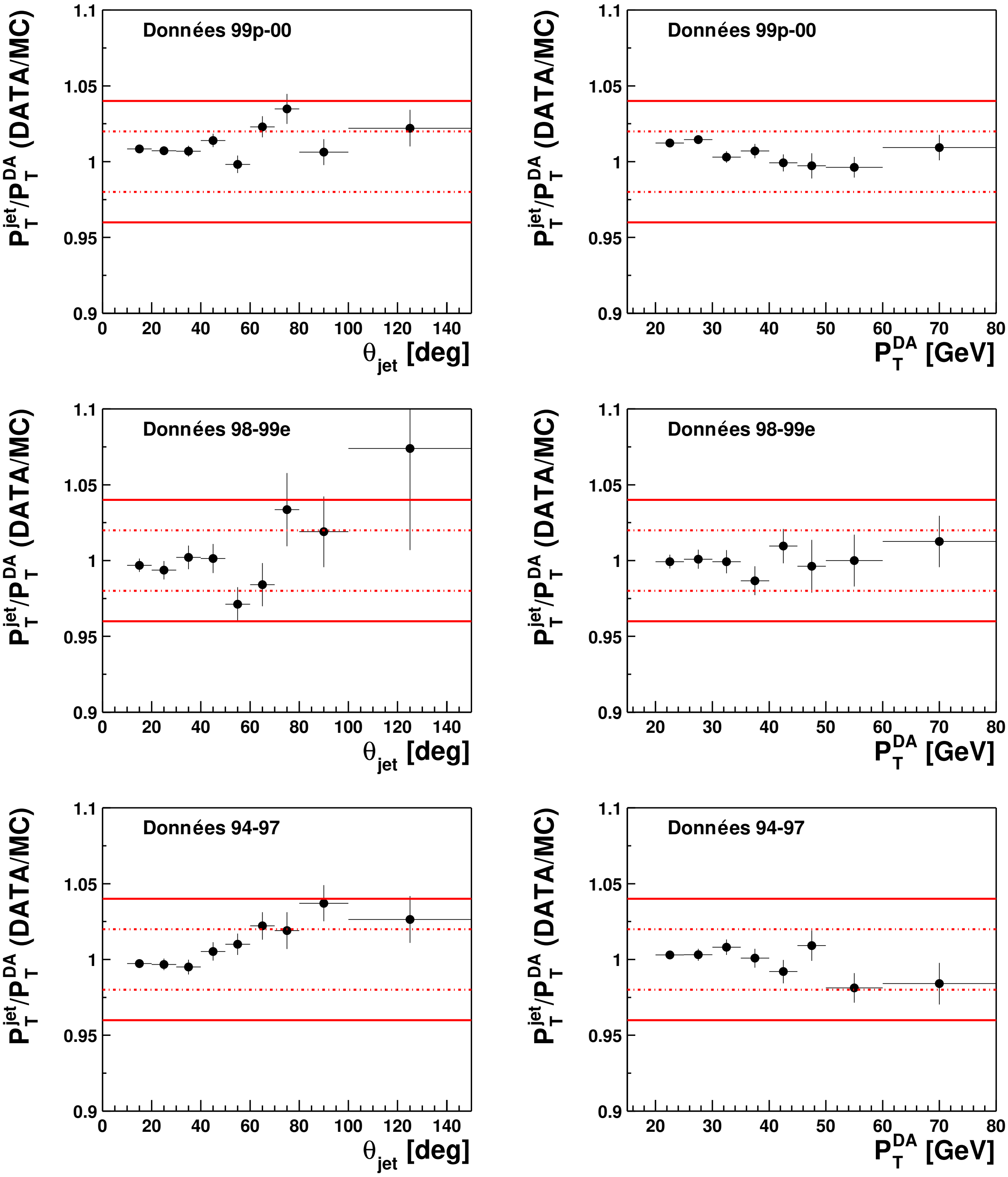


FIG. 4.16 – Distribution de la valeur moyenne du double rapport P_T^{Jet}/P_T^{DA} (données/MC) en fonction de θ_{jet} et de P_T^{DA} (données e^+p 99-00, e^-p 98-99 et e^+p 94-97).

Conclusion

Au cours des deux chapitres précédents, nous avons vu qu'un nouvel environnement orienté objet a été développé au sein de la collaboration. La structure finale a été présentée ainsi que les performances et contraintes liées à son développement. Les informations sur les particules nécessaires à la plupart des analyses sont maintenant disponibles de manière officielle et standard. De plus, afin de réduire la période d'apprentissage pour les nouveaux venus, un effort a été fourni au niveau de la documentation.

Nous avons également présenté les principaux identificateurs de particules. Trois algorithmes de reconstruction hadronique ont été implémentés et validés, dont un nouveau, $HADROO$. Les performances de celui-ci montrent une amélioration de la résolution en énergie de l'état hadronique final à bas P_T. Un algorithme de reconstruction de jet K_T a également été implémenté et validé. Les tests des identificateurs à grand P_T, domaine qui nous intéresse par la suite dans l'analyse, a permis de mettre en évidence les erreurs d'implémentation et de les corriger. Maintenant, les performances physiques des algorithmes sont équivalentes au fortran et les algorithmes d'identification de hadrons et de jets développés sont intégrés dans la reconstruction standard.

Ainsi, le but de la première étape de développement des différents identificateurs qui était d'obtenir des algorithmes au moins aussi performants en terme de résultats de physique que les anciens programmes fortran a été rempli : ces algorithmes sont déjà utilisés dans les analyses physiques et les contrôles de qualité des nouvelles données de la phase II de HERA. Cet environnement servira de base pour des futurs améliorations ainsi que pour le développement de nouveaux algorithmes. Ainsi, de nouvelles optimisations de l'identificateur de hadrons par exemple sont en cours afin d'essayer de gagner encore davantage sur la résolution en énergie, principalement à bas P_T. De même, une première version d'un identificateur de lepton tau est sur le point d'être intégrée.

Nous allons pouvoir utiliser ces nouveaux outils dans le cadre de notre analyse de physique qui fera l'objet des chapitres suivants.

Chapitre 5

Etude des topologies à grand E_T

Si non è vero, è ben trovato.

Giordano Bruno
Gli eroici furori

De nombreux canaux ont déjà fait l'objet d'études à grande impulsion transverse dans des analyses dédiées : la recherche de leptons isolés [16], l'analyse multi-électrons [11] et les études de jets dans les processus de photoproduction [8] ou DIS [4] n'en sont que des exemples.
Cette analyse tente une approche plus générale en étudiant toutes les topologies possibles à grande énergie transverse, suivant un principe proche des analyses [27, 28, 29]. Le but est de rechercher des phénomènes au-delà du Modèle Standard dans une analyse qui soit :

- générique, c'est-à-dire avec la même définition de particules pour tous les canaux étudiés
- indépendante des modèles de nouvelle physique
- sensible à plusieurs modèles prédisant des phénomènes au-delà du Modèle Standard

Dans ce chapitre, nous décrivons l'analyse développée pour étudier les topologies à grande énergie transverse de manière générique sur l'ensemble des données prises sur H1 pendant la phase I de HERA. Après une brève description des particules recherchées, nous discutons leurs critères de sélection et leur efficacité d'identification. Ensuite, nous décrivons les effets expérimentaux, comme par exemple les efficacités liées au système de déclenchement, dont il faut tenir compte, et déterminons les erreurs systématiques et statistiques. Nous présentons finalement les résultats en comparant les données aux prédictions du Modèle Standard. Cette analyse utilise les algorithmes et outils de reconstruction que nous avons développés en langage orienté objet et qui ont été décrits dans les chapitres 3 et 4.

5.1 Définition des objets physiques

Nous recherchons les topologies définies à partir des cinq types de particules suivantes :

- L'électron (abrégé e)
- Le muon (abrégé μ)
- Le jet (abrégé j)
- Le neutrino (abrégé ν)
- Le photon (abrégé γ)

Par la suite, nous utiliserons l'appellation “corps” pour désigner des particules, parmi celles définies ci-dessus, qui sont isolées et reconstruites avec une énergie transverse importante. Les critères précis sont décrits dans les parties dédiées aux corps respectifs. Le potentiel de découverte dépend à la fois de notre capacité à reconstruire les corps recherchés et des erreurs d'identification. Nous avons donc étudié, d'une part l'efficacité de sélection individuelle de chaque corps, et d'autre part la contamination dans chaque canal venant d'une mauvaise reconstruction de l'un des corps. L'objectif de pureté conduit à appliquer une sélection plus stricte sur les objets physiques que celle décrite dans le chapitre 4. Nous identifierons les corps dans la région centrale du détecteur, là où l'identification est la plus fiable. Pour la détermination des critères, nous nous sommes appuyés sur les analyses précédentes de H1. L'ensemble des critères de sélection utilisés sont résumés dans le tableau 5.1. Les notations des variables utilisées pour décrire les coupures sont celles introduites dans la partie 4.1.2.

5.1.1 Critères du corps électron

L'efficacité de reconstruction de l'électron dépend fortement de l'efficacité d'associer une trace à l'amas électromagnétique du calorimètre, sachant que l'efficacité de reconstruction d'un amas électromagnétique est proche de 100 %. Ainsi, suivant les critères développés au cours de l'analyse multi-électrons [11], nous avons séparé la reconstruction de l'électron en deux parties distinctes : la partie avant avec un angle polaire θ contenu entre $10^o < \theta < 20^o$ et une partie centrale avec $20^o \leq \theta < 150^o$ (θ étant reconstruit à partir de l'amas électromagnétique). De façon générale, nous demandons un amas électromagnétique isolé dans le calorimètre LAr avec les critères décrits dans le chapitre 4.2.1. Pour distinguer l'électron du photon, nous recherchons les traces associées à l'amas électromagnétique en calculant la distance de l'amas par rapport à la direction des traces au point d'impact dans le calorimètre. Cette distance doit être inférieure à 12 cm pour un électron.
A l'avant, on demande que l'amas électromagnétique soit associé à au moins une trace ajustée au point d'interaction (notée $DTRA$).
Dans la partie centrale, où l'efficacité de reconstruction des trajectographes est plus élevée, on demande l'association à une seule trace, isolée dans un cône de 0,5 ra-

Electron général :	Isolation calorimétrique : $E_{cone}^{ele} \leq 3$ % $E_{electron}$
$20^o < \theta_{amas} < 150^o$:	1 particule chargée isolée
	pointant vers l'amas ($DCA < 12$ cm)
	$R_{start}^{trace} < 30$ cm
	$1/P_T^{trace} - 1/P_T^{amas} < 0,02$
$10° < \theta_{amas} \leq 20°$:	$\geq$1 trace ajustée au vertex
	pointant vers l'amas ($DCA < 12$ cm)
Photon	Isolation calorimétrique : $E_{cone}^{ele} \leq 3$ % $E_{electron}$
	Pas de traces (DTNV/DTRA) pointant vers l'amas
	($DCA < 12$ cm)
	Aucune particule chargée avec $D_{trace}^{amas} < 0{,}5$
Muon	Grade 1-3
	Isolation par rapport aux traces et jets :
	$D_{muon}^{traces\ DTRA} > 0{,}5$
	$D_{muon}^{jet} > 1{,}0$
Jet général :	$R_{jet} > 0{,}1$ ou fraction électromagnétique $< 0{,}9$
$20^o < \theta_{jet} < 150^o$:	$N_{tracesDTRA}^{jet} \geq 1$
	($N_{tracesDTRA}^{jet} \geq 2$ dans les interstices en Φ du détecteur)
Neutrino	Défini seulement si $(E - P_z) \leq 55{,}2$ GeV

TAB. 5.1 – Résumé des critères de sélection pour chaque corps recherché.

dian dans le plan (η-ϕ), qui doit être identifiée comme particule chargée selon les critères définis dans la partie 4.1. Nous imposons des critères supplémentaires dans la partie centrale afin de diminuer le taux d'électrons mal identifiés. Ainsi, la quantité $1/P_T^{trace}$ - $1/P_T^{amas}$, imposant la compatibilité (notée "*Pt match*" sur la figure 5.1) entre la mesure de l'impulsion transverse de la trace et celle de l'amas, doit être petite ($<$ 0,02 GeV^{-1}). De plus, on demande aussi que les dépôts dans le premier trajectographe central (CJC 1) associés à cette trace soient reconstruits à une distance radiale R_{start} inférieure à 30 cm (noté "$R(Start)$" sur la figure 5.1). Ceci permet d'éliminer les reconversions de photon en paire e^+e^- qui donnent naissance à des traces pointant vers des amas électromagnétiques.

Nous définissons, de manière générale pour un corps i, l'efficacité de sélection ϵ_i par la relation suivante :

$$\epsilon_i = \frac{N_i^{sel}}{N_i^{ini}} \tag{5.1}$$

où N_i^{sel} est le nombre d'événements sélectionnés avec les critères serrés du corps i et N_i^{ini} le nombre d'événements initial dans le lot.

L'efficacité de sélection associée à ces critères serrés des électrons, ϵ_{ele}, a été déterminée en sélectionnant un lot de processus à courant neutre avec les coupures suivantes :

- Coupures de qualités :
  - Utilisation des identificateurs de bruit de fond décrits dans [56, 57]

– $|z_{vtx}| < 40$ cm
– $P_T^{had} > 15$ GeV
– $N_{jets} \geq 1$
– $43 < E - P_z < 67$ GeV
– $N_{ele}^{amas} = 1$
– $P_T^{ele} > 20$ GeV
– $10^o < \theta^{ele} < 150^o$

où P_T^{had} est l'impulsion transverse du système hadronique. A partir du lot e^+p 99-00, dont les propriétés sont décrites dans la partie 5.2.3, nous obtenons un lot de 20994 événements pour 22709 ± 227 prédits. L'absence de critères sur les traces associées à l'électron permet, en appliquant successivement les coupures de sélection, de déterminer l'inefficacité associée à chaque critère. La figure 5.1 montre l'efficacité ϵ_{ele} après application des deux derniers critères, en fonction de l'impulsion transverse et de l'angle polaire de l'électron.

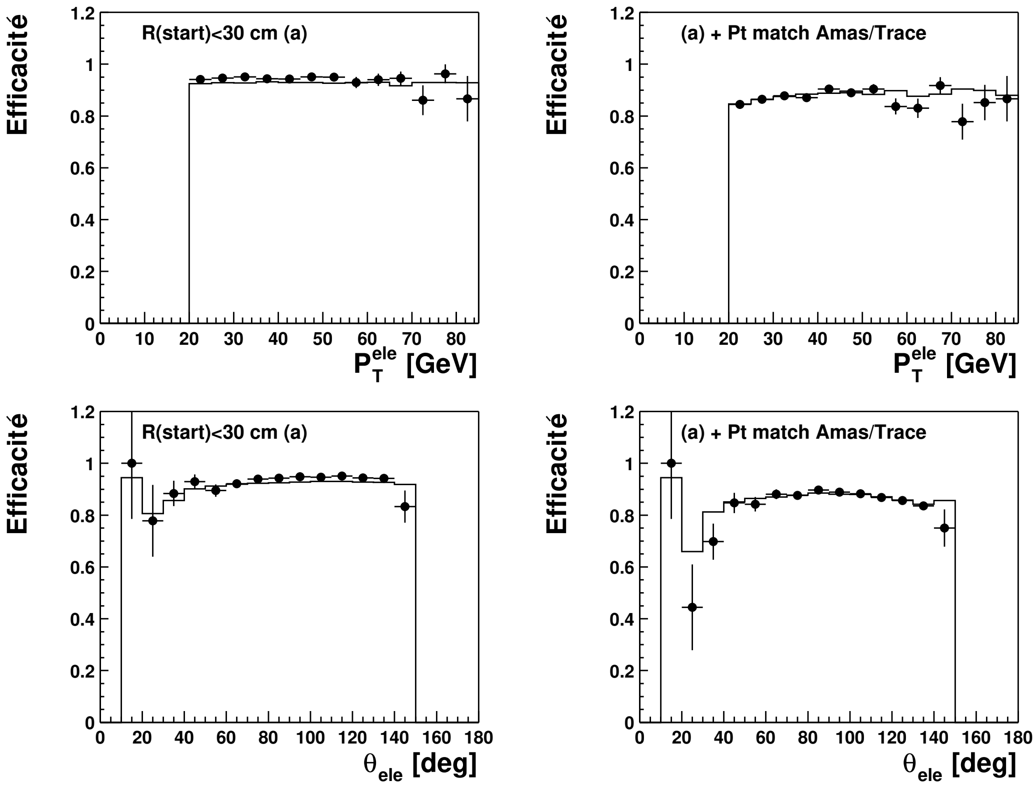


FIG. 5.1 – Efficacité de reconstruction de l'électron avec les coupures finales (figures de droite) en fonction de l'impulsion transverse et de l'angle polaire de l'électron. Les prédictions brutes (histogrammes) ont été rabaissées de 5% pour obtenir un bon accord avec les données (points).

Pour obtenir un bon accord entre les données et les simulations Monte-Carlo, il faut diminuer les efficacités de ces dernières de 5%. Cet effet a été corrigé de la même manière dans l'analyse dédiée aux multi-électrons [11]. Après cet ajustement global, nous remarquons d'après les figures 5.1 une bonne description de l'efficacité entre les données et les simulations Monte-Carlo. Après toutes les coupures, l'efficacité varie de 45% à 90% pour la région centrale ($\theta_{ele} > 20^o$) et vaut 90% dans la partie avant ($\theta_{ele} \leq 20^o$), où l'erreur statistique du lot de données est néanmoins très importante. L'efficacité dépend peu de l'impulsion transverse.

5.1.2 Critères du corps photon

L'efficacité d'identification d'un photon dépend d'une part de l'efficacité de reconstruction d'un amas dans la partie électromagnétique du calorimètre qui est quasiment 100%, et d'autre part de la bonne description des traces et de leur points d'impact dans le calorimètre.
Pour qu'un dépôt électromagnétique soit identifié comme photon, nous requierons qu'aucune trace, ajustée au vertex ($DTRA$) ou non ($DTNV$), ne pointe vers l'amas (noté "0 $DTNVassoc$" sur la figure 5.2). Pour augmenter la pureté d'identification, il ne doit pas y avoir de particule chargée dans un cône d'un rayon de 0,5 radian dans le plan ($\eta - \phi$) en partant du point d'interaction autour du dépôt électromagnétique (noté "0 $Part.\ char.$" sur la figure 5.2). Afin d'estimer l'efficacité d'identification du corps photon, ϵ_γ, nous sélectionnons un lot d'événements Compton élastiques ayant les caractéristiques suivantes :

- Coupures de qualité :
  - Utilisation des identificateurs de bruit de fond décrits dans [56, 57].
  - $|z_{vtx}| < 40$ cm
- $N_{amas} = 2$ dont un électron serré
- $45 < E - P_z < 70$ GeV
- $P_T^{amas1,2} > 15$ GeV , $10^o < \theta^{amas1,2} < 150^o$
- $D_{jet}^{amas1,2} > 0{,}5$
- $E^{amas1} + E^{amas2} > 95\%\ E_{tot}^{calo}$

L'électron serré correspond à la définition du corps électron définit précédement et N_{amas} au nombre d'amas électromagnétiques. $D_{jet}^{amas\ i}$ est défini comme la distance de l'amas i au jet le plus proche dans le plan (η-ϕ). La coupure sur $E^{amas1} + E^{amas2}$ permet de sélectionner un lot d'événements Compton élastiques peu contaminé par les événements de diffusion profondément inélastique. Le lot d'événements ayant été sélectionné sans critère de trace sur le photon, l'application de ces critères permet de mesurer l'efficacité de sélection des photons qui y est associée. Avec cette sélection sur le lot de données $e^{\pm}p$ 94-00, nous observons 265 événements de données pour 298 ± 12 prédits. Nous avons négligé dans la simulation la contribution subsistante de processus de production de paires de leptons, qui vaut 2%.

Sur la figure 5.2 sont présentées les distributions de l'efficacité ϵ_γ en fonction de l'angle polaire et de l'impulsion transverse du photon pour les deux derniers critères appliqués.

Les distributions en fonction de l'angle polaire de la figure 5.2 montrent que l'efficacité d'identification est de l'ordre de 60 % dans la partie avant et à peu près 90% dans la région centrale du détecteur. Nous observons un bon accord entre les données et les simulations dans la limite de la faible statistique. Comme la section

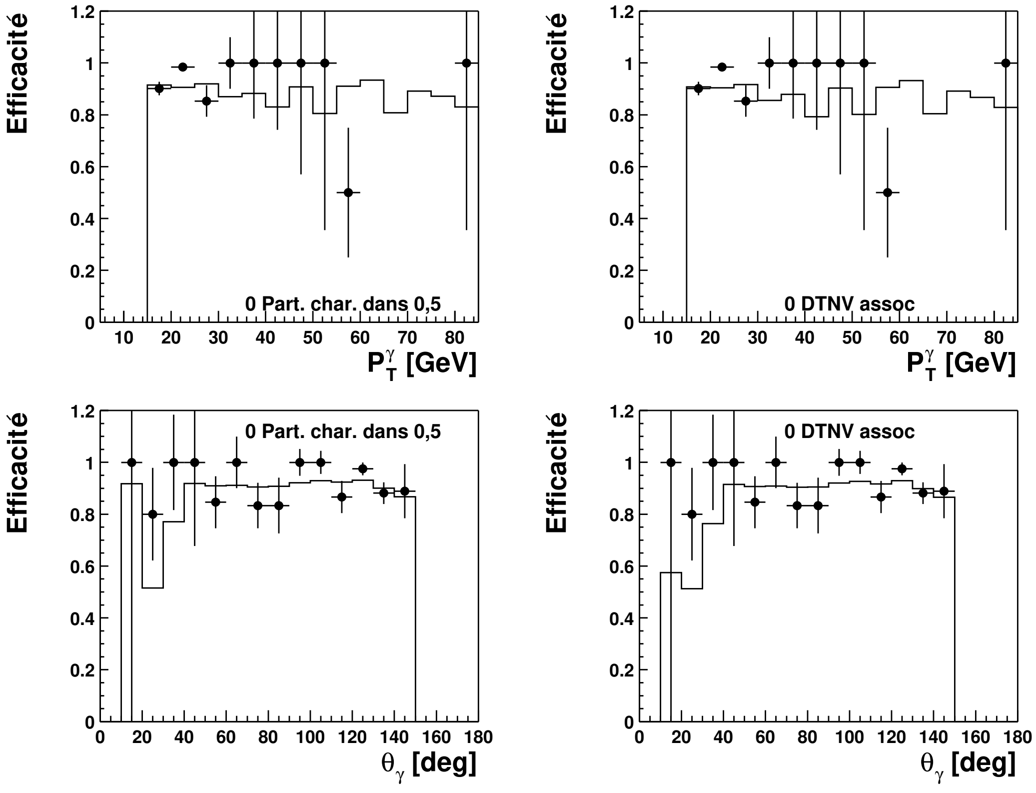


FIG. 5.2 – Efficacité de reconstruction avec les coupures finales du photon en fonction de P_T^γ et θ_γ.

efficace chute rapidement en fonction de l'impulsion transverse, il ne subsiste qu'une trentaine d'événements de données au-delà de 30 GeV en P_T avec les critères finaux : la mesure d'efficacité est dominée dans ce domaine cinématique par les erreurs statistiques.

5.1.3 Critères du corps muon

Les muons sont sélectionnés à partir des cinq critères (ou "grades" en anglais) décrits dans la partie 4.3.1. Par la suite, nous considérons uniquement les muons de grade 1 à 3, donc ceux qui sont bien identifiés comme particule chargée ayant déposé une partie de leur énergie dans le fer instrumenté. De plus, ils doivent être isolés dans le calorimètre suivant la définition donnée au chapitre 4.3.1.
Nous requierons une isolation du muon par rapport à toutes les traces ajustées au vertex dans un cône d'un rayon de 0,5 radian dans le plan $(\eta - \phi)$ autour de sa trace centrale associée. De plus, le muon doit être isolé par rapport aux jets dans le plan $(\eta - \phi)$ autour de la trace à plus de 1,0 radian. Ces restrictions sont nécessaires, car dans les événements contenant des jets, on reconstruit souvent des muons issus de la désintégration de quarks b ou de désintégration de hadrons. De la même manière que pour les photons, nous pouvons déterminer l'efficacité de reconstruction du muon

ϵ_μ en utilisant un lot de paires de muons élastiques. Le but est de sélectionner un lot élastique contenant deux particules chargées, dont au moins un muon, et de déterminer l'efficacité d'identification du second muon. Nous avons sélectionné un lot sur les données e^-p 98-99 et e^+p 99-00 avec les coupures suivantes :

- Coupures de qualité :
  - Utilisation des identificateurs de bruit de fond décrits dans [56, 57].
  - Utilisation des déclencheurs pertinents décrits dans la partie 5.2.4.
  - $|z_{vtx}| < 40$ cm
- $N_{part.\ char.} = 2$
- $N_{ele} = 0$
- $E_{calo} < 2$ GeV
- $N_\mu \geq 1$
- $P_T^{part.\ char.} > 5$ GeV, $10^o < \theta_{part.\ char.} < 150^o$
- $\alpha_{part.\ char.1}^{part.\ char.2} < 165^o$

où E_{calo} est l'énergie calorimétrique qui n'est pas associée aux muons. L'indice "*part.char.*" désigne les particules chargées et $\alpha_{part.\ char.1}^{part.\ char.2}$ l'angle d'ouverture entre les deux particules chargées sélectionnées. La coupure sur α permet, comme cela sera montré dans la partie 5.3.1, de rejeter le bruit de fond des muons cosmiques. La coupure sur E_{calo} permet de sélectionner un lot élastique.

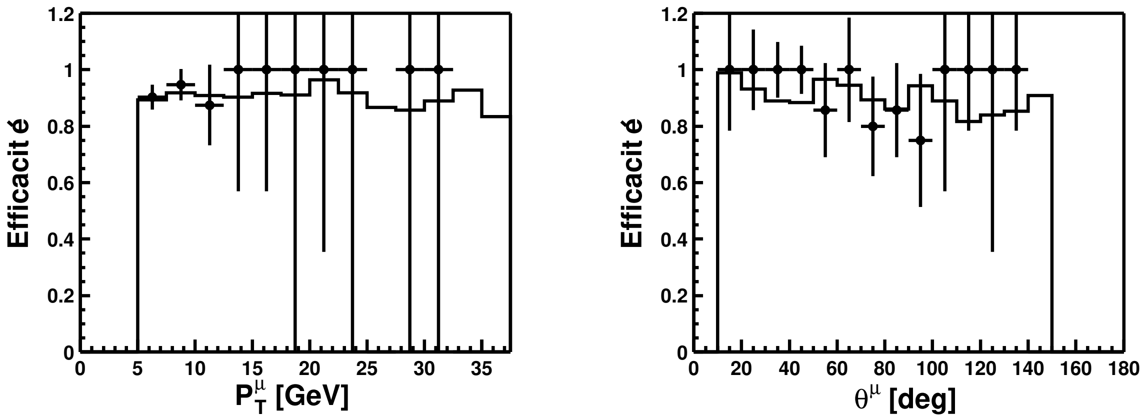


Fig. 5.3 – Efficacité de reconstruction du muon avec les coupures finales en fonction de P_T^μ et θ_μ.

Nous observons 102 événements pour 98 $\pm$ 14 événements prédits. L'efficacité d'identification avec les coupures finales de muon est représentée sur la figure 5.3, en fonction de l'angle polaire et de l'impulsion transverse de celui-ci. Elle varie peu en fonction de P_T^μ et θ_μ et vaut à peu près 90%. Nous observons un bon accord entre données et simulations Monte-Carlo pour des P_T^μ compris entre 5 et 12,5 GeV. Au-delà, seuls 13 événements sont reconstruits dans les données, ce qui est insuffisant pour étudier le comportement dans le domaine cinématique de l'analyse. Les efficacités sont décrites avec une précision de l'ordre 5% qui sera prise en compte dans les erreurs systématiques.

5.1.4 Critères du corps jet

Dans la partie 4.4, nous avons vu que la reconstruction des jets est effectuée à partir des hadrons. Ces hadrons sont reconstruits après l'identification des leptons et plus spécialement des électrons de l'événement. Or, parfois, des électrons peuvent être identifiés comme objet hadronique lorsqu'ils ne sont pas trouvés par l'identificateur d'électrons. Dans ce cas, ils peuvent être reconstruits comme un jet. Afin d'y rémédier, nous imposons trois critères supplémentaires aux jets reconstruits. Nous demandons que leur taille dans le plan $(\eta - \phi)$ soit supérieure à 0,1 ou que leur fraction électromagnétique soit inférieure à 0,9 (noté "$R > 0,1 OUEM < 0,9$" sur la figure 5.4) pour éliminer par exemple les électrons qui se sont enfuis dans les trous du détecteur. Finalement, les jets doivent contenir, lorsqu'ils sont reconstruits dans la partie centrale, au moins une trace ajustée au vertex (noté "$N(traces) >= 1$" sur la figure 5.4) et deux lorsqu'ils pointent vers des trous du détecteur . La validité de l'algorithme de jets et la comparaison entre les jets reconstruits et générés de la simulation Monte-Carlo ont été montrés dans la partie 4.5.2. Pour mesurer à quel point les coupures plus serrées déteriorent l'efficacité d'identification, nous sélectionnons des événements à courant neutre pour lesquels il n'existe, à l'ordre dominant, qu'un seul jet provenant du quark éjecté du proton. En appliquant les critères de sélection supplémentaires sur le jet, nous pouvons étudier leur impact sur l'efficacité d'identification ϵ_{jet}. Nous avons choisi la sélection courant neutre utilisée pour l'étude des critères de sélection du corps électron de la partie 5.1.1 en appliquant des coupures supplémentaires :

- Sélection courant neutre de la partie 5.1.1.
- $N_{jet} \geq 1$
- $P_T^{jet} > 20$ GeV
- $10^o < \theta_{jet} < 150^o$

Nous obtenons un lot avec 20957 événements observés pour 20582 $\pm$ 4507 prédits. L'efficacité d'identification en fonction de l'angle polaire et de l'impulsion transverse du jet de plus grand P_T en appliquant les critères supplémentaires est montrée dans la figure 5.4. Elle avoisine 98% et reste stable avec P_T^{jet}, jusqu'à des valeurs de P_T^{jet} = 80 GeV.

5.1.5 Critères du corps neutrino

Nous reconstruisons l'impulsion transverse manquante à partir des quadri-vecteurs impulsion-énergie de toutes les particules de l'événement. Nous associons l'énergie manquante au neutrino qui n'est pas détecté. Pour traiter sur un pied d'égalité le neutrino et les autres objets, il faut reconstruire l'angle polaire et azimutal de celui-ci. Pour cela, nous utilisons la quantité conservée $(E - P_z)_{vis}$ reconstruite avec les quadri-vecteurs de toutes les particules visibles. L'impulsion transverse, l'angle

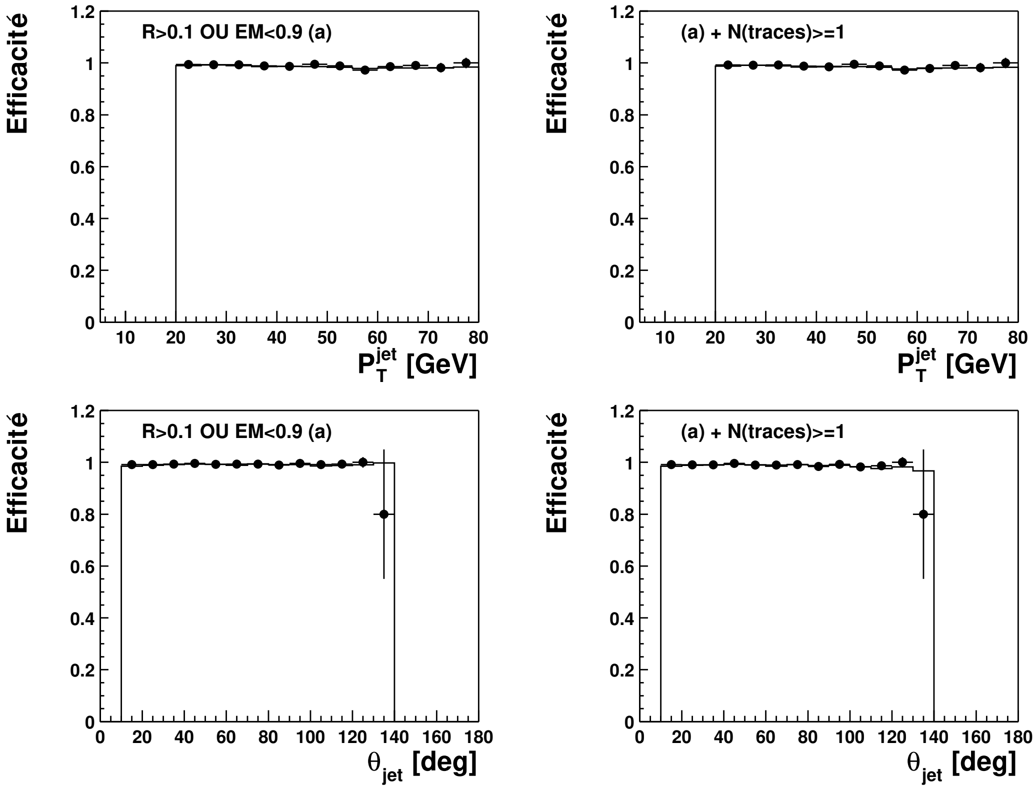


FIG. 5.4 – Efficacité des critères supplémentaires de reconstruction des jets en fonction de P_T^{jet} et θ_{jet}.

polaire et azimutal se calculent alors à partir de :

$$\begin{aligned}\overrightarrow{P_T^{\nu}} &= \overrightarrow{P_T^{miss}} \\ \theta_\nu &= 2 \cdot atan\left(\frac{55,2-(E-P_z)_{vis}}{P_T^{miss}}\right) \qquad (5.2) \\ & \qquad (5.3)\end{aligned}$$

5.2 Sélection de base

5.2.1 Définition cinématique des canaux

Dans la partie précédente, nous avons passé en revu les cinq types de corps recherchés et étudié les critères de qualité permettant de les identifier. Ces cinq corps seront traités sur un pied d'égalité en ne faisant à priori aucune hypothèse sur la topologie et les caractéristiques du signal exotique recherché. Un événement est retenu dans le lot final lorsqu'au moins deux corps bien isolés sont identifiés avec :

- L'énergie transverse E_T de chaque corps supérieure à 20 GeV
- $10^o < \theta_{corps} < 150^o$
- La distance dans le plan (η-ϕ) entre les corps plus grande que 1,0 radian

Les corps identifiés peuvent être du même type ou de type différent.
La coupure angulaire jusqu'à 10^o permet d'étendre la recherche dans la partie avant où les particules sont attendues dans des événements exotiques. Le critère d'isolation est nécessaire car nous voulons classifier l'événement en fonction des corps primaires qui seraient par exemple produits par la désintégration d'une particule lourde et non pas en fonction d'éventuelles particules secondaires radiées.
Ces coupures, résumées dans le tableau 5.2, ne permettent pas d'associer dans tous les cas un événement à un canal précis. En effet, il faut définir une prescription pour classer les événements sélectionnés. Lorsqu'un événement contient 3 corps isolés les uns par rapport aux autres, nous l'associons au canal qui correspond aux types des corps : par exemple un événement contenant un électron, un muon et un jet est retenu dans le canal e-μ-j.

Corps	Critères
e, μ, j, ν, γ	critères du paragraphe 5.1 $10^o < \theta_{corps} < 150^o$ $E_T^{corps} > E_T^{MIN} = 20$ GeV $R_{corps1}^{corps2} > R_{MIN} = 1{,}0$ radian

TAB. 5.2 – Coupures cinématiques de sélection.

Par contre, lorsque seuls deux parmi ces trois corps sont isolés entre eux, comme dans la figure 5.5, nous devons définir une convention.

Nous avons choisi d'ordonner les corps avec une énergie transverse décroissante, c'est-à-dire avec $E_T^{corps1} > E_T^{corps2} > E_T^{corps3}$. Ainsi, l'événement de la figure 5.5 va être classé dans le canal e-μ si l'énergie transverse du muon est plus grande que celle du jet et dans e-j dans le cas contraire. Cette convention nous permet de définir des canaux à partir des objets physiques d'énergie transverse décroissante.
Nous pouvons nous demander si cette convention ne biaise pas les résultats finaux. Afin d'estimer l'impact de cette prescription, nous avons comparé les estimations Monte-Carlo obtenues avec la prescription définie ci-dessus et une prescription dans laquelle on ordonne les corps suivant une préférence leptonique, avec les priorités décroissantes suivantes : muon $\rightarrow$ électron $\rightarrow$ photon $\rightarrow$ jet $\rightarrow$ neutrino. Les différences en taux absolus sont comprises entre 0 % pour les canaux e-μ et μ-μ, et atteignent 20% pour le canal μ-j-ν. La similitude des résultats s'explique par le fait que seuls les événements contenant une topologie avec trois corps (de section efficace réduite par rapport aux deux corps) sont susceptibles de migrer d'un canal à l'autre. Comme le taux d'événements est faible par rapport aux canaux à deux corps, ces migrations d'un canal à deux corps vers un autre sont petites.

Nous avons vu comment classer les événements contenant des corps de type différent. Dans le cas où nous identifions des corps de même type dans un canal, comme par exemple deux jets, nous appellerons $jet1$ le jet de plus grand E_T et $jet2$ celui de moins grand E_T : les objets de même type dans un canal sont classés selon un ordre décroissant en E_T.

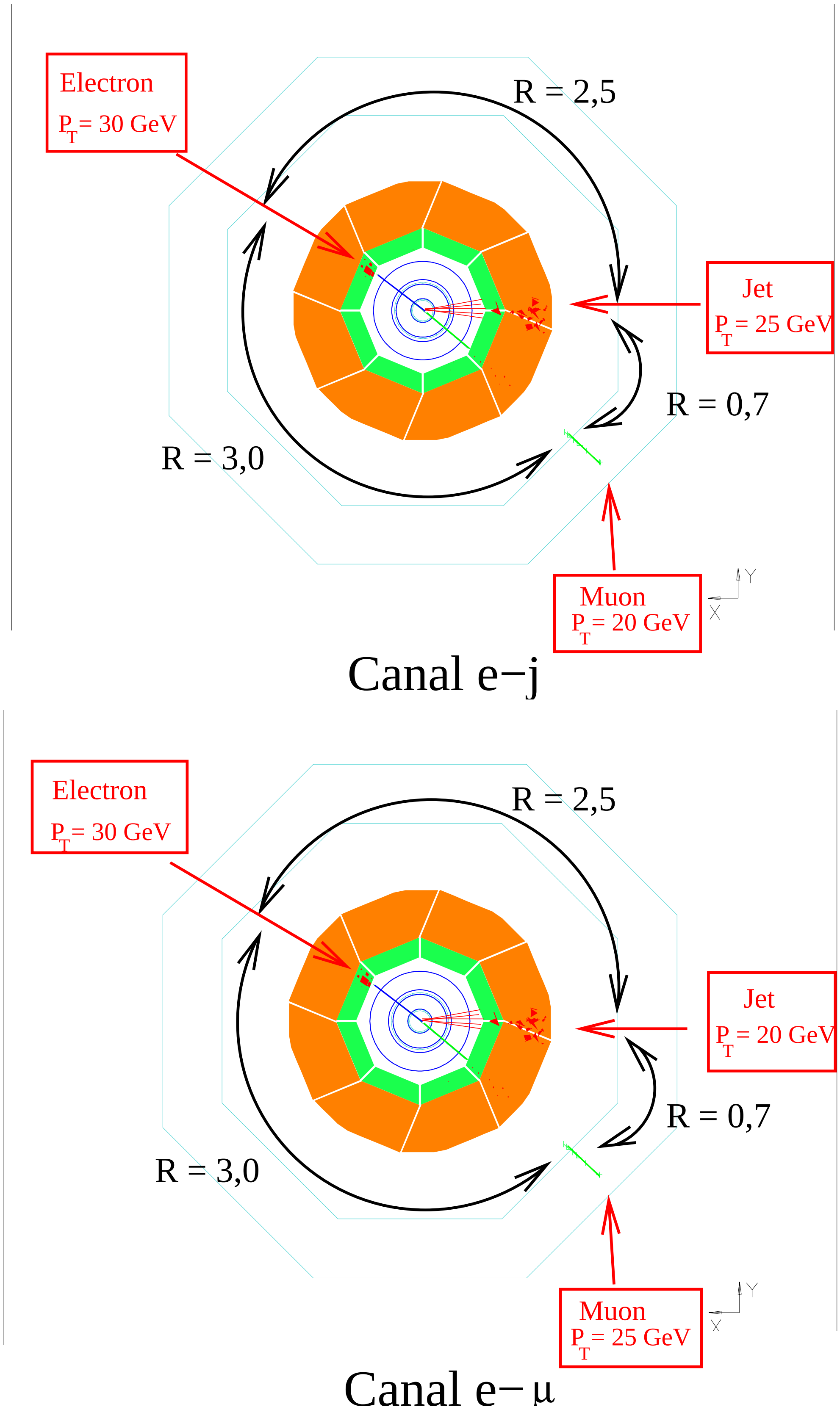


FIG. 5.5 – Définition des canaux : Si on présélectionne 3 corps de grande énergie transverse, dont deux ne sont pas isolés entre eux, l'événement est attribué au canal contenant les corps de plus grande énergie transverse.

5.2.2 Définition des critères de qualité

Afin de ne sélectionner que des événements physiques venant d'interactions *electron−proton*, nous demandons que :

- $|z_{vtx}| < 40$ cm
- $E - P_z < 75$ GeV
- Non activation des identificateurs "0" à "5" et "8" de bruit de fond décrits dans [56, 57]

L'utilisation des identificateurs topologiques de bruit de fond permet de rejeter les événements de muons cosmiques et les muons accompagnant le faisceau de protons qui traversent le détecteur. Les muons issus de rayonnement cosmique ou accompagnant le faisceau de protons (appelés "beam-halo") peuvent interagir dans le calorimètre et donner des gerbes électromagnétiques qui pourraient être identifiées comme électron. Lorsque ces gerbes s'étalent en plus dans la section hadronique du calorimètre LAr, des jets et/ou de l'impulsion manquante peuvent être reconstruits. Dans tous les canaux, à part le canal e-μ, nous utilisons en plus les identificateurs numérotés "6" et "7" rejetant les muons cosmiques à partir d'un dépôt d'énergie dans le calorimètre LAr aligné avec des segments de traces ou un dépôt d'énergie dans le fer instrumenté. Dans le canal e-μ, ces identificateurs apportent une inefficacité totale d'environ 16% [56, 57] et ne peuvent donc être utilisés. Dans les canaux contenant un neutrino, où des dépôts calorimétriques dus aux halos de muons peuvent simuler de l'impulsion transverse manquante, nous appliquons aussi le dernier identificateur numéroté "9" rejetant des événements à partir de leur signature dans le calorimètre LAr uniquement. Les identificateurs présentés permettent donc de réduire le bruit de fond non physique. Ils sont également appliqués aux simulations Monte-Carlo afin de tenir compte de leur inefficacité.

Pour renforcer la réjection des muons cosmiques et de halo dans les données, nous utilisons aussi l'information temporelle déterminée à partir du temps de dérive dans les trajectographes CJC et du calorimètre LAr. Les événements sont gardés lorsque leur information temporelle correspond à un croisement de faisceau : les tolérances de $\pm$ 0,7 croisement de faisceau pour le calorimètre LAr et $\pm$ 0,1 pour les chambres à dérive CJC sont choisies pour maintenir une haute efficacité de sélection.

Lot	Critères
Données/MC	- $\|z_{vtx}\| < 40$ cm - $E - P_z < 75$ GeV - Non activation des identificateurs "0" à "8" de bruit de fond (à part le canal e-μ) - Non activation de l'identificateur "9" de bruit de fond (canaux à neutrino uniquement)
Données	- Information temporelle : $\|T_{LAr}\| < 0,7$ & $\|T_{CJC}\| < 0,1$ - HT LAr, SpaCal, Lumi, TOF, CJC

TAB. 5.3 – Coupures de qualité.

Enfin, nous appliquons des critères de qualité sur l'état des différents sous-détecteurs. Nous demandons que les chambres à dérive centrales CJC 1 et CJC 2, le système de luminosité (abrégé $Lumi$), les calorimètre LAr et SpaCal et le système de temps de vol (abrégé TOF) soient en marche (HT pour "Haute Tension"). Tous les critères utilisés sont résumés dans le tableau 5.3.

5.2.3 Lots de données et simulations Monte-Carlo

Nous avons divisé le lot global de données de la période HERA I, d'une luminosité de 118,3 pb^{-1}, en trois sous-lots dont les propriétés sont reportées dans le tableau 5.4 : nous distinguons les lots e^+p 99-00, e^-p 98-99 et e^+p 94-97.

Période	Type de collision	Luminosité (pb^{-1})	$\sqrt{s}$ (GeV)
1994-1998	e^+p	37,5	300
1998-1999	e^-p	14,2	318
1999-2000	e^+p	66,6	318
2002-2003	e^+p	2,4	318

TAB. 5.4 – Pour les différentes périodes de prise de données, l'énergie au centre de masse ainsi que la luminosité collectée par H1.

Les données sont comparées aux simulations Monte-Carlo des processus suivants :
- Photoproduction
- Courants neutres
- Courants chargés
- Production de paire de leptons
- Diffusion QED Compton
- Production de boson $W^\pm$

Les processus ont été générés dans différents domaines cinématiques. Pour le processus de photoproduction, les domaines cinématiques sont séparés suivant leur valeur de $\widehat{P_T}$, l'impulsion transverse du parton final dans le centre de masse eq. Les processus à courant neutre et à courant chargé générés avec le générateur DJANGO sont divisés en régions de Q^2. Dans le lot de courants neutres généré avec RAPGAP, une coupure en impulsion transverse manquante P_T^{miss} ou en impulsion transverse du jet généré de plus grand P_T a été appliquée. Les lots de QED Compton sont divisés en régions de masse invariante du système électron-photon. Pour les processus de production de paires de leptons, une coupure sur l'impulsion transverse des leptons générés a été appliquée.

Tous les processus ont été simulés avec les conditions de prises de données respectives aux périodes de run. Les caractéristiques des différents lots Monte-Carlo utilisés ainsi que la luminosité simulée associée sont résumées dans le tableau 5.5.

Processus	Générateur	Variable	Luminosité (fb^{-1})
Photoproduction	PYTHIA 6.1	$\widehat{P_T} > 10$ GeV	0,1
		$\widehat{P_T} > 15$ GeV	1
		$\widehat{P_T} > 25$ GeV	10
		$\widehat{P_T} > 40$ GeV	10
DIS NC	DJANGO 6.2	$Q^2 > 100$ GeV^2	0,3
		$Q^2 > 1000$ GeV^2	1,1
		$Q^2 > 10000$ GeV^2	57
DIS NC	RAPGAP 2.8	$P_T^{miss} > 8$ GeV $\|\|$	
		$P_T^{jet} > 12$ GeV	1,2
DIS CC	DJANGO 6.2	$Q^2 > 100$ GeV^2	7
		$Q^2 > 10000$ GeV^2	360
Production de paires de leptons	GRAPE ee	$P_T^{ele} > 8$ GeV	60
	GRAPE $\mu\mu$	$P_T^{\mu} > 8$ GeV	100
QED Compton	WABGEN	$M_{e,\gamma} > 10$ GeV	3,4
		$M_{e,\gamma} > 50$ GeV	124
		$M_{e,\gamma} > 100$ GeV	1900
Production de $W^{\pm}$	EPVEC	-	100

TAB. 5.5 – Caractéristiques des lots généres : pour chaque processus, le générateur utilisé et la luminosité intégrée simulée sont reportés.

5.2.4 Etude des efficacités du système de déclenchement

Pour chaque canal analysé, il nous faut prendre en compte l'efficacité du système de déclenchement de H1. L'efficacité de déclenchement pour un dépôt électromagnétique avec un $P_T > 20$ GeV est quasiment 100% (d'après [2]). Donc, on attribuera à tous les canaux contenant au moins un objet électromagnétique de type électron ou photon une efficacité de déclenchement de 10%. Il reste à déterminer l'efficacité de déclenchement d'un muon, d'un jet et de l'impulsion manquante. Nous avons utilisé les efficacités déterminées dans les analyses [1, 2] et [12] pour l'impulsion manquante et le muon. Nous avons mesuré l'efficacité de déclenchement pour des jets à partir des lots de données et vérifié l'efficacité de déclenchement de muons et d'impulsion manquante, étudiée dans des analyses dédiées, à partir de lots de simulations Monte-Carlo. Les efficacités de déclenchement sont prises en compte dans notre analyse en pondérant les lots Monte-Carlo simulés avec les efficacités mesurées. Les efficacités appliquées dans les différents canaux sont résumées dans le tableau 5.6.

5.2.4.1 Efficacité de déclenchement de jets

Pour étudier l'efficacité de déclenchement, il faut déterminer les sous-déclencheurs susceptibles de déclencher et un sous-déclencheur moniteur indépendant pour la mesure de l'efficacité de déclenchement. Nous utiliserons les sous-déclencheurs décrits dans la partie 2.2.5.1 et utilisés lors des analyses dédiées de jets. Il s'agit des sous-déclencheurs appelés S64, S67, S75 et S77 qui déclenchent essentiellement à partir d'un dépôt d'énergie d'un certain seuil (quelques GeV) dans le calorimètre LAr. Nous allons estimer l'efficacité de déclenchement ϵ_j à partir de la relation suivante :

$$\epsilon_j = \frac{Nombre\ d\acute{}\acute{e}venements\ d\acute{e}clench\acute{e}s\ avec\ \mathbf{S}\ et\ \mathbf{M}}{Nombre\ d\acute{}\acute{e}venements\ d\acute{e}clench\acute{e}s\ avec\ \mathbf{M}} \tag{5.4}$$

où S correspond aux sous-déclencheurs considérés dans notre analyse, et M au sous-déclencheur moniteur. Nous comparerons par la suite deux méthodes différentes que nous appellerons "Méthode courant neutre" et "Méthode Electron tagger", correspondant à l'utilisation de deux sous-déclencheurs moniteur différents et donc à deux sous-lots de données différents, pour estimer l'efficacité.

Méthode courant neutre

La première méthode, employée lors des mesures de sections efficaces de production de di-jets [62], est basée sur l'utilisation du sous-déclencheur S3 au niveau L4 en tant que sous-déclencheur moniteur. Le S3 déclenche lors de la détection d'un dépôt électromagnétique dans le calorimètre arrière SpaCal et est donc indépendant des sous-déclencheurs S64, S67, S75 et S77. Nous allons constituer un lot de contrôle de courants neutres, avec un électron reconstruit dans le calorimètre arrière SpaCal, permettant d'estimer l'efficacité ϵ_j^{NC}. Cette méthode suppose que la topologie des jets est la même pour les courants neutres et les processus de photoproduction [7]. S correspond aux quatre sous-déclencheurs S64, S67, S75 et S77 et M au sous-déclencheur moniteur S3 au niveau L4. Nous définissons un lot de contrôle à partir des événements reconstruits dans le canal j-j dans lequel nous sélectionnons plus de 30000 événements pour les lots e^+p 99-00, e^-p 98-99 et e^+p 94-97 cumulés. Le lot de contrôle satisfait les critères suivants :
- Sélection de l'analyse générique (canal j-j)
- Reconstruction d'un dépôt électromagnétique dans le calorimètre SpaCal

Nous nous plaçons ainsi dans la configuration où les propriétés cinématiques des jets de ce lot sont similaires aux propriétés des jets dans le lot final.

Méthode Electron Tagger

La seconde méthode utilise le déclencheur moniteur S83 qui déclenche lorsqu'il y a coïncidence entre un dépôt dans l'"Electron-Tagger" (ET), la présence d'un vertex et d'un certain nombre de traces. Ce déclencheur a été utilisé dans des mesures de sections efficaces de photoproduction à grande impulsion transverse [63] où l'électron a déposé de l'énergie dans le ET. Nous allons constituer un lot de contrôle remplissant les critères suivant :
- L'énergie déposée dans le détecteur de photon du luminomètre doit être nulle
- La position du point d'impact du positron diffusé sur la fenêtre du détecteur ET doit être telle que $|X_{ET}| \leq 6{,}5$ cm

– La valeur de y reconstruite à partir de l'énergie déposée dans le ET doit satisfaire la relation : 0,3 < y_{ET} < 0,65, avec y_{ET} = 1 - E_{ET}/E_e où E_e est l'énergie des électrons incidents.

Ce lot de contrôle permet d'estimer l'efficacité ϵ_j^{ET}, où S correspond aux quatre sous-déclencheur S64, S67, S75 et S77 et M au sous-déclencheur moniteur S83 au niveau L1.

Résultats

Nous avons représenté l'efficacité de déclenchement obtenue avec la méthode ET en fonction de P_T^{jet} et de θ_{jet} dans la figure 5.6, en la comparant à celle obtenue avec la méthode NC. Pour les lots e^+p 94-97, l'information du ET n'était pas disponible pour la totalité de la période de prises de données : seule l'efficacité de la méthode NC est présentée.

De manière générale, l'efficacité de déclenchement atteint 100% dès 30 GeV en P_T^{jet} avec les deux méthodes. Nous observons que l'efficacité, proche de 100% à grand P_T, chute à plus petit P_T. Pour le lot de données e^+p 94-97, l'efficacité dépasse 92% dès 20 GeV en P_T^{jet}. Pour les lots e^-p 98-99 et e^+p 99-00, elle vaut environ 98%. Les deux méthodes présentées donnent des résulats similaires à 2% près. Nous tenons compte de cette efficacité de déclenchement et de cette incertitude systématique dans le canal j-j.

5.2.4.2 Efficacité de déclenchement de muons

Pour le canal μ-μ, nous utilisons les sous-déclencheurs S19, S22, S34 et S56 au niveau L1 décrits dans la partie 2.2.5.1, basés sur la coïncidence de haute multiplicité de traces et de signaux dans le système à muons. Leur efficacité a été mesurée dans l'analyse dédiée à la production de paires de muons [12] et il a été montré qu'elle était bien reproduite dans les simulations Monte-Carlo. La figure 5.7 montre l'efficacité en fonction de l'impulsion transverse du muon de plus grand P_T à partir des sous-déclencheurs simulés dans le Monte-Carlo : elle dépend peu de l'impulsion transverse du muon, P_T^{μ}, et vaut 70%.

Nous en tenons compte dans les canaux μ-μ, μ-j et μ-j-ν.

5.2.4.3 Efficacité de déclenchement d'impulsion manquante (neutrino)

L'efficacité de déclenchement d'impulsion manquante $\epsilon_{P_T^{miss}}$ a été mesurée dans les analyses dédiées aux courants chargés [1, 2]. Les sous-déclencheurs utilisés sont le S66, S67, S71, S75 et S77, décrits dans la partie 2.2.5.1. Le comportement en fonction de P_T^{miss} est montré sur la figure 5.8.

Elle dépasse 85% dès 20 GeV en impulsion manquante et atteint quasiment 100% à partir de 40 GeV.

Nous en tenons compte dans les canaux à neutrino j-ν, j-j-ν et μ-j-ν. Dans le canal μ-j-ν, il suffit que l'un des sous-déclencheurs dédiés aux muons ou l'un des sous-déclencheurs dédiés aux courants chargés ait déclenché. Dans le canal μ-j, nous utilisons aussi les sous-déclencheurs courant chargé car leur efficacité est plus grande que celle des déclencheurs muons.

Canal	Efficacité de déclenchement
μ-μ	ϵ_μ
μ-j	$\epsilon_\mu / \epsilon_{P_T^{miss}}$
j-j	ϵ_j^{NC}
j-ν	$\epsilon_{P_T^{miss}}$
j-j-ν	$\epsilon_{P_T^{miss}}$
μ-j-ν	$\epsilon_\mu / \epsilon_{P_T^{miss}}$

TAB. 5.6 – Canaux pour lesquels nous avons tenu compte des efficacités de déclenchement. Lorsque deux efficacités sont mentionnées pour le même canal, un événement est retenu lorsqu'un des sous-déclencheurs correspondants a déclenché. Pour tous les autres canaux, l'efficacité de déclenchement est estimée égale à 100%.

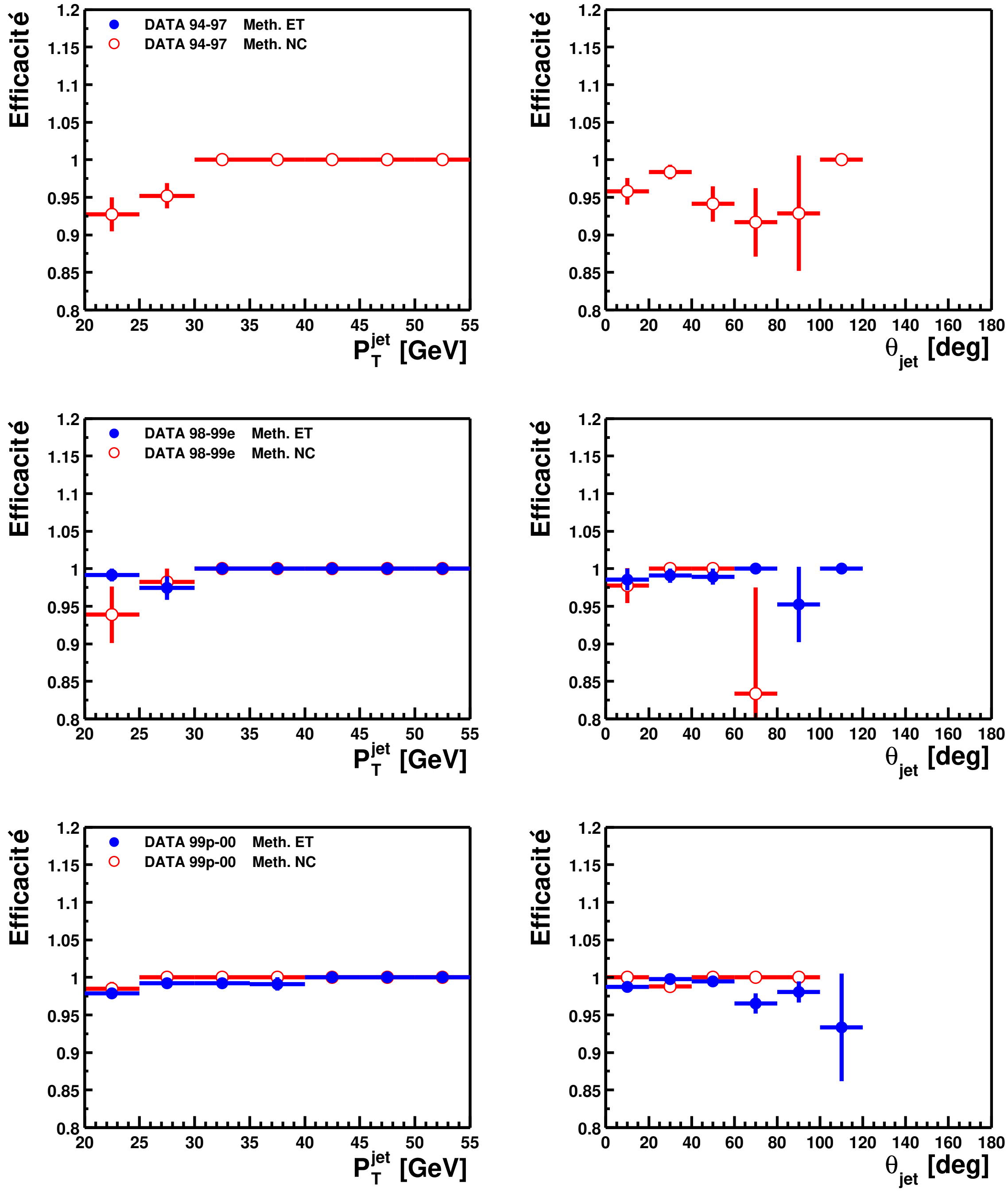


FIG. 5.6 – Efficacités du système de déclenchement ϵ_j^{ET} et ϵ_j^{NC} en fonction de P_T^{jet} et de θ_{jet} pour les lots e^+p 94-97, e^-p 98-99 et e^+p 99-00.

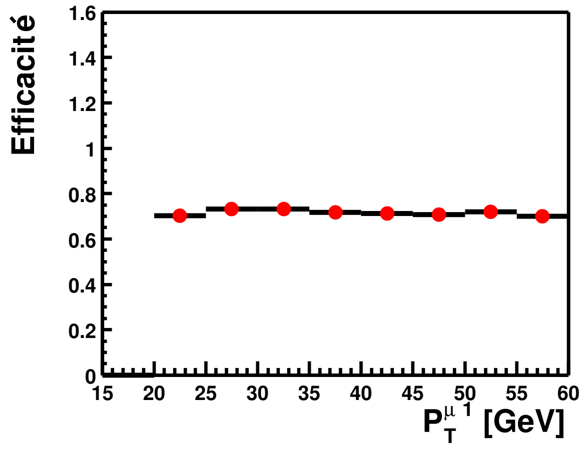


FIG. 5.7 – Efficacité du système de déclenchement ϵ_μ en fonction de P_T^μ.

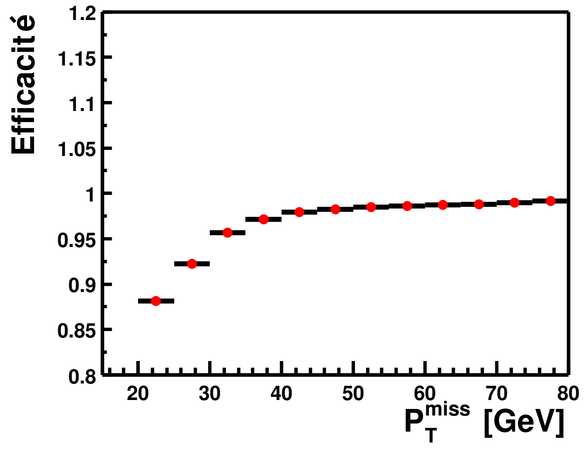


FIG. 5.8 – Efficacité du système de déclenchement $\epsilon_{P_T^{miss}}$ en fonction de P_T^{miss}.

5.2.5 Résultats bruts

A partir des critères cinématiques et de qualité (tableaux 5.2 et 5.3) de la sélection de base, appelée dans la suite sélection brute, et en tenant compte des efficacités de déclenchement, nous avons constitué notre lot de base. Au total, nous avons considéré cent un canaux. Il s'agit de tous les canaux à deux, trois et quatre corps, ainsi que de deux canaux à cinq corps. Dans les données, comme le montre la figure 5.9, aucun événement à six corps ou plus n'a été trouvé.

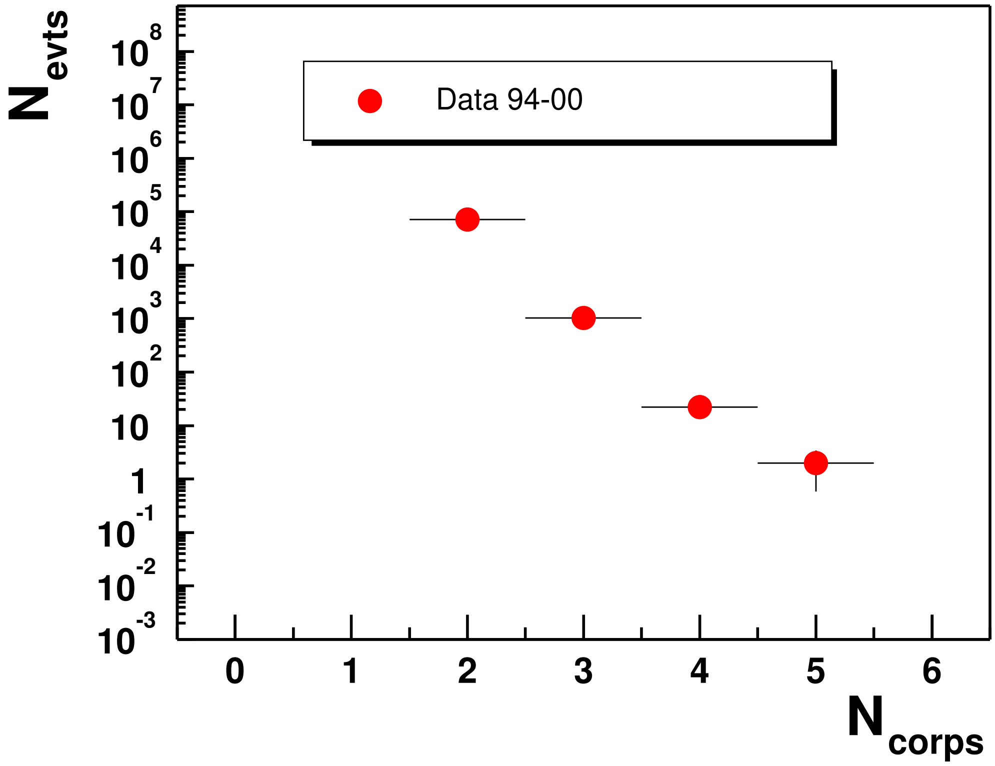


FIG. 5.9 – Nombre d'événements sélectionnés en fonction du nombre de corps pour la totalité des données (lots e^+p 99-00, e^-p 98-99 et e^+p 94-97).

La figure 5.10 nous montre le taux d'événements observés et prédits dans les canaux à deux, trois et quatre corps pour les données et les simulations Monte-Carlo en appliquant la sélection brute.

Nous observons en tout 24 canaux ayant un nombre d'événements de données supérieur ou égal à 1, et 24 canaux avec un nombre d'événements prédits supérieur à 0,5. Nous allons dans l'analyse finale retenir uniquement les canaux satisfaisant l'une de ces deux conditions. Il reste finalement 27 canaux à étudier. Parmi ces 27 canaux, il faut écarter, dans un second temps, les trois canaux suivants de l'analyse complète :

- μ-ν :
  Ce canal est dominé, dans les données, par un bruit de fond de photoproduction

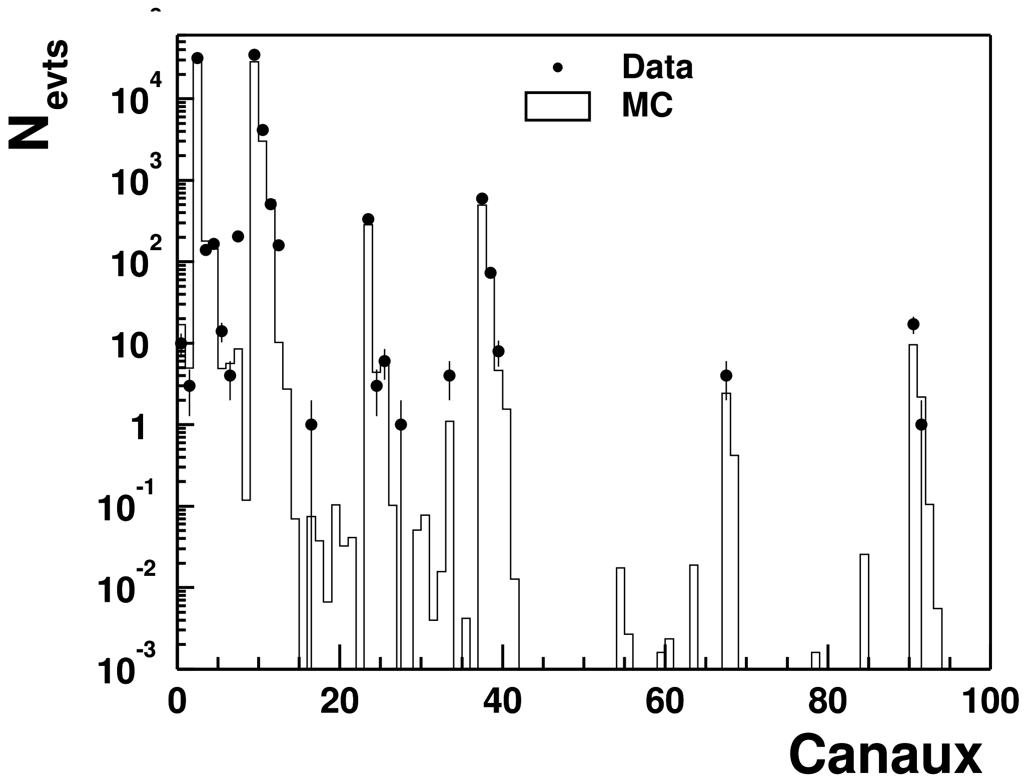


FIG. 5.10 – Taux d'événements prédits et observés dans les canaux à deux, trois et quatre corps après la sélection brute (lots e^+p 99-00, e^-p 98-99 et e^+p 94-97). Les différents canaux sont indexés en abscisse par nombre de corps croissants.

de basse impulsion transverse : des hadrons dits de "punch through" de basse impulsion transverse traversent à l'avant le calorimètre pour déposer la plupart de leur énergie dans le fer instrumenté. Ils sont alors identifiés comme muons. Nous écartons ce canal car ce bruit n'est pas prédictible de manière fiable.

– **ν-γ, γ-γ :**
Il faut écarter ces canaux car nous recherchons des événements avec un vertex reconstruit, ce qui n'est pas assuré dans ces canaux car aucun des corps ne possède de trace chargée.

Il subsiste 24 canaux qui ont été scrupuleusement analysés : nous étudions 10 canaux à deux corps, 9 canaux à trois corps, 3 canaux à quatre corps et 2 canaux à cinq corps. Après la sélection brute, ces canaux sont encore affectés par deux types de bruits de fond détériorant l'accord entre les données et les simulations Monte-Carlo :
– Les bruits de fond non physiques
– Les bruits de fond physiques venant des interactions *ep*

Ces deux types de bruit de fond et la manière de les réduire sont discutés dans la suite.

5.3 Sélection raffinée

Les coupures décrites dans la sélection brute permettent de rejeter la plupart des bruits de fond rencontrés. Cependant, comme nous allons le voir, certains canaux sont encore affectés de bruit de fond non physique, comme par exemple la reconstruction de muons cosmiques dans le canal μ-μ, ou par de la contamination venant d'autres canaux, due à une mauvaise identification de l'un des corps.

5.3.1 Les bruits de fond non physiques

Dans le canal μ-μ des événements de bruit de fond non physique subsistent. Des muons cosmiques sont reconstruits en coïncidence avec un croisement de faisceau. Ceux-ci peuvent être rejetés, comme dans l'analyse de di-muons [12], à partir de l'angle d'ouverture entre les deux muons reconstruits. En effet, l'angle d'ouverture, appelé α, entre les muons reconstruits doit être proche de 180^o pour des muons cosmiques. La figure 5.11 montre la distribution de α avant la coupure plaçée à 165^o. Nous rejetons 19 événements (86%) pour environ 0,01 de signal attendu.

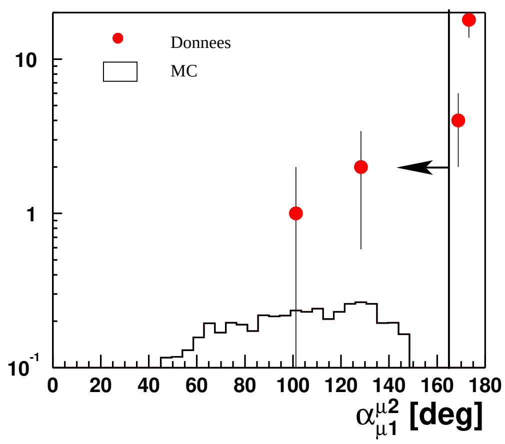


FIG. 5.11 – Canal μ-μ : Angle d'ouverture α dans le plan $(\eta - \phi)$ entre les deux muons identifiés.

Par ailleurs, nous avons visualisé tous les événements des canaux à faible statistique ($N_{evts} < 35$), et des queues de spectre en masse invariante situées entre 100 et 280 GeV selon les canaux. En plus du canal μ-μ, trois canaux présentent des événements de bruit de fond non physique que nous rejettons par inspection visuelle : il s'agit des canaux j-ν, μ-j et j-γ.

Dans le canal j-ν, nous trouvons des muons cosmiques dont les gerbes isolées dans la partie hadronique du calorimètre, comme cela est montré sur la figure 5.12, simulent de l'énergie manquante.

Dans le canal μ-j, des muons cosmiques traversant le détecteur sont identifiés comme muon dans le fer instrumenté, et leur gerbe dans le calorimètre dans la région opposée en ϕ, comme jet.

Des halos du faisceau de protons peuvent donner des dépôts isolés dans la partie électromagnétique du calorimètre. Ceux-ci peuvent alors être identifiés comme photon et contaminer le canal j-γ.

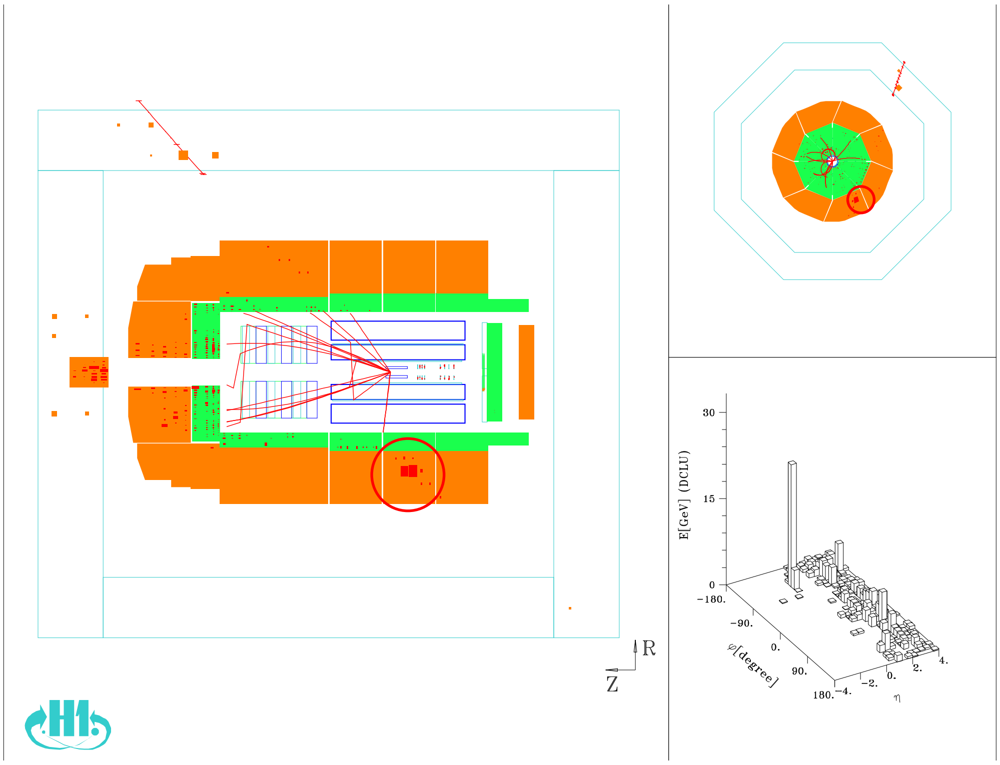


FIG. 5.12 – Canal j-ν : un exemple d'événement cosmique passant toutes les coupures. Dans la vue longitudinale et transverse, le cercle rouge marque le dépôt calorimétrique du muon dans la section hadronique.

Le tableau 5.7 résume le taux d'evénements par canal obtenus sur les trois lots e^+p 99-00, e^-p 98-99 et e^+p 94-97, avant et après rejection. Aucun événement à grande masse invariante n'est rejeté après la visualisation.

Canal	Nombre d'événements		Type d'événement rejeté
	Initial	Après visualisation/coupure	
μ-μ	22	3	Cosmique
μ-j	5	4	Cosmique
j-ν	3024	2826	Cosmique Halo
j-γ	528	504	Halo

TAB. 5.7 – Taux d'événements des canaux à bruit de fond non physique avant et après réjection complémentaire (données 1994-2000).

5.3.2 Les bruits de fond physiques

La seconde catégorie de bruit de fond résiduel est due à une identification erronée de particules. Dans certains canaux, que nous allons spécifier ci-dessous, une mauvaise identification de l'un des corps peut donner une contribution importante à un processus pour lequel la prédiction du Modèle Standard est quasi-nulle. Nous connaissons les processus du Modèle Standard susceptibles de donner les différentes topologies possibles. Ainsi, intrisinquèment, le processus de courant neutre ne peut contribuer au canal j-ν : ces processus étant équilibrés dans le plan transverse, une identification de neutrino ne serait due qu'à une mauvaise reconstruction des particules finales.

Afin d'évaluer la proportion de ces processus par canal, nous définissons la pureté d'identification P_i^{id} (i=canal1 .. cannalN) à partir de la relation suivante :

$$P_i^{id} = \frac{N_{signal,i}}{N_{signal,i} + N_{conta.,i}} \tag{5.5}$$

où $N_{signal,i}$ représente le nombre d'événements correspondant aux processus pouvant induire la topologie du canal i et $N_{conta.,i}$ le nombre d'événements de contamination provenant des autres processus dans le canal i. Les processus vrais et contaminants sont résumés dans le tableau 5.8 pour chaque canal. Pour cette étude, nous avons écarté les trois canaux e-γ-γ, e-j-j-j-j et j-j-j-j-j : la contribution totale des simulations Monte-Carlo est inférieure à 10^{-3} dans le canal e-j-j-j-j et est nulle dans les canaux e-γ-γ et j-j-j-j-j.

La distribution 5.13 montre qu'après la sélection brute, huit canaux possèdent une pureté de moins de 75% (e-e, e-ν, e-j-ν, j-ν, j-γ, j-j-ν, j-j-γ et j-j-j-ν). Afin d'être sûr qu'une possible déviation dans un canal ne soit pas due à une contamination, nous allons étudier les raisons des contaminations dans les canaux cités et discuter les coupures additionnelles pouvant être utilisées en vue d'améliorer la pureté de ces canaux. Cette étude sera principalement limitée aux canaux à deux corps afin de maintenir une sensibilité suffisante aux processus au-delà du Modèle Standard dans les canaux à 3 corps.

Canal	Processus signal du MS	Processus contaminants principaux du MS
e-e	$\gamma\gamma$	NC, *Compton*
e-μ	$\gamma\gamma$, $W^{\pm}$	NC
e-j	NC	γp, $\gamma\gamma$
e-ν	$W^{\pm}$	NC
e-γ	NC,*Compton*	$\gamma\gamma$
μ-μ	$\gamma\gamma$	-
μ-j	$\gamma\gamma$, $W^{\pm}$	NC, γp
j-j	NC, $W^{\pm}$, γp	NC
j-ν	CC, $W^{\pm}$	NC, γp
j-γ	γp, *Compton*	NC
e-j-j	NC,$W^{\pm}$	γp
e-j-ν	$W^{\pm}$	NC, CC, γp
e-j-γ	NC	$\gamma\gamma$
e-γ-γ	NC,*Compton*	$\gamma\gamma$
μ-j-ν	$W^{\pm}$	NC, CC, γp
j-j-j	NC, γp, $W^{\pm}$	NC
j-j-ν	CC, $W^{\pm}$	NC, γp
j-j-γ	γp, *Compton*	NC, CC
j-ν-γ	CC	NC, γp
e-j-j-j	NC, $W^{\pm}$	γp
j-j-j-j	NC, γp, $W^{\pm}$	NC
j-j-j-ν	CC, $W^{\pm}$	γp, NC
e-j-j-j-j	NC, $W^{\pm}$	γp
j-j-j-j-j	NC, γp, $W^{\pm}$	NC

TAB. 5.8 – Processus du MS induisant les vraies topologies et celles qui sont dues à une mauvaise identification.

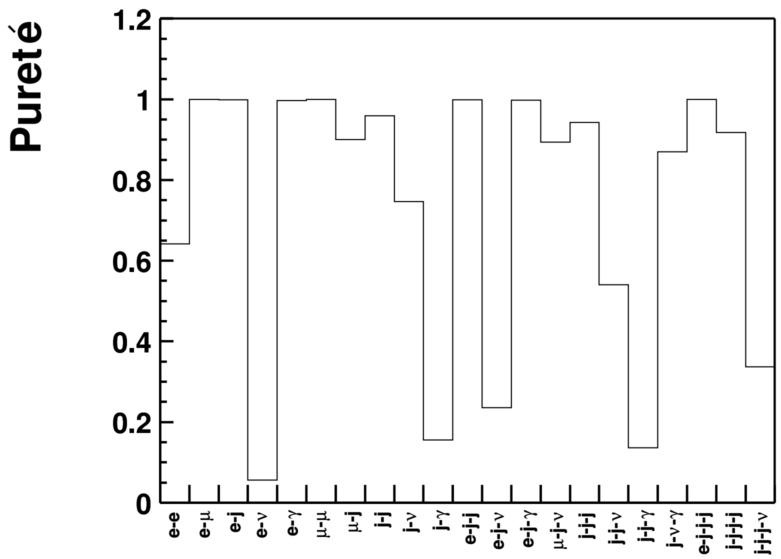


FIG. 5.13 – Pureté d'identification en fonction des canaux après la sélection brute.

Dans le canal j-γ, la contamination principale vient des courants neutres. En effet, ces événements migrent dans les canaux à photons à cause de l'inefficacité des trajectographes : des électrons sont identifiés comme photon car aucune trace n'est mesurée. Comme la section efficace de production de courants neutres est très importante à HERA, ces canaux présentent un fort taux de contamination. Dans les courants neutres, la topologie typique donne un électron diffusé à grand angle polaire, donc vers l'arrière du détecteur, alors que le jet est projeté vers l'avant. Dans le centre de masse du système j-γ, cette configuration se résume à l'inéquation

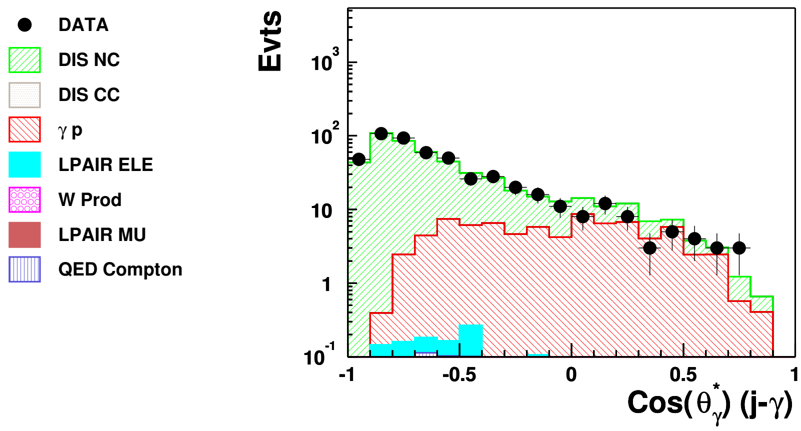


FIG. 5.14 – Canal j-γ : Distribution de la quantité $cos(\theta^*_\gamma)$.

suivante : $cos(\theta^*_\gamma) < 0{,}0$, où $cos(\theta^*_\gamma)$ est défini comme l'angle polaire de l'amas électromagnétique. La distribution est représentée dans la figure 5.14. Pour diminuer la contamination, nous allons considérer dans l'analyse finale uniquement la région définie par $cos(\theta^*_\gamma) > 0{,}0$.

Dans le canal e-e, la distribution (figure 5.15) de l'angle polaire du second électron (celui de plus faible P_T) montre une forte contribution de processus à courant neutre à l'avant ($\theta < 20^o$). Comme nous relâchons les critères d'isolation de la trace dans la partie avant (pour maintenir une haute efficacité d'identification), on identifie à tord des électrons dans les jets. Afin d'y remédier, nous restreignons dans ce canal l'acceptance angulaire à $20^o < \theta_{ele1,ele2} < 150^o$. La région retenue est celle utilisée dans l'analyse dédiée aux multi-électrons [11].

Dans les canaux contenant un neutrino, nous allons voir que deux types de processus simulent de l'impulsion manquante : nous avons d'une part les processus de courants neutres, dominants dans les canaux e-ν et e-j-ν, et d'autre part le processus de photoproduction, dominant dans les canaux contenant un jet (ou plus) en plus du neutrino.

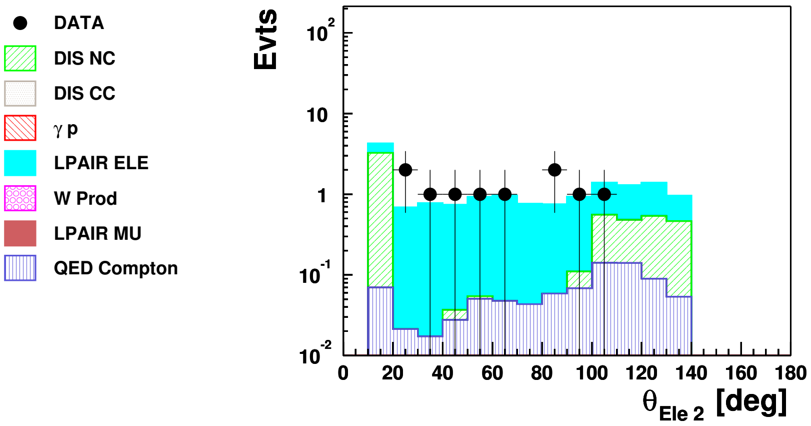


FIG. 5.15 – Canal e-e : distribution de l'angle polaire de diffusion du second candidat électron de plus grande impulsion transverse. A l'avant ($\theta_{ele} < 20^o$), des électrons faussement identifiés dans les jets constituent la prédiction principale du Modèle Standard.

Pour s'affranchir de la contribution de photoproduction et de courants neutres, on utilise une coupure développée dans les analyses dédiées aux courants chargés [1, 3] : la quantité $V = \frac{V_{ap}}{V_p}$ décrit l'équilibre azimutal de l'événement. V_{ap} et V_p correspondent aux projections des quantités d'énergie antiparallèle et parallèle à la direction de l'énergie transverse manquante. Pour les calculer, il faut séparer le plan transverse aux faisceaux en deux demi-plans, la frontière étant perpendiculaire à la direction de l'énergie transverse manquante (symbolisée par la flèche sur la figure 5.16).

L'énergie transverse de l'ensemble des dépôts calorimétriques est alors sommée dans chacun de ces deux demi-plans. Le demi-plan contenant la direction de l'énergie transverse manquante est défini comme le demi-plan anti-parallèle, la somme de l'énergie transverse des dépôts calorimétrique de ce plan est appelée V_{ap}. L'autre plan est appelé plan parallèle permettant de définir la somme V_p. Ces deux variables permettent de calculer $V = \frac{V_{ap}}{V_p}$. Dans un événement qui devrait être équilibré en énergie transverse mais où une partie de celle-ci semble manquante parce qu'un objet est mal reconstruit, les dépôts d'énergie seront repartis uniformément dans le plan transverse et le rapport $\frac{V_{ap}}{V_p}$ sera proche de 1. Si au contraire un neutrino est présent dans cet événement, les dépôts d'énergie seront plus importants dans la direction opposée à celle du neutrino. V_p étant alors nettement plus grand que V_{ap}, le rapport $\frac{V_{ap}}{V_p}$ sera plus proche de 0.

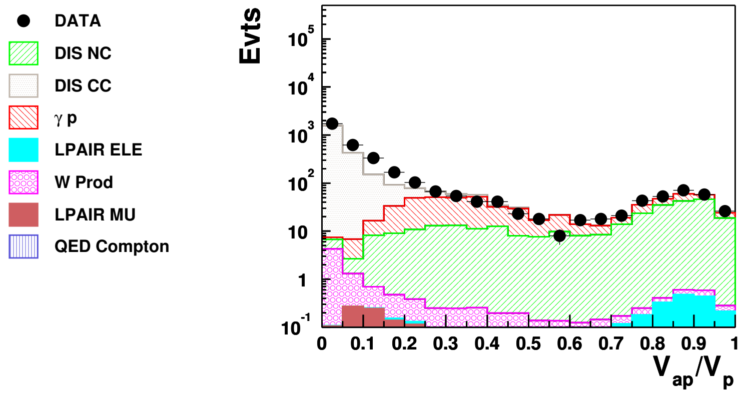


FIG. 5.17 – Canal j-ν : Distribution de la variable $V = V_{ap}/V_p$.

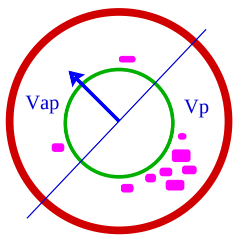


FIG. 5.16 – La variable $V = \frac{V_{ap}}{V_p}$

La figure 5.17 montre la distribution de la variable V dans le canal j-ν : les courants neutres et les processus de photoproduction sont concentrés vers de grandes valeurs de V. La coupure à 0,15 permet de diminuer le bruit de fond de 80%. La topologie des canaux e-ν et e-j-ν correspond à celle des événements dits leptons isolés décrits dans la publication [16] ; cependant, ils sont dominés par des événements courant neutre où nous avons, à tord, reconstruit de l'impulsion manquante importante et où, dans le cas du canal e-ν, l'impulsion transverse du jet reconstruit est trop faible pour passer la sélection. Dans ces canaux, nous relâchons la coupure de V à 0,5. Dans le cas particulier du canal e-j-ν, nous avons choisi d'appliquer cette coupure, comme dans [16], afin d'augmenter la sensibilité. Cette coupure permet de réduire la contribution de courants neutres mal identifiés de 40% dans le canal e-ν et de 80% dans le canal e-j-ν. La contribution des courants neutres peut être réduite davantage en demandant que la quantité $E - P_z$ soit inférieure à 45 GeV (car le neutrino a emporté de l'énergie) ; L'impureté élevée subsistante dans le canal e-ν est due aux courants neutres et courants chargés mal identifiés.

Dans le canal j-j-γ, le bruit de fond de courants neutres est également très important. Une solution possible serait d'appliquer une coupure sur la quantité $E - P_z$ permettant de réduire cette contribution. Néanmoins, une telle coupure serait beaucoup trop restrictive en terme d'efficacité vis à vis de processus au-delà du Modèle Standard pouvant être équilibrés en $E - P_z$. Elle n'a donc pas été retenue.

Canal	Coupures anti-bruit
e-e	$20^o < \theta_{ele1,2} < 150^o$
e-ν e-j-ν	$E - P_z < 45$ GeV $V = \frac{V_{ap}}{V_p} < 0{,}5$
j-ν	$V = \frac{V_{ap}}{V_p} < 0{,}15$
j-γ	$cos(\theta^*_\gamma) > 0{,}0$

TAB. 5.9 – Coupures supplémentaires appliquées pour réduire le bruit de fond de contamination.

En résumé, les critères de sélection supplémentaires rassemblés dans le tableau 5.9, permettent d'obtenir une pureté de 93% dans le canal e-e , de 98% dans le canal j-ν et de 84% dans le canal e-j-ν. Dans les canaux e-ν et j-γ, le taux de contamination reste élevé avec respectivement 78% et 36%. Le tableau 5.10 résume les taux avant et après application des coupures supplémentaires.

Canal	e-e	e-ν	j-ν	j-γ	e-j-ν
Pureté avant (%)	64	6	74	15	24
Pureté après (%)	93	22	98	64	84

TAB. 5.10 – Pureté avant et après les coupures supplémentaires dans les canaux pourvus d'une forte contamination.

5.4 Discussion des erreurs systématiques

Nos mesures sont entachées de plusieurs sources d'erreurs qu'il faut déterminer. Nous les discutons dans la suite et étudions leur impact dans l'analyse. Il y a 3 catégories principales d'erreurs systématiques :

- Les incertitudes expérimentales.
- Les incertitudes statistiques dues au nombre d'événements générés pour chaque processus.
- Les incertitudes théoriques des modèles utilisés par les programmes de Monte-Carlo pour décrire les différents processus.

Les effets individuels et combinés sont résumés dans la table 5.13.

5.4.1 Incertitudes statistiques des simulations

Nos échantillons de simulation des processus du Modèle Standard comportent un nombre fini d'événements : ils sont donc affectés par des fluctuations statistiques. Ces fluctuations constituent ce qu'on appelle les erreurs statistiques des simulations du Modèle Standard. Pour chaque canal, elles sont calculées et reportées dans la table 5.13. Elle sont plus faibles pour les canaux à taux d'événements importants, comme par exemple les canaux à deux corps j-j ou e-j. Elle ne vaut que 0,4% pour le dernier. Comme l'erreur est directement reliée au nombre d'événements prédits, nous retrouvons une erreur atteignant 75% dans les canaux à cinq corps, comme par exemple e-j-j-j-j. Cette incertitude pourrait en principe être réduite en simulant des lots plus importants, mais cela n'a pas été possible dans les délais de la thèse. Pour les canaux e-γ-γ et j-j-j-j-j, il faut noter que la prédiction nulle, décrite dans 2.3.1, ne nous permet pas d'estimer les erreurs.

5.4.2 Incertitudes expérimentales

Cinq types d'erreurs expérimentales ont été traitées dans cette analyse. Elles sont résumées dans la table 5.11 :

- **L'erreur sur l'échelle d'énergie :**
  D'après les vérifications de la calibration que nous avons effectuées, elle vaut entre 0,5% et 3% (paragraphe 4.2.2.1) pour les dépôts électromagnétiques (électron/photon), et 4% (paragraphe 4.5.2.1) pour les dépôts hadroniques (jets). Pour le muon, l'erreur mesurée de la trace associée est de 5% [16]. Nous associons une erreur de 4% à la mesure de l'énergie du neutrino. Nous remarquons d'après la table 5.13 que les erreurs sur l'échelle d'énergie constituent la source principale d'incertitude : elle est très importante pour les canaux contenant des jets et de façon générale pour les canaux à plus de 2 corps. Elle atteint 100% dans le canal e-j-j-j-j.
- **L'incertitude sur les angles polaires :**
  Elle a été mesurée pour les électrons, muons et jets dans la thèse [64]. Pour les dépôts électromagnétiques, une valeur de 5 mrad a été trouvée. L'erreur pour le muon ne vaut que 3 mrad. L'erreur systématique sur la mesure angulaire des jets a été déterminée dans des événements courant neutre en comparant l'angle polaire du jet à celui que l'on peut déduire à partir des grandeurs cinématiques de l'électron diffusé. Dans la région centrale, elle vaut en moyenne 10 mrad et est inférieure à 5 mrad à l'avant.
  La table 5.13 montre que la contribution de ce type de source d'erreurs est comprise entre 1,4% pour le canal e-γ et atteint 100% pour le canal e-j-j-j-j. Cette valeur importante est due à deux effets : le nombre d'événements prédits est faible et le nombre de jets produits à l'avant est très important. Donc, de toutes petites corrections des angles polaires des jets induisent des migrations d'événements dans ce canal. Nous remarquons que ce phénomène est prononcé dans tous les canaux à jets.
- **L'incertitude sur l'efficacité d'identification :**

Dans la partie 5.1.1, nous avons vu qu'il faut corriger de 5% les simulations Monte-Carlo pour décrire l'efficacité d'identification des électrons. L'erreur approximative de cette méthode est 2%. Pour les muons, l'incertitude sur l'efficacité d'identification est 5% suivant la publication [16]. Pour les photons l'erreur associée est de 2%.

- **L'incertitude de l'efficacité de déclenchement :**
  Dans la partie 5.2.4, nous avons exposé les efficacités de déclenchement associées aux muons, au neutrino et au jet. Dans le cas des muons, l'erreur a été obtenue en comparant l'accord de description de l'efficacité entre les données et les simulations Monte Carlo. Elle a été évaluée à 5% [12]. L'erreur sur l'efficacité de déclenchement pour les événements à impulsion manquante est estimée à 3%. Dans le cas de dépôts électromagnétiques, elle est très faible et négligeable par rapport aux autres sources d'erreurs systématiques. Elle n'est donc pas prise en compte.
- **L'erreur sur la mesure de luminosité :**
  L'incertitude sur les mesures de luminosité integrée est de 1,5%.

Corps	Echelle d'énergie / P_T	θ (mrad)	Efficacité d'identification	Efficacité de déclenchement
Electron	0,5%-3%	5	2 %	–
Muon	5%	3	5 %	5 %
Jet	4%	5-10	–	3%
Photon	0,5%-3%	5	2 %	–
Neutrino	4%	20	–	3 %

TAB. 5.11 – Résumé des erreurs systématiques associées aux différents corps.

5.4.3 Incertitudes théoriques

Nous discutons ici les incertitudes théoriques des processus pour lesquels nous disposons de simulations Monte-Carlo. Ces incertitudes (table 5.12) ne sont pas négligeables devant les erreurs expérimentales et les erreurs statistiques :

- **PYTHIA** : L'incertitude vient du fait que $PYTHIA$ génère les processus QCD uniquement à l'ordre dominant en α_s. Nous sommes alors amené à introduire un facteur appelé κ de 1,2 qui va être utilisé dans tous les canaux contenant au moins deux jets pour pondérer le spectre en P_T simulé : en effet, pour décrire les processus de photoproduction qui donnent deux jets dans l'état final, il faut corriger $PYTHIA$ par ce facteur de normalisation pour tenir compte des corrections d'ordres supérieurs. Comme cette procédure repose sur l'hypothèse que la forme du spectre est identique à tous les ordres, nous associons une erreur théorique de 10% à $PYTHIA$.
- **DJANGO, RAPGAP** : De le même manière que $PYTHIA$, ces générateurs n'incluent pas les ordres supérieurs en α_s dans les calculs. Nous associons une

incertitude de 10% à ces générateurs dans les canaux contenant au moins deux jets en sus d'un électron, et nous appliquons à ces canaux un facteur de renormalisation de $\kappa = 1,2$.

- **EPVEC** : Des études antérieures [65] ont montré que la section efficace du processus $ep \rightarrow eW^{\pm}X$ dépend de la fonction de structure du photon, du proton et en partie de l'échelle partonique en Q^2. La somme de ces erreurs vaut à peu près 30%, ce qui correspond à l'incertitude que nous associons au calcul de *EPVEC*.
- **GRAPE** : *GRAPE* calcule les diagrammes correspondant aux collisions $\gamma\gamma$ en tenant compte des termes d'interférence entre l'électron diffusé et les électrons produits. De plus, *GRAPE* inclut les processus Cabibbo-Parisi, de conversion interne de photon et de production de Z^o. L'interaction avec le proton est divisée en trois régions de l'espace de phase : une région élastique, une région quasiélastique et une région inélastique. L'incertitude attribuée à ce calcul, identique à l'analyse des états finaux à multi-électrons [11], est de 5%.
- **WABGEN** : *WABGEN* calcule les diagrammes des processus élastiques et quasiélastiques de "bremsstrahlung" à grand angle. Dans un premier temps, aucune incertitude systématique n'y a été associée.

Monte-Carlo	erreur
PYTHIA	10 %
CC, NC dans canaux multi-jets	10 %
GRAPE	5 %
EPVEC	30 %

TAB. 5.12 – Erreurs systématiques associées aux générateurs Monte-Carlo utilisés.

Canal	Erreur en %			
	Totale	Stat.	Theta	Energie
e-e	16.8	2.0	9.9	12.1
e-μ	14.4	1.8	1.5	11.0
e-j	8.2	0.4	4.0	6.7
e-ν	21.7	5.7	4.9	17.3
e-γ	4.1	2.0	1.4	2.6
μ-μ	21.3	2.1	1.6	17.8
μ-j	20.5	4.3	5.9	16.7
j-j	21.9	0.4	7.3	18.7
j-ν	4.9	0.4	1.6	3.1
j-γ	26.1	7.2	5.8	23.6
e-j-j	21.3	1.6	9.2	17.1
e-j-ν	30.5	10.6	12.3	5.2
e-j-γ	29.0	17.2	3.7	20.6
e-γ-γ	-	-	-	-
μ-j-ν	29.8	2.8	5.8	9.6
j-j-j	33.0	2.1	11.5	29.7
j-j-ν	28.4	3.3	13.6	23.6
j-j-γ	39.0	11.9	29.2	21.3
j-ν-γ	14.8	11.4	4.3	8.2
e-j-j-j	32.0	14.7	18.0	27.1
j-j-j-j	62.3	16.7	27.4	52.6
j-j-j-ν	33.9	14.3	21.0	25.2
e-j-j-j-j	163.0	75.3	100.0	100.0
j-j-j-j-j	-	-	-	-

TAB. 5.13 – Erreurs totales et les différentes contributions ajoutées quadratiquement pour les canaux analysés (données 1994-2000).

5.5 Résultats : données HERA I

Dans les parties précédentes, nous avons déterminé les efficacités d'identification de chaque corps, les erreurs statistiques, théoriques et systématiques ainsi que les efficacités de déclenchement dont nous tenons compte dans les différents canaux. Nous exposons dans cette partie les résultats obtenus avec l'ensemble des données e^+p 99-00, e^-p 98-99 et e^+p 94-97 cumulées entre 1994 et 2000 avec la sélection finale, résumée dans le tableau 5.14. Nous recherchons la signature de phénomènes de nouvelle physique qui se manifesteraient essentiellement à grand E_T. Dans un premier temps, nous donnons les taux globaux dans chaque canal, puis discutons l'accord au Modèle Standard en s'appuyant sur la variable $\Sigma E_T^{canal\ i}$. Elle correspond à la somme des énergies transverses de tous les corps du canal i et est donc appropriée pour l'étude des phénomènes à grand E_T.

Lot	Critères de qualité
Données/MC	- $\lvert z_{vtx}\rvert < 40$ cm - $E - P_z < 75$ GeV - Non activation des identificateurs de bruit de fond
Données	- Information temporelle : $\lvert T_{LAr}\rvert < 0{,}7$ & $\lvert T_{CJC}\rvert < 0{,}1$ - HT LAr, SpaCal, Lumi, TOF, CJC

Corps	Critères cinématiques
e, μ, j, ν, γ	$10^o < \theta_{corps} < 150^o$ $E_T^{corps} > E_T^{MIN} = 20$ GeV $R_{corps1}^{corps2} > R_{MIN} = 1{,}0$ radian

Canal	Critères supplémentaires anti-bruit
e-e	$20^o < \theta_{ele1,2} < 150^o$
e-ν, e-j-ν	$E - P_z < 45$ GeV, $V = \frac{V_{ap}}{V_p} < 0{,}5$
μ-μ	$\alpha < 165^o$ + visualisation
μ-j	visualisation
j-ν	$V = \frac{V_{ap}}{V_p} < 0{,}15$ + visualisation
j-γ	$cos(\theta_\gamma^*) > 0{,}0$ + visualisation

TAB. 5.14 – Ensemble des critères de sélection.

5.5.1 Les taux globaux

Pour chacun des 24 canaux analysés, nous avons comparé le nombre d'événements observés aux prédictions de tous les processus du Modèle Standard. La figure 5.18 montre les taux avec les erreurs statistiques et systématiques : les canaux à deux corps sont présentés en premier, puis les canaux à trois, quatre et cinq corps. Pour les canaux à faible statistique, les erreurs sont calculées suivant une distribution poissonienne.

Nous remarquons les taux importants dans les canaux e-j, j-ν et j-j dominés respectivement par les processus de courants neutres et chargés ainsi que par le processus de photoproduction. Il s'agit des processus dominants à HERA, à section efficace importante. Dans les canaux à trois corps ou plus et les canaux purement leptoniques, les taux sont naturellement plus faibles car les processus impliqués sont d'ordre supérieur. Pour les canaux e-γ-γ et j-j-j-j-j, pour lesquels les taux de simulations Monte-Carlo sont nuls, une limite supérieure d'événements prédits a été calculé : en supposant une distribution poissonienne des événements et en tenant compte de la luminosité simulée du processus dominant, une limite supérieure à 68% de niveau de confiance a été calculée.

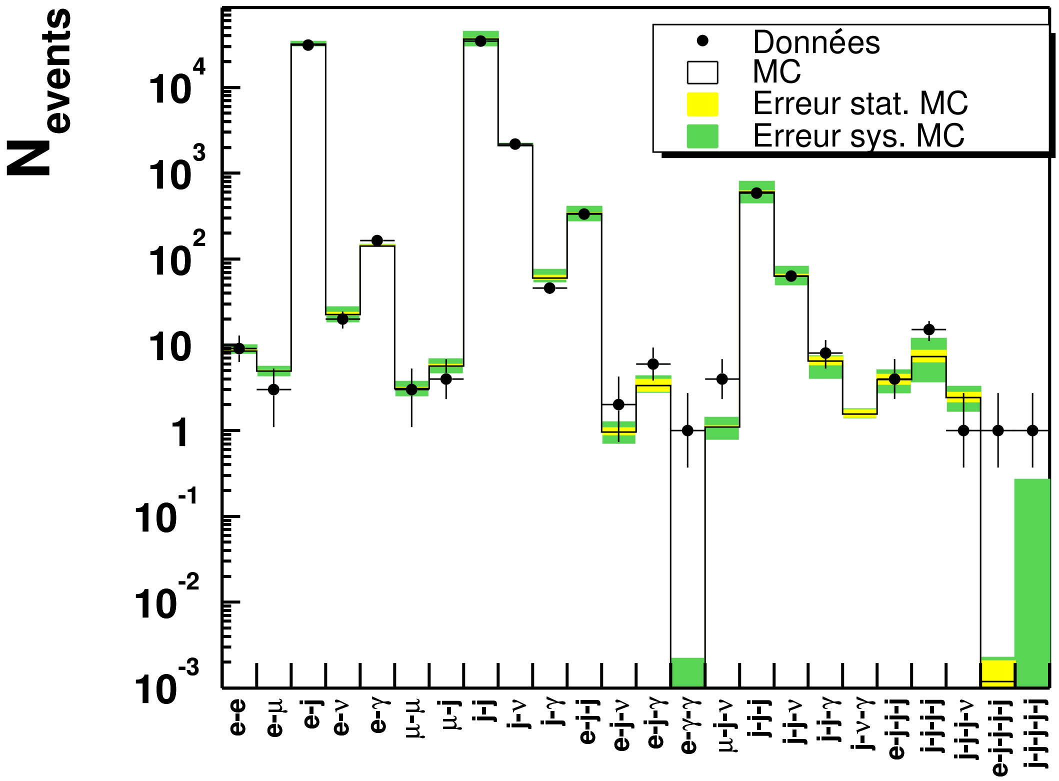


FIG. 5.18 – Taux d'événements par canal pour la totalité des données prises entre 1994 et 2000.

Pour mieux illustrer la comparaison aux prédictions, nous avons étudié le rapport R_{acc} défini par :

$$R_{acc} = \frac{N_{data} - N_{MC}}{N_{MC}} \tag{5.6}$$

où N_{data} est le taux d'événements observés et N_{MC} le taux d'événements prédits.

Dans la majorité des canaux, comme le montre la figure 5.19, les mesures sont en bon accord avec les prédictions du Modèle Standard dans la limite des erreurs calculées. Néanmoins, quelques déviations sont visibles : notamment dans les canaux e-γ-γ, μ-j-ν, e-j-j-j-j et j-j-j-j-j qui seront discutés par la suite.

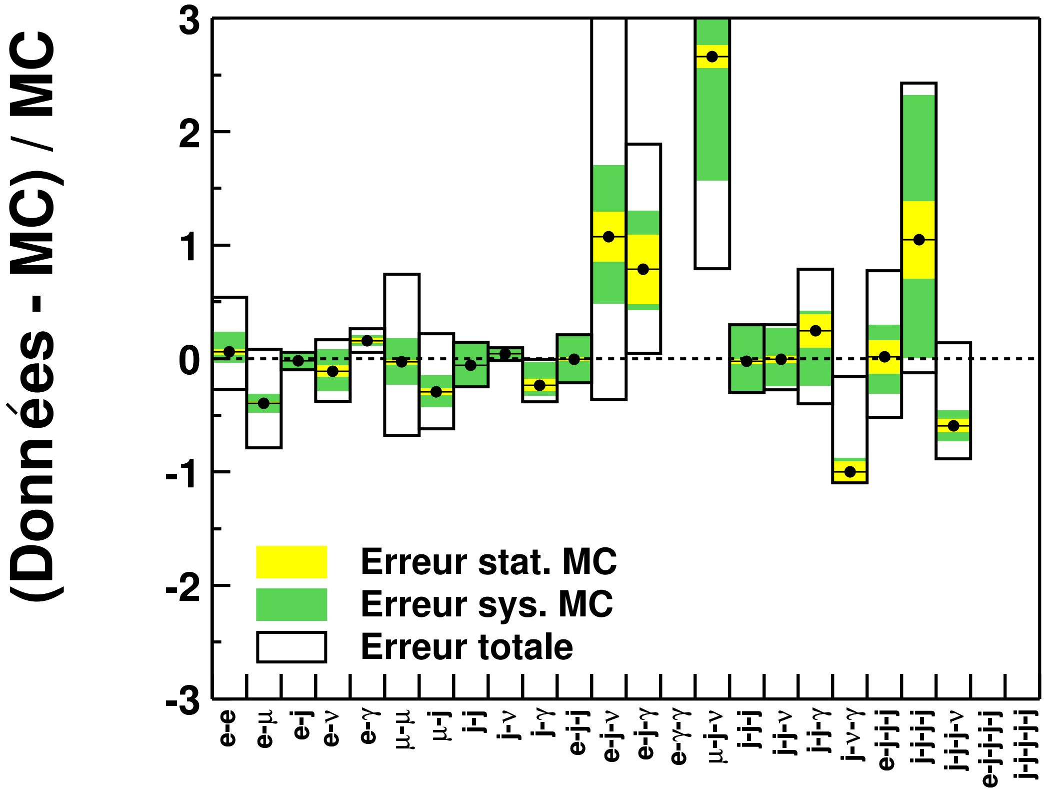


FIG. 5.19 – Taux d'événements observés ramenés aux taux d'événements prédits pour la totalité des données prises entre 1994 et 2000. Les erreurs statistiques du Monte-Carlo et systématiques sont montrées. L'erreur totale représentée inclut de plus l'erreur statistique correspondant au nombre d'événements observés dans les données. Les barres correspondantes aux canaux e-γ-γ, e-j-j-j-j et j-j-j-j-j se trouvent en dehors de la figure, l'écart au Monte-Carlo étant supérieur à 3.

Pour tenir compte des processus du Modèle Standard et au-delà dont la section efficace dépend fortement du signe du lepton incident (positron ou électron), nous avons étudié indépendamment les deux lots e^+p 94-97/99-00 et e^-p 98-99, dont les taux sont reportés dans les tableaux 5.15 et 5.16. Nous obtenons un bon accord global pour les lots e^+p 94-97/99-00 et e^-p 98-99, dans lequel, en raison de la faible luminosité intégrée de 14 pb^{-1}, les fluctuations statistiques sont importantes.
Nous remarquons que l'unique événement e-γ-γ est enregistré dans les données e^+p 94-97/99-00. Les déviations dans le canal μ-j-ν sont concentrées également dans le lot e^+p 94-97/99-00. Dans le canal j-j-j-j, un excès est visible dans le lot e^+p 94-97/99-00 qui n'est pas confirmé par les résultats du lot e^-p 98-99. Les seuls événements des canaux e-j-j-j-j et j-j-j-j-j sont enregistrés dans les données e^+p 94-97/99-00.

Canal	Données	MC ± stat. ± sys.
e-e	6.	$7.4 \pm 0.1 \, ^{+0.6}_{-1.4}$
e-μ	3.	$4.3 \pm 0.1 \, ^{+0.6}_{-0.6}$
e-j	27580.	$27934.2 \pm 114.1 \, ^{+2324.5}_{-2124.7}$
e-ν	19.	$19.5 \pm 1.1 \, ^{+3.7}_{-4.0}$
e-γ	148.	$119.9 \pm 2.6 \, ^{+4.7}_{-4.2}$
μ-μ	3.	$2.7 \pm 0.1 \, ^{+0.6}_{-0.6}$
μ-j	4.	$5.0 \pm 0.2 \, ^{+0.9}_{-1.0}$
j-j	30524.	$32174.9 \pm 121.3 \, ^{+6453.1}_{-7063.7}$
j-ν	1691.	$1618.4 \pm 6.8 \, ^{+82.7}_{-78.6}$
j-γ	40.	$54.3 \pm 3.7 \, ^{+6.0}_{-13.0}$
e-j-j	297.	$294.8 \pm 4.8 \, ^{+58.6}_{-62.5}$
e-j-ν	2.	$0.8 \pm 0.1 \, ^{+0.2}_{-0.3}$
e-j-γ	4.	$3.0 \pm 0.5 \, ^{+0.3}_{-0.7}$
e-γ-γ	1.	-
μ-j-ν	4.	$1.0 \pm 0.0 \, ^{+0.3}_{-0.3}$
j-j-j	498.	$524.5 \pm 10.7 \, ^{+145.3}_{-172.5}$
j-j-ν	51.	$52.2 \pm 1.8 \, ^{+12.6}_{-15.1}$
j-j-γ	6.	$5.5 \pm 0.7 \, ^{+1.9}_{-0.5}$
j-ν-γ	0.	$1.0 \pm 0.2 \, ^{+0.0}_{-0.2}$
e-j-j-j	3.	$3.5 \pm 0.5 \, ^{+1.0}_{-0.9}$
j-j-j-j	14.	$6.3 \pm 1.0 \, ^{+3.0}_{-3.7}$
j-j-j-ν	0.	$2.0 \pm 0.3 \, ^{+0.7}_{-0.6}$
e-j-j-j-j	1.	$0.0 \pm 0.0 \, ^{+0.0}_{-0.0}$
j-j-j-j-j	1.	-

TAB. 5.15 – Taux d'événements prédits et observés par canal avec les erreurs systématiques et statistiques associées pour le lot e^+p 94-97/99-00.

Par la suite, nous allons montrer et discuter les résultats dans chacun des canaux (i) en nous appuyant sur les distributions en ΣE_T^i. Pour chaque canal, les prédictions des processus du Modèle Standard seront comparées aux données. Nous allons insister sur les événements reconstruits à grand ΣE_T^i, et lorsqu'il y a lieu, donner des interprétations théoriques pour des événements atypiques. Les canaux ont été classés en six catégories selon le processus dominant du Modèle Standard. L'ensemble des distributions d'impulsion transverse et d'angle polaire pour les canaux ayant un taux d'événements prédits supérieur à 10^{-3} sont regroupées dans l'appendice B.

Channel	Data	MC ± stat. ± sys.
e-e	3.	$1.1 \pm 0.1\ ^{+0.1}_{-0.1}$
e-μ	0.	$0.7 \pm 0.0\ ^{+0.1}_{-0.1}$
e-j	3811.	$4050.4 \pm 21.5\ ^{+299.7}_{-297.5}$
e-ν	1.	$3.1 \pm 0.2\ ^{+0.6}_{-0.7}$
e-γ	17.	$22.8 \pm 0.8\ ^{+0.6}_{-0.9}$
μ-μ	0.	$0.4 \pm 0.0\ ^{+0.1}_{-0.1}$
μ-j	0.	$0.7 \pm 0.04\ ^{+0.14}_{-0.13}$
j-j	4037.	$4712.4 \pm 25.5\ ^{+941.8}_{-1029.3}$
j-ν	506.	$492.2 \pm 2.2\ ^{+20.1}_{-19.7}$
j-γ	6.	$5.9 \pm 0.7\ ^{+0.5}_{-2.2}$
e-j-j	38.	$43.2 \pm 0.9\ ^{+8.5}_{-9.3}$
e-j-ν	0.	$0.1 \pm 0.0\ ^{+0.0}_{-0.0}$
e-j-γ	2.	$0.4 \pm 0.1\ ^{+0.0}_{-0.1}$
e-γ-γ	0.	-
μ-j-ν	0.	$0.1 \pm 0.0\ ^{+0.0}_{-0.0}$
j-j-j	90.	$80.6 \pm 2.3\ ^{+22.3}_{-27.0}$
j-j-ν	12.	$11.4 \pm 0.4\ ^{+2.3}_{-2.8}$
j-j-γ	2.	$1.0 \pm 0.1\ ^{+0.5}_{-0.1}$
j-ν-γ	0.	$0.6 \pm 0.1\ ^{+0.0}_{-0.0}$
e-j-j-j	1.	$0.5 \pm 0.1\ ^{+0.1}_{-0.1}$
j-j-j-j	1.	$1.0 \pm 0.2\ ^{+0.5}_{-0.7}$
j-j-j-ν	1.	$0.5 \pm 0.1\ ^{+0.1}_{-0.1}$
e-j-j-j-j	0.	$0.0 \pm 0.0\ ^{+0.0}_{-0.0}$
j-j-j-j-j	0.	-

TAB. 5.16 – Taux d'événements prédits et observés par canal avec les erreurs sytématiques et statistiques associées pour le lot e^-p 98-99.

5.5.2 Canaux de processus à di-leptons

Nous classons les canaux e-e, e-μ, μ-μ et μ-j dans cette catégorie.
Le canal e-e est dominé à plus de 93% par le processus de production de paires d'électrons. Les contaminations résiduelles, dues à une mauvaise identification, sont les courants neutres (<1%) et le processus de diffusion QED Compton (6%). Dans les événements courant neutre, des pions neutres se désintégrant dans la section électromagnétique du calorimètre en deux photons, ou d'autres parties du jet peuvent simuler la présence d'un électron. L'identification de deux électrons de grand E_T dans des processus de QED Compton est due à la conversion du photon en deux électrons très proches qui ne sont pas discriminés par les trajectographes.

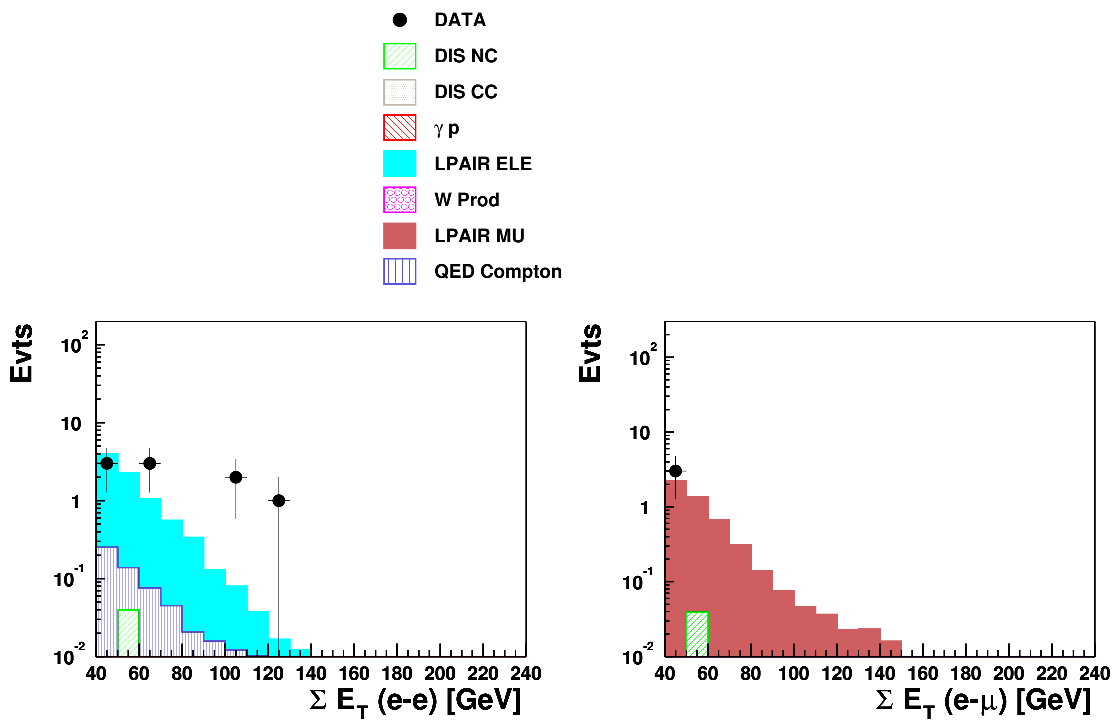


FIG. 5.20 – Distributions de ΣE_T^{e-e} et $\Sigma E_T^{e-\mu}$.

La distribution à gauche sur la figure 5.20 montre le spectre en ΣE_T^{e-e}. Nous observons un bon accord global avec 9 événements observés pour 8,5 $\pm$ 1,5 prédits. Néanmoins, trois événements avec $\Sigma E_T^{e-e} > 100$ GeV se détachent. Ces événements ont également des masses invariantes au-dessus de 100 GeV et ont été trouvés auparavant dans l'analyse dédiée aux multi-électrons [11]. Un événement parmi ces trois est représenté en vue longitudinale dans la figure 5.21. Nous ne trouvons aucun événement dans le canal e-e-e, dont l'un serait l'électron diffusé, avec des E_T au-dessus de 20 GeV. La prédiction est en-dessous de 10^{-3}. Dans l'analyse dédiée aux multi-électrons [11], trois événements contenant trois électrons sont reconstruits avec une masse invariante supérieure à 100 GeV. Ces événements ne sont pas identifiés dans notre analyse car au moins un des électrons ne satisfait pas les critères d'impulsion transverse minimale.

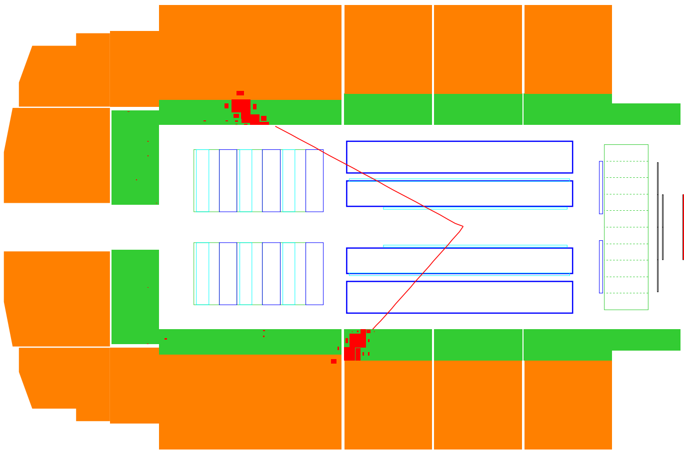


FIG. 5.21 – Vue longitudinale d'un événement dans le canal e-e à grand ΣE_T^{e-e}.

Dans le canal e-μ, nous trouvons trois événements reconstruits pour 4,9 $\pm$ 0,7 prédits. La prédiction est constituée à 98% de processus de production de paires de muons, dans lequel l'un des deux muons produits ne satisfait pas les critères cinématiques et dans lequel nous reconstruisons l'électron incident diffusé. 2% des événements sont des courants neutres dans lesquels nous reconstruisons un muon issu d'une désintégration d'une particule du jet. Le spectre en $\Sigma E_T^{e-\mu}$ (distribution à droite sur la figure 5.20) ne nous indique pas de production de particule lourde dans ce canal.
Le canal μ-μ est dominé à 100 % par la production élastique et inélastique de paires de muons dans les collisions $\gamma\gamma$ avec 3 événenements observés pour 3,1 $\pm$ 0,6 prédits. Dans l'analyse [12], dans laquelle les énergies transverses minimales des muons sont de 2 GeV, il a été montré que la contribution venant d'autres processus ($Z^o \rightarrow \mu\mu$, $c\bar{c}, b\bar{b} \rightarrow \mu\mu$ ...) est inférieure à 1%. Elle est donc négligeable dans notre cas. La distribution $\Sigma E_T^{\mu-\mu}$ à gauche sur la figure 5.22 montre un bon accord entre les données et les prédictions du Modèle Standard.

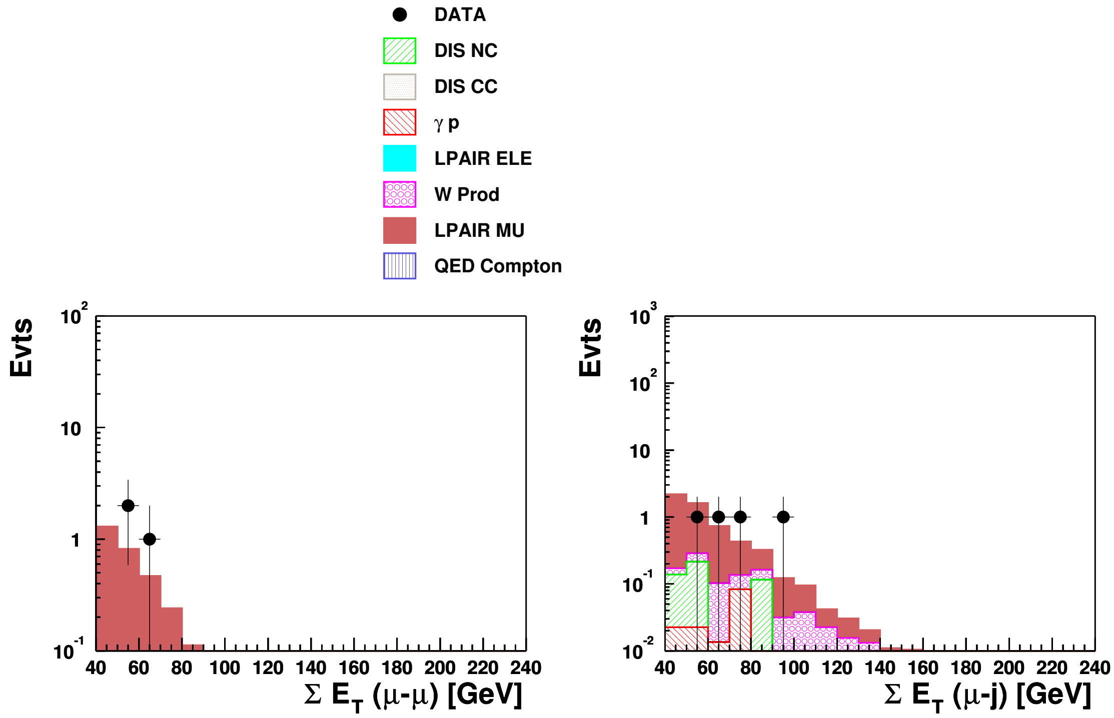


FIG. 5.22 – Distributions de $\Sigma E_T^{\mu-\mu}$ et $\Sigma E_T^{\mu-j}$.

Un événement est reconstruit avec un $\Sigma E_T^{\mu-\mu} = 68{,}3$ GeV et une masse invariante de 80 GeV. La figure 5.23 montre la vue longitudinale de cet événement.

Dans le canal μ-j, nous observons 4 événements pour 5,7 $\pm$ 1,2 prédits. Les contributions majeures sont bien sûr la production inélastique de paires de muons (82%), où l'un des muons a échappé à la sélection, et la production de boson $W^\pm$ (8%), où le neutrino ne passe pas les critères de sélection. Nous avons également des contributions du processus de courant neutre (8 %) et de photoproduction (2%) où des quarks lourds issus de jets se sont désintégrés en muons. La distribution $\Sigma E_T^{\mu-j}$ de la figure 5.22 montre que nous obtenons un bon accord entre les données et les prédictions. La figure 5.24 visualise l'événement observé de plus grande énergie transverse.

En résumé, nous observons un bon accord global dans les quatre canaux e-e, e-μ, μ-μ et μ-j. Néanmoins, dans le canal e-e, trois événements à une masse invariante au-delà de 100 GeV sont trouvés. Une explication possible avancée était la désintégration d'un boson de Higgs exotique en paire d'électrons étudiée dans [66], mais les dernières limites préliminaires [67] obtenues par l'expérience OPAL rendent cette hypothèse très peu probable.

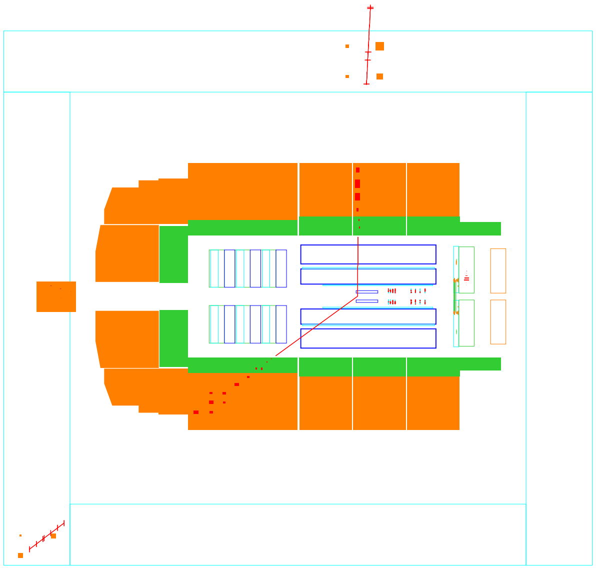


FIG. 5.23 – Vue longitudinale d'un événement dans le canal μ-μ avec $\Sigma E_T^{\mu-\mu} = 68$ GeV.

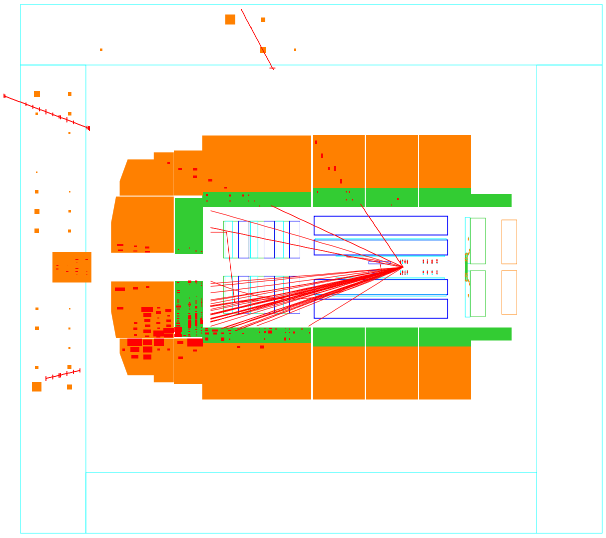


FIG. 5.24 – Vue longitudinale d'un événement du canal μ-j avec $\Sigma E_T^{\mu-j} = 91$ GeV. Le jet est projeté vers l'avant alors que le muon est identifié dans la région centrale. Un second muon de basse impulsion transverse est visible dans la région avant.

5.5.3 Canaux de leptons isolés avec grande impulsion transverse manquante

Nous classons les canaux e-ν, e-j-ν et μ-j-ν parmi cette catégorie. Les topologies typiques de la production de boson $W^{\pm}$ avec désintégration leptonique correspondent à ces canaux. La production de $W^{\pm}$ se manifeste préférentiellement dans le canal e-ν car, la plupart du temps, l'énergie transverse du jet est inférieure au critère de sélection de 20 GeV.

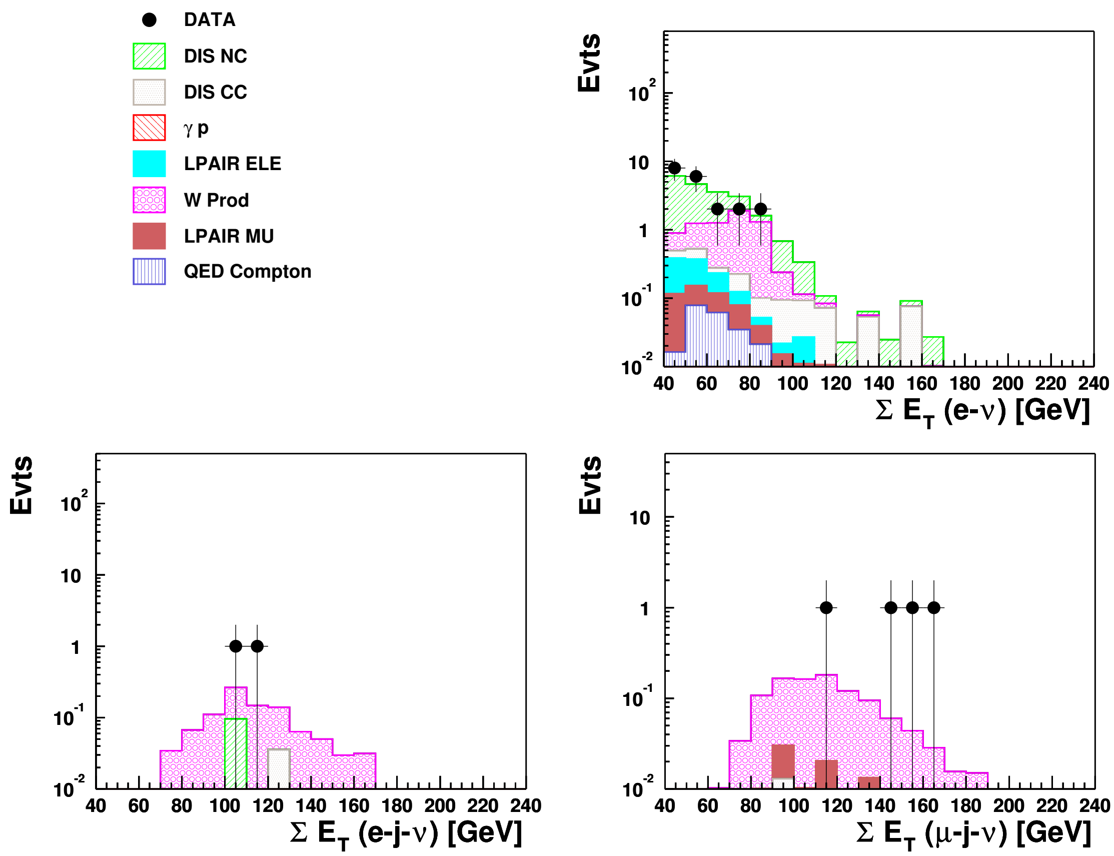


FIG. 5.25 – Distributions de $\Sigma E_T^{e-\nu}$, $\Sigma E_T^{e-j-\nu}$ et $\Sigma E_T^{\mu-j-\nu}$.

Dans ce canal, dont la distribution en $\Sigma E_T^{e-\nu}$ est montrée en haut à droite de la figure 5.25, 20 événements sont observés pour 22,5 $\pm$ 4,8 prédits. Les contributions principales sont alors la production de $W^{\pm}$ (22%), les courants neutres (65%) et chargés (4%). Le taux de mauvaise identification de courants neutres est important surtout à basse énergie transverse manquante, c'est-à-dire pour E_T^{ν} compris entre 20 et 25 GeV : la plupart du temps, il s'agit d'une reconstruction erronée de l'impulsion transverse de l'état hadronique qui simule de l'énergie manquante. La contribution de courants chargés est concentrée à l'avant du détecteur : en effet, des pions neutres issus du jet peuvent se désintégrer dans la section électromagnétique du calorimètre LAr. Nous n'observons aucun événement au-delà de 100 GeV en $\Sigma E_T^{e-\nu}$.

Nous retrouvons la même topologie dans le canal e-j-ν, où le jet est reconstruit à grand E_T. La reconstruction explicite de celui-ci dans ce canal réduit considérablement le bruit de fond venant d'une mauvaise identification : dans l'analyse dédiée [16], la coupure sur l'impulsion transverse de l'état hadronique final à plus de 12 GeV remplit la même fonction. Finalement, nous observons, comme le montre la distribution $\Sigma E_T^{e-j-\nu}$ (distribution en bas à gauche de la figure 5.25), 2 événements pour 1,0 $\pm$ 0,3 prédits. Le processus dominant est la production de $W^{\pm}$ avec 84%. La contamination la plus importante vient des courants neutres (10 %). La vue longitudinale de l'un des événements observés est montrée dans la figure 5.26.

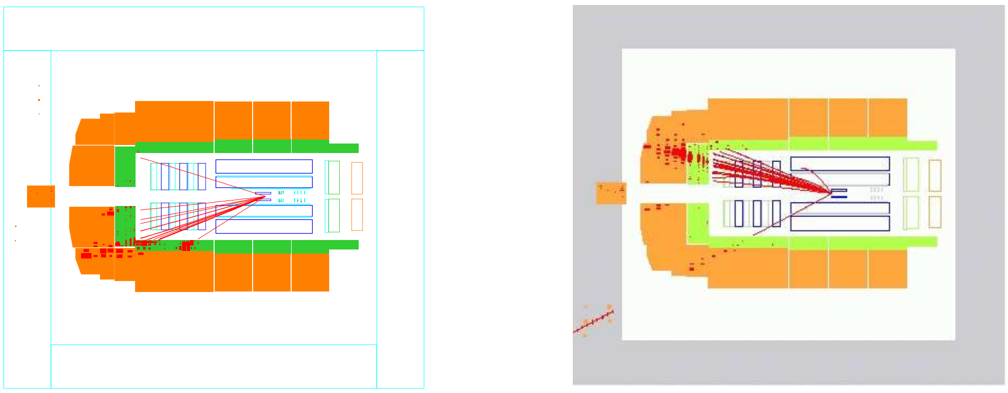


FIG. 5.26 – Vue longitudinale d'un événement du canal e-j-ν (à gauche) et d'un événement dans le canal μ-j-ν (à droite).

Le canal μ-j-ν est similaire au canal e-j-ν car les prédictions théoriques sont égales. Seules les contributions de bruit de fond diffèrent. Ainsi, μ-j-ν est dominé par la production de boson $W^{\pm}$ (89%). Il subsiste de faibles contributions de production inélastiques de paire de muons (10%) pour lesquels l'un des muons a échappé à la détection. Des contributions négligeables viennent des courants chargés (moins de 1%) où des quark lourds se sont désintégrés en muons. La distribution en $\Sigma E_T^{\mu-j-\nu}$ en bas à droite de la figure 5.25 montre que nous reconstruisons 4 événements pour 1,1 $\pm$ 0,3 prédits. Deux de ces événements ont des $\Sigma E_T^{\mu-j-\nu}$ au-dessus de 150 GeV. La vue longitudinale de l'un d'entre eux est montrée dans la figure 5.26 (le muon est reconstruit à l'avant).

Finalement, nous avons vérifié que les événements interessants avec P_T^{X}[1] > 25 GeV, tabulés dans [16], sont bien retrouvés dans la présente analyse (voir tableau 5.17). Quatre parmi six événements à muons isolés sont sélectionnés dans le canal μ-j-ν. Les deux autres sont reconstruits dans les canaux μ-ν et j-ν : pour le premier, l'angle

[1] P_T^X est l'impulsion transverse de tous les amas calorimétriques, sauf ceux associés à l'électron et au muon.

Leptons isolés

run	événement	état	canal	critères de "rejet"
236176	3849 (e)	sélectionné	j-ν	$P_T^e = 10$ GeV
186729	702 (μ)	sélectionné	μ-j-ν	-
188108	5066 (μ)	sélectionné	μ-ν	$\theta^{jet} = 9,6°$
192227	6208 (μ)	sélectionné	μ-j-ν	-
195308	16793 (μ)	sélectionné	μ-j-ν	-
248207	32134 (e)	sélectionné	e-j-ν	-
252020	30485 (e)	sélectionné	e-j-ν	-
266336	4126 (μ)	sélectionné	j-ν	$P_T^\mu = $ 19,7 GeV
268338	70014 (e)	sélectionné	j-ν	Qualité de trace de l'électron insuffisante
270132	73115 (μ)	sélectionné	μ-j-ν	-
275991	29613 (e)	non sélectionné	-	$\theta^{jet} = 5°$ Qualité de trace de l'électron insuffisante

TAB. 5.17 – Evénements dits à "leptons isolés" sélectionnés dans l'analyse [16] avec $P_T^X > 25$ GeV. P_T^X est l'impulsion transverse de tous les amas calorimétriques, sauf ceux associés à l'électron et au muon.

polaire du jet (θ_{jet} = 9,6^o) est inférieur à notre critère de 10^o, et pour le second, l'impulsion transverse du muon (P_T^μ = 19,6 GeV) est inférieure à 20 GeV.
Parmi les cinq événements dits à électrons isolés de l'analyse [16], seuls deux sont reconstruits dans le canal e-j-ν : parmi les trois restants, deux sont reconstruits, comme nous pouvons nous convaincre d'après le tableau 5.17, dans le canal j-ν car l'électron ne satisfait pas nos critères de sélection de 20 GeV en impulsion transverse ou de qualité sur la trace associée ; pour le dernier, les critères de qualité de la trace associée à l'électron et l'angle polaire du jet (θ_{jet} = 5^o) sont insuffisants pour être retenus dans notre analyse.

En résumé, nous observons dans les deux canaux e-ν et e-j-ν une bonne description de la distribution ΣE_T par les prédictions. Dans le canal μ-j-ν, plusieurs événements intéressants sont reconstruits à $\Sigma E_T^{\mu-j-\nu} > 150$ GeV.

5.5.4 Canaux à courant neutre

Les processus à courant neutre sont reconstruits principalement dans les canaux e-j, e-j-j, e-j-j-j et e-j-j-j-j.
Dans le canal e-j, la fraction de courants neutres est supérieure à 99,7% avec en tout 31391 événements observés pour 31984 $\pm$ 2622 prédits. Des contributions supplémentaires viennent d'événements de photoproduction (0,1%) dans lesquels l'un des jets a été identifié comme électron (essentiellement pour des angles polaires $\theta < 20^o$), d'événements de production inélastique de paires de leptons où l'un des leptons (où les deux pour la production de paires de muons) a échappé à la détection (0,1%).

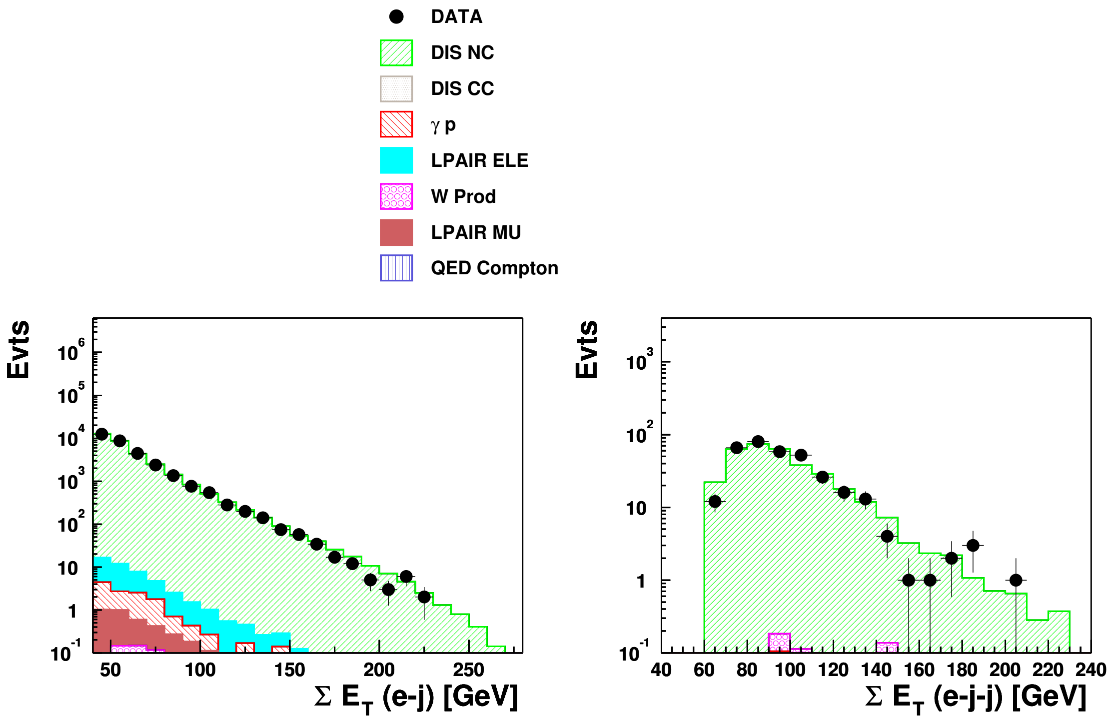


FIG. 5.27 – Distributions de ΣE_T^{e-j} et ΣE_T^{e-j-j}.

Nous obtenons la distribution en ΣE_T^{e-j} (à gauche de la figure 5.27) montrant un très bon accord entre les données et les prédictions Monte-Carlo jusqu'à des ΣE_T^{e-j} supérieurs à 225 GeV. Nous remarquons un léger déficit peu significatif de données pour $175 < \Sigma E_T^{e-j} < 215$ GeV, déjà identifié auparavant dans l'analyse [29]. L'événement de plus grand ΣE_T^{e-j} est montré dans la figure 5.28.

Comme nous l'avons vu au chapitre 1, le canal e-j-j est tout à fait identique au canal e-j de par son processus, à la correction radiative QCD près. Il permet d'étudier le processus de DIS à l'ordre supérieur (NLO). La distribution ΣE_T^{e-j-j} (à droite de la figure 5.27) montre un accord global entre les données et les prédictions Monte-Carlo, dont la majorité vient de courants neutres (99,3%) avec 335

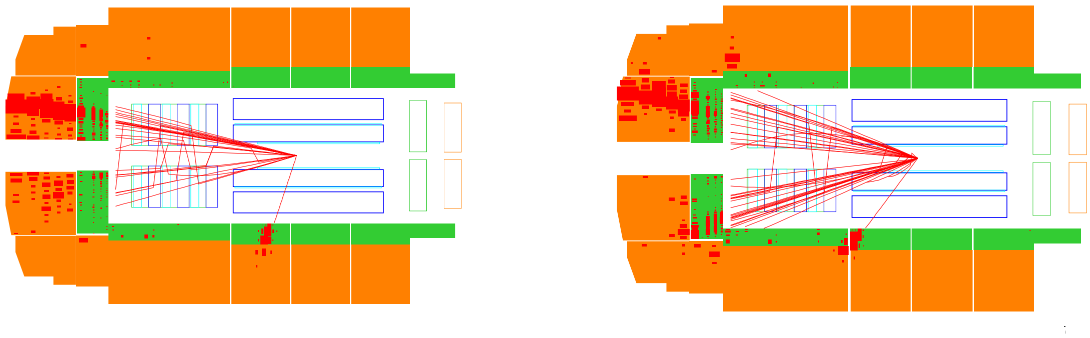


FIG. 5.28 – Vue longitudinale des événements du canal *e-j* (à gauche) et *e-j-j* (à droite) de plus grande énergie transverse.

événenements observés pour 337 ± 70 prédits. Nous retrouvons les mêmes contributions minoritaires (0,7%) d'autres processus que dans le canal *e-j*. L'événement ayant $\Sigma E_T^{e-j-j} > 200$ GeV est aussi retrouvé à grande masse invariante. Une vue longitudinale en est montrée dans la figure 5.28.
Le canal à quatre corps *e-j-j-j* est plus difficile à analyser car d'une part, les prédictions de taux du Modèle Standard sont faibles, ce qui nécessite donc une statistique Monte-Carlo très importante et engendre une erreur statistique importante de plus de 30%, et d'autre part les prédictions exactes ne sont pas connues : les générateurs Monte-Carlo utilisés n'incluent pas les corrections d'ordre supérieur NLO. Nous retrouvons, comme cela est montré sur la figure 5.29, 4 événements pour 3,94 ± 1,3 prédits avec un événement ayant $\Sigma E_T^{e-j-j-j} > 180$ GeV. Néanmoins les prédictions nulles dans cette région de l'espace de phases ne nous permettent pas de conclure. Dans le canal *e-j-j-j-j*, un événement est reconstruit pour une prédiction d'environ 10^{-3} de processus de production de $W^{\pm}$. Une vue longitudinale d'un événement typique du canal *e-j-j-j* et de l'événement du canal *e-j-j-j-j* est montrée dans la figure 5.30.

En résumé, nous avons vu que les spectres dans les canaux *e-j*, *e-j-j*, *e-j-j-j* et *e-j-j-j-j* sont bien décrits dans l'ensemble. Les erreurs systématiques et statistiques dans les canaux *e-j-j*, *e-j-j-j* et *e-j-j-j-j* sont importantes (jusqu'à 160%). Nous avons relevé des événements intéressants à grand ΣE_T, mais leur contribution dans les canaux *e-j* et *e-j-j* est prédite par les prédictions du Modèle Standard. Dans les canaux *e-j-j-j* et *e-j-j-j-j*, les incertitudes totales importantes, tant expérimentales que théoriques, ne permettent pas de valider la déviation observée. Pour deux raisons, les deux derniers canaux n'ont jamais fait l'objet d'études détaillés auparavant : d'une part, il faudrait générer un nombre très important d'événements, et d'autre part, les corrections radiatives ne sont pas inclues dans les générateurs.

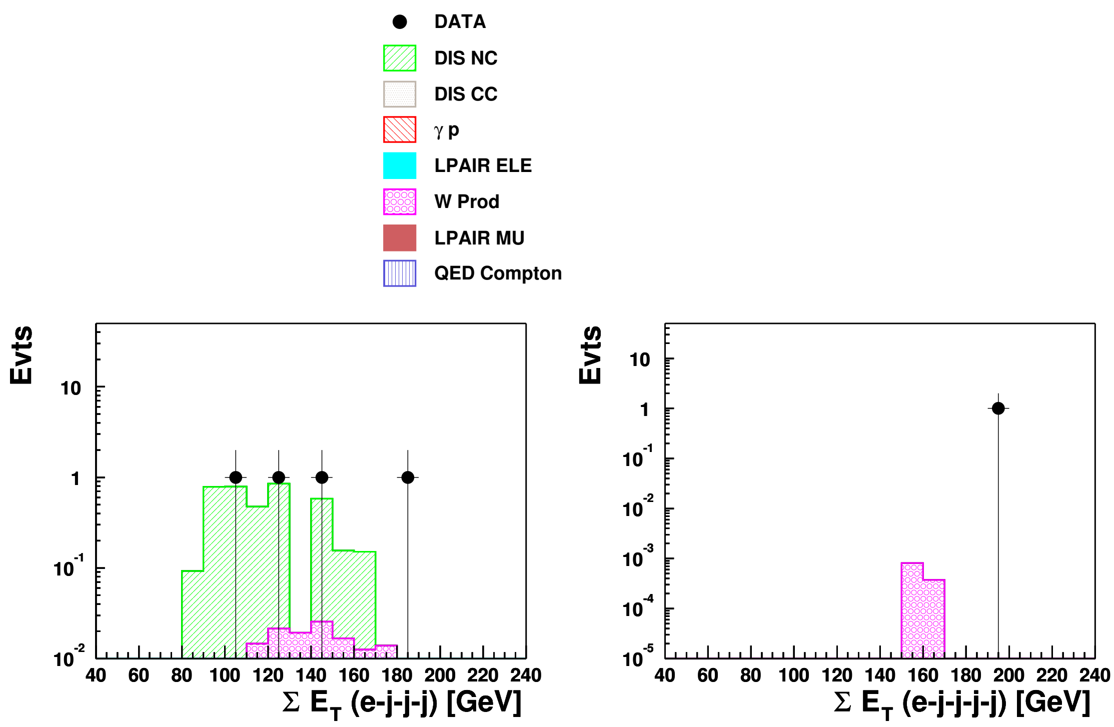


FIG. 5.29 – Distributions de $\Sigma E_T^{e-j-j-j}$ et $\Sigma E_T^{e-j-j-j-j}$.

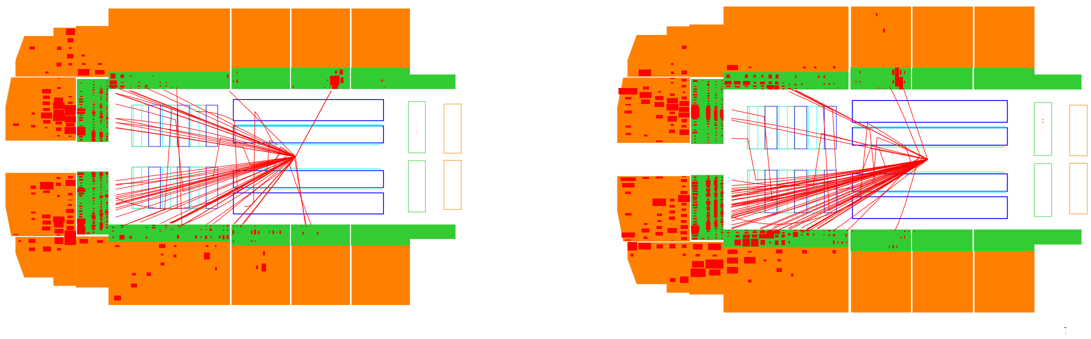


FIG. 5.30 – Vue longitudinale d'un événement du canal e-j-j-j (à gauche) et de l' événement du canal e-j-j-j-j (à droite).

5.5.5 Canaux à courant chargé

De la même manière que pour les courants neutres, nous trouvons la topologie classique de courant chargé dans certains canaux. Il s'agit des canaux contenant au moins un jet et de l'énergie transverse manquante identifiée comme un neutrino : ce sont les canaux j-ν, j-j-ν et j-j-j-ν. Le canal j-ν est dominé par des courants chargés (2197 événements observés pour 2110 $\pm$ 105 prédits) à plus de 98%. Nous trouvons aussi les contributions de production de boson $W^{\pm}$ (contribuant à 0,3%) où l'un des leptons n'est pas détecté (d'après nos critères de sélection) et celle des processus de photoproduction et de courant neutre, où la mesure erronée des jets (et/ou de l'électron pour les NC) simulent de l'énergie transverse manquante (contribution totale de 1%). La distribution $\Sigma E_T^{j-\nu}$ de la figure 5.31 est bien décrite par les prédictions du Modèle Standard jusqu'à des valeurs de 250 GeV. Seul un déficit notable de 8 événements observés pour 17 prédits est visible pour $160 < \Sigma E_T^{j-\nu} <$ 170 GeV.

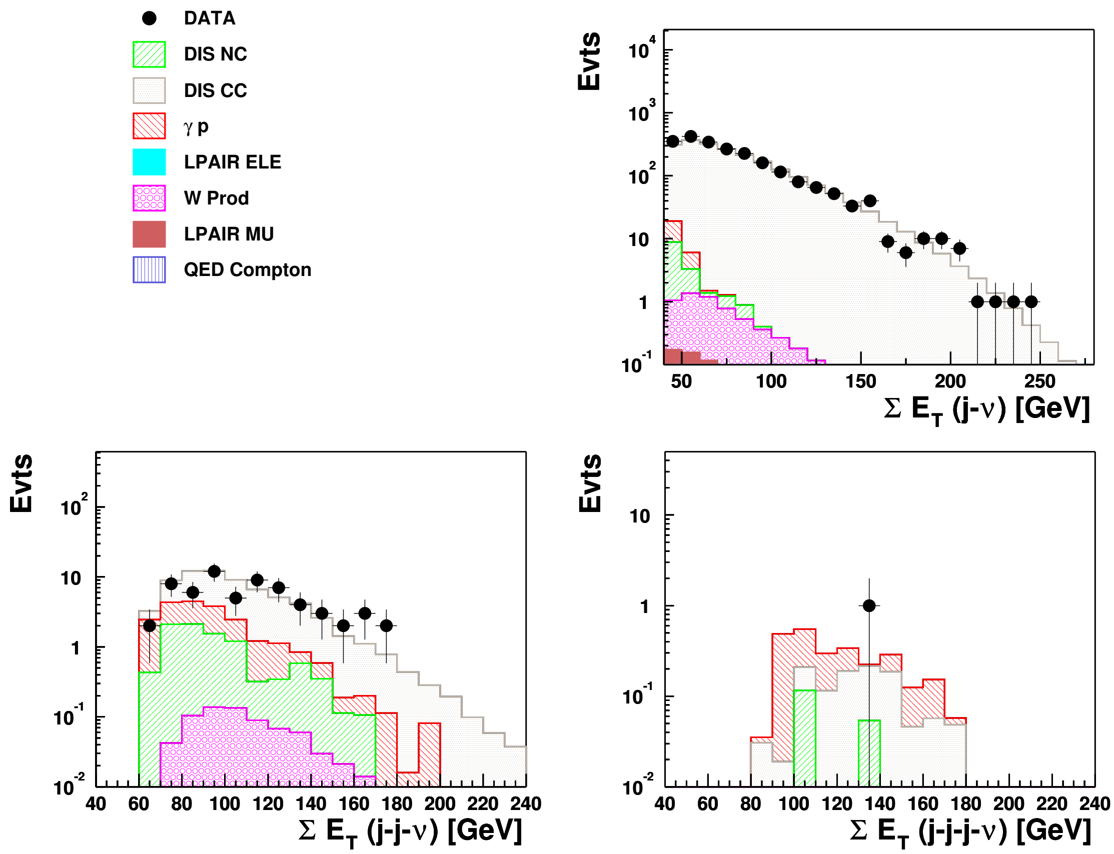
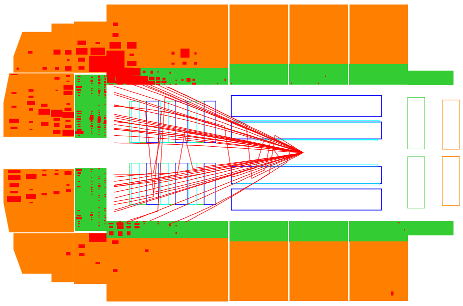


FIG. 5.31 – Distributions de $\Sigma E_T^{j-\nu}$, $\Sigma E_T^{j-j-\nu}$ et $\Sigma E_T^{j-j-j-\nu}$.

Un événement est trouvé avec un $\Sigma E_T^{j-\nu} > 240$ GeV (pour 0,6 prédit) ; comme nous pouvons le voir sur la figure 5.32, le jet y est projeté à l'avant.

Dans le canal j-j-ν, les contributions dues à une mauvaise identification sont



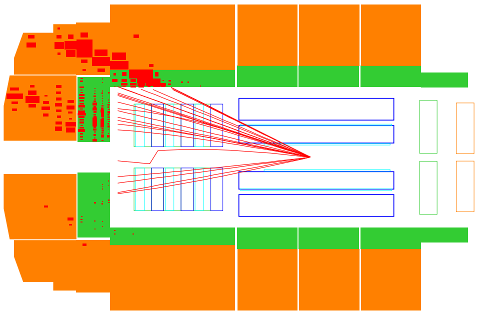


FIG. 5.32 – Vue longitudinale des événements du canal j-ν (à gauche) et j-j-ν (à droite) de plus grande énergie transverse. Dans l'événement j-ν, des jets supplémentaires à basse impulsion transverse sont reconstruits à l'avant.

un peu plus élevées. Nous observons 63 événements pour 63 $\pm$ 18 prédits dont 65% de courants chargés. Les contributions résiduelles viennent essentiellement des courants neutres avec 11% et du processus de photoproduction (21%). La distribution $\Sigma E_T^{j-j-\nu}$ montre un bon accord entre les données et les prédictions (figure 5.31). Aucun événement ne se détache à $\Sigma E_T^{j-j-\nu} > 200$ GeV. L'événement de plus grand $\Sigma E_T^{j-j-\nu}$ est montré sur la figure de droite de 5.32.

Dans le canal j-j-j-ν, nous retrouvons exactement un événement pour 2,4 $\pm$ 0,8 prédits dont 33% de courants chargés. Ce canal, affecté par une forte contamination d'événements de photoproduction (55%), est dominé par des erreurs statistiques et systématiques importantes : l'erreur totale est de 34%. Nous remarquons d'après la distribution $\Sigma E_T^{j-j-j-\nu}$, que l'événement observé et montré sur la figure 5.33, se situe dans une région attendue de l'espace de phase.

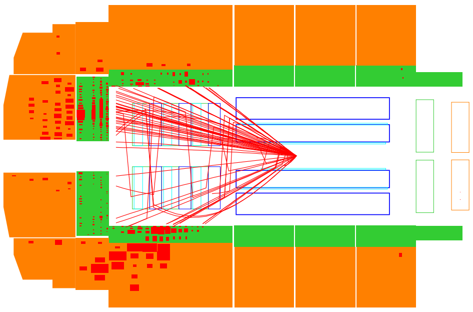


FIG. 5.33 – Vue longitudinale de l'événement du canal j-j-j-ν.

En résumé, nous avons observé un bon accord entre prédictions et données dans les canaux j-ν et j-j-ν. Dans le canal j-j-j-ν, le faible taux prédit de 2,4 événements est tout à fait compatible avec l'événement observé. Nous n'observons pas d'écart important par rapport aux prédictions.

5.5.6 Canaux à photons

Nous classons les canaux radiatifs e-j-γ et j-ν-γ, les canaux j-γ et j-j-γ ainsi que les deux canaux e-γ et e-γ-γ de diffusion QED Compton à grand E_T parmi cette catégorie.

Le canal e-γ est bien sûr dominé par les processus de diffusion QED Compton (78%), mais la contribution de courants neutres est également importante (21%). Nous observons aussi une contribution faible de production de paires de leptons (< 1%) et de photoproduction. Ces contributions sont dues à la radiation d'un photon d'un quark ou d'un électron. Nous obtenons 165 événements pour 142 $\pm$ 6 prédits.

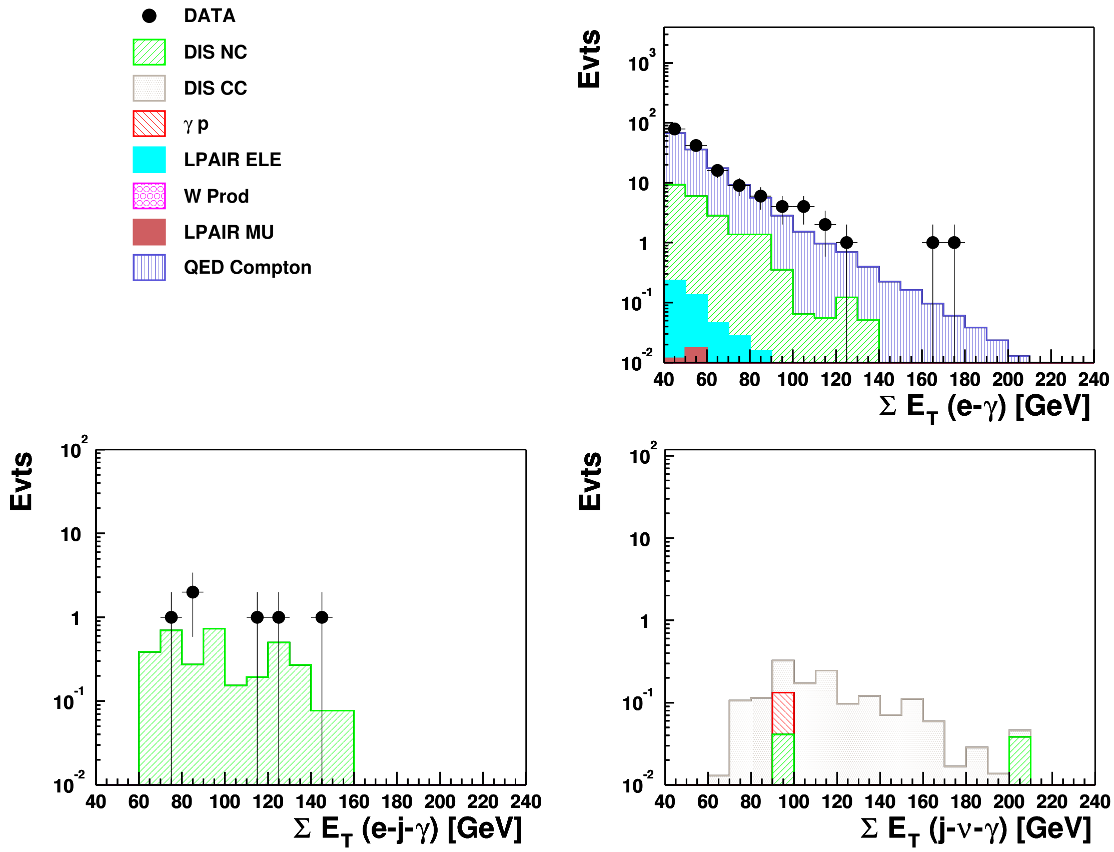


FIG. 5.34 – Distribution de $\Sigma E_T^{e-\gamma}$, $\Sigma E_T^{e-j-\gamma}$ et $\Sigma E_T^{j-\nu-\gamma}$.

L'allure de la distribution en $\Sigma E_T^{e-\gamma}$ est en bon accord avec les prédictions (figure 5.34). Deux événements à $\Sigma E_T^{e-\gamma} > 160$ GeV ressortent du lot. L'un d'entre eux est montré sur la figure 5.35.
Dans le canal e-γ-γ, nous observons un événement pour 0 événement prédit. Son énergie transverse totale est de 99 GeV. Comme cette topologie ne peut pratiquement être produite que par diffusion QED Compton avec radiation d'un photon par l'électron, la contribution potentielle d'autres processus est réduite. Or, nous avons vu dans le paragraphe 2.3.1 qu'au moment de l'analyse, le générateur $WABGEN$

utilisé n'incluait pas encore les corrections radiatives électromagnétiques à grande impulsion transverse. Cependant, nous pouvons estimer l'ordre de grandeur de la contribution du Modèle Standard en faisant l'hypothèse qu'à énergie transverse égale, le taux d'événements prédits de diffusion QED Compton avec radiation de photon est à peu près égal à celui du processus de diffusion QED Compton simple, diminué d'un facteur α_{em} (α_{em} est la constante de structure fine habituelle). Dans notre cas, cela se traduit par la relation suivante :

$$N^{e-\gamma-\gamma}_{E_T>60} \approx \alpha_{em}\, N^{e-\gamma}_{E_T>60} \tag{5.7}$$

où $N^{e-\gamma-\gamma}_{E_T>60}$ est le taux d'événements prédits dans le canal e-γ-γ et $N^{e-\gamma}_{E_T>60}$ le taux d'événements observés dans le canal e-γ à $\Sigma E_T > 60$ GeV. Nous en déduisons donc un taux attendu d'environ 0,3 événement dans le canal e-γ-γ, ce qui est compatible avec l'événement observé. Une vue longitudinale de l'événement observé est montrée sur la figure 5.35.

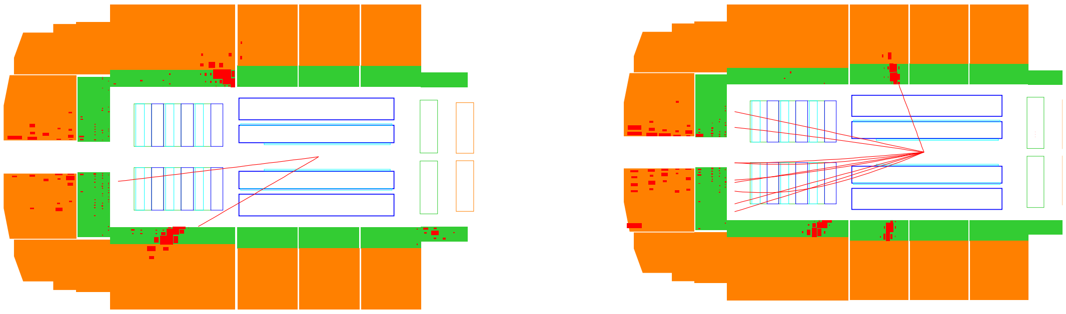


FIG. 5.35 – Vue longitudinale de l'événement e-γ de plus grand $\Sigma E_T^{e-\gamma}$(à gauche) et de l'événement e-γ-γ (à droite).

Les canaux e-j-γ et j-ν-γ correspondent respectivement à la topologie de courant neutre (99%) et de courant chargé (87%) avec une radiation de photon. Dans le canal e-j-γ, nous remarquons une faible contribution (1%) de processus de production inélastique d'électrons où l'un des électrons ne satisfait pas les critères de sélection. Dans le canal j-ν-γ, le taux de mauvaise identification est élevé : nous observons des contributions de processus de photoproduction (6%) où la reconstruction erronée de jets simulent de l'impulsion manquante et également des courant neutres (5%). Les distributions en $\Sigma E_T^{e-j-\gamma}$ et $\Sigma E_T^{j-\nu-\gamma}$ (figure 5.34) sont en accord global compte tenu du faible taux d'événements prédits. Dans le canal e-j-γ, 6 événements sont observés pour 3,4 $\pm$ 1,0 prédits. Dans le canal j-ν-γ, aucun événement n'est observé pour

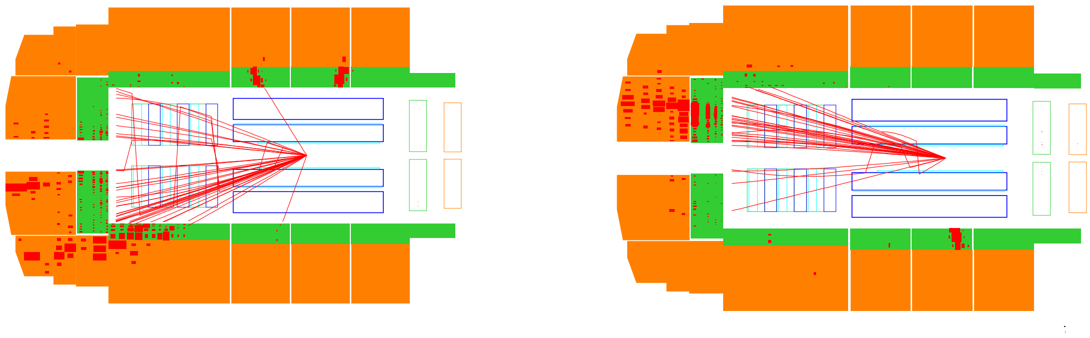


FIG. 5.36 – Vue longitudinale d'un événement du canal e-j-γ (à gauche) et j-γ (à droite).

1,5 ± 0,2 prédits. Une vue longitudinale d'un événement typique du canal e-j-γ est montrée dans la figure 5.36.

Nous avons également analysé les canaux à jets contenant un photon dans l'état final : il s'agit des canaux j-γ et j-j-γ.
Dans le canal j-γ, nous observons 46 événements pour 60 ± 16 prédits, dont seulement 60% de processus de photoproduction contenant un jet et un photon dans l'état final. 36% des contributions viennent de courants neutres mal identifiés. Cette contribution est due à l'inefficacité de reconstruction de la trace de l'électron, qui bien que de l'ordre de 1% uniquement, donne une contamination significative au canal j-γ car la section efficace des courants neutres est environ de deux ordres de grandeur plus importante que celle de la production réelle de j-γ. Le Monte-Carlo a été ajusté pour décrire correctement cette inefficacité. La distribution en $\Sigma E_T^{j-\gamma}$ présentée sur la figure 5.37 est bien décrite. Un événement typique est montré sur la figure 5.36.

Dans le canal j-j-γ, dont la distribution en $\Sigma E_T^{j-j-\gamma}$ est présentée sur la figure 5.37, nous observons 8 événements pour 6,5 ± 2,5 prédits. Les contributions principales sont les courants neutres radiatifs avec 86% et le processus de photoproduction radiatif avec 13%. Les événements de données sont reconstruits dans une région de l'espace de phase peuplée par les prédictions Monte-Carlo. La vue longitudinale d'un événement typique est présentée dans la figure 5.38.

En résumé, nous avons montré que la distribution en $\Sigma E_T^{e-\gamma}$ est bien décrite dans le canal e-γ . Deux événements ayant des masses supérieures à 160 GeV sont trouvés. Dans le canal e-γ-γ, 1 événement est observé, mais nous sommes limités par les spécificités du générateur qui ne nous permettent pas d'estimer la prédiction du Modèle Standard précisément. Les canaux radiatifs e-j-γ et j-ν-γ sont en bon accord

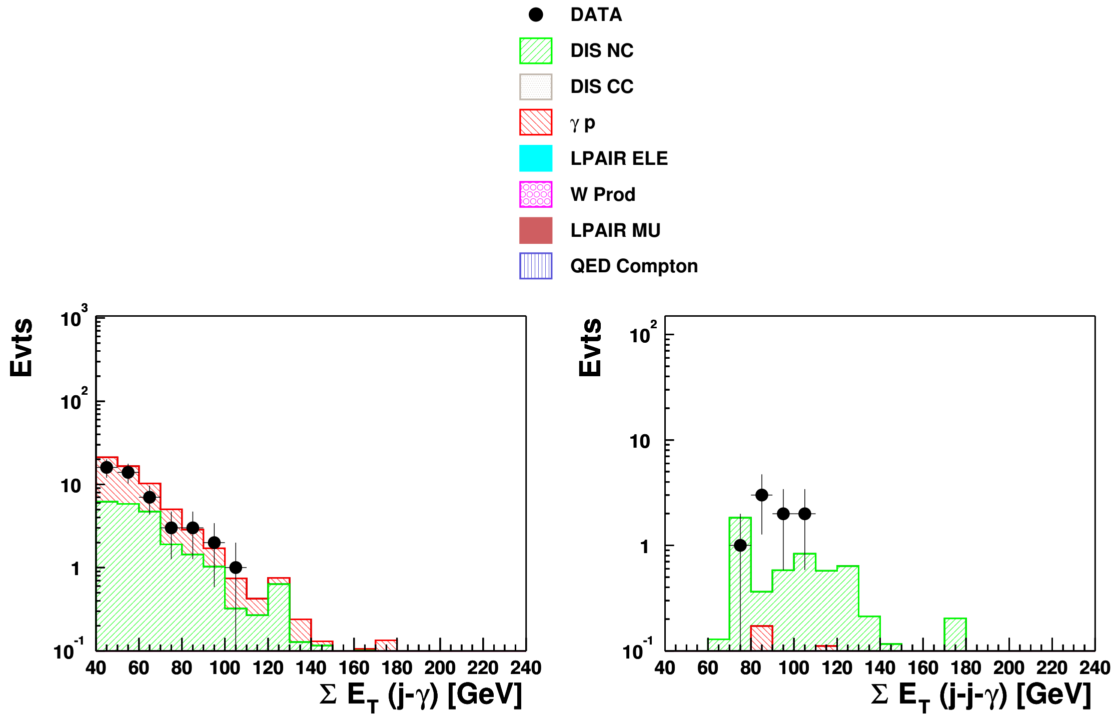


FIG. 5.37 – Distributions de $\Sigma E_T^{j-\gamma}$ et $\Sigma E_T^{j-j-\gamma}$.

avec les prédictions. Dans les canaux j-γ et j-j-γ, l'accord est aussi correct, mais nous observons une contamination importante par le processus de courant neutre.

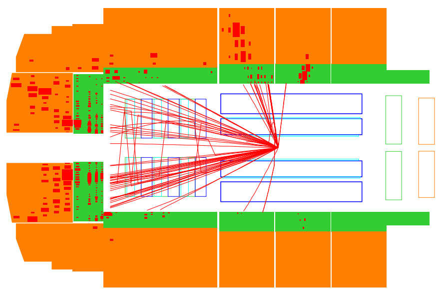


FIG. 5.38 – Vue longitudinale d'un événement typique du canal j-j-γ.

5.5.7 Canaux de jets en photoproduction

Nous classons les canaux j-j, j-j-j, j-j-j-j et j-j-j-j-j parmi cette catégorie. Il s'agit de canaux dominés par le processus de photoproduction. Dans le canal j-j, nous trouvons 34561 événements observés pour 36812 ± 8098 prédits dont une majorité de processus photoproduction (85%). Une contribution plus faible de 14% vient de courants neutres, dans lesquels l'électron ne remplit pas les critères de sélection des corps. Il subsiste des événements courant chargé et de production de boson $W^{\pm}$ où l'une des particules ne passe pas les critères de sélection. Une contribution négligeable provient de la production inélastique de paires d'électrons où l'on a reconstruit l'un des électrons comme jet.

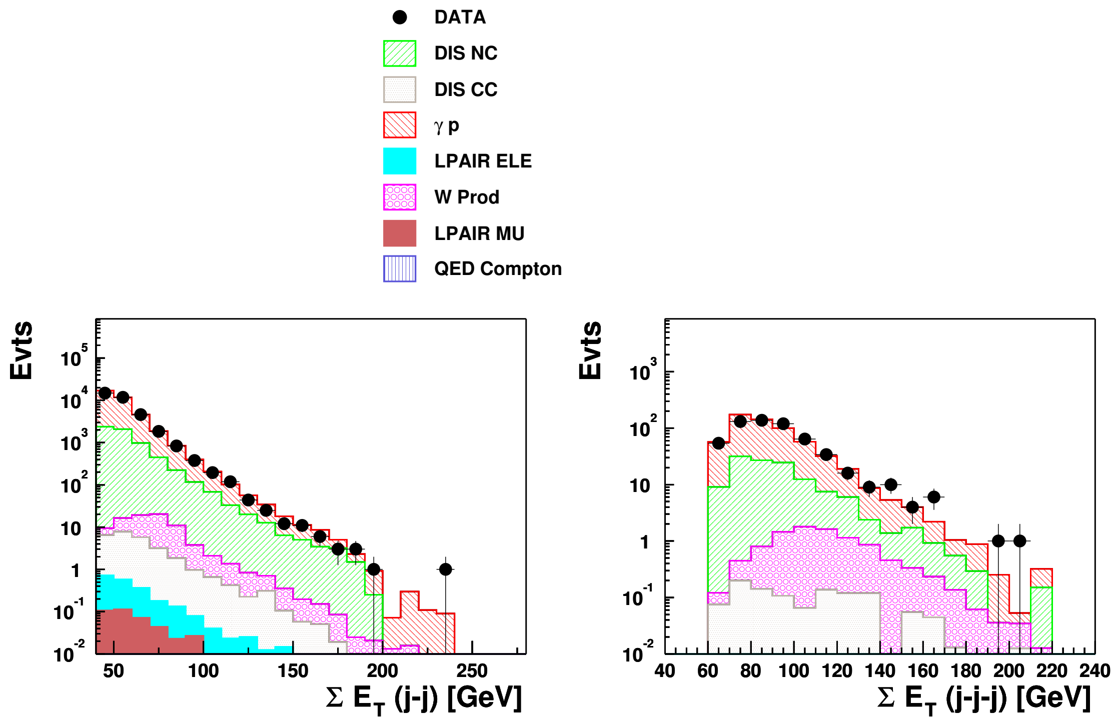


FIG. 5.39 – Distributions de ΣE_T^{j-j} et ΣE_T^{j-j-j}.

De façon générale, la distribution de ΣE_T^{j-j} (figure de gauche de 5.39) est bien décrite jusqu'à des $\Sigma E_T^{j-j} = 200$ GeV sur six ordres de grandeur. L'événement se détachant dans la distribution en ΣE_T^{j-j} est également reconstruit à grande masse. Une vue longitudinale de celui-ci est présentée dans la figure 5.40.

Dans le canal j-j-j, la distribution de ΣE_T^{j-j-j} (distribution à droite de la figure 5.39) est également bien décrite. Nous obtenons 588 événements de données pour 604 ± 199 prédits. Comme dans le canal j-j, les contributions principales sont les processus de photoproduction (82%) et de courant neutre (16%). Des contributions plus importantes (2%) de production de boson $W^{\pm}$ apparaîssent à grand ΣE_T^{j-j-j}. Au-dessus de 190 GeV, nous trouvons deux événements pour 0,7 prédits. L'un est montré en vue longitudinale sur la figure 5.40.

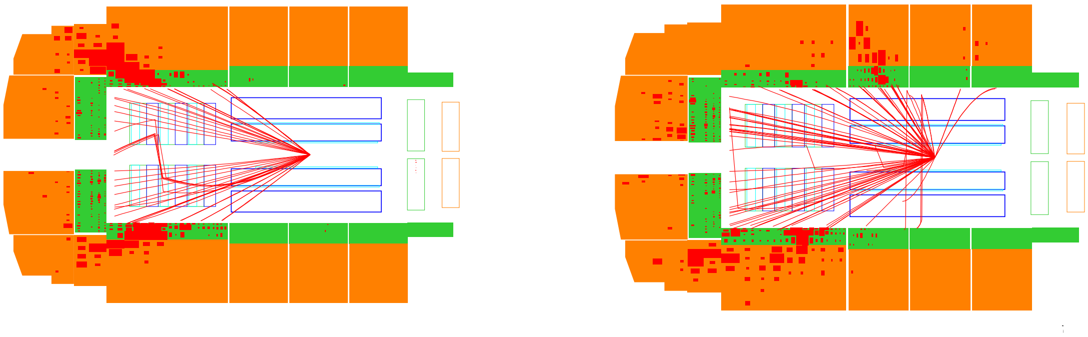


FIG. 5.40 – Vue longitudinale des événements du canal j-j (à gauche) et j-j-j (à droite) de plus grande énergie transverse.

Les prédictions dans les canaux à quatre corps sont beaucoup plus faibles. Nous sommes confrontés dans le canal j-j-j-j à des incertitudes très grandes de plus de 60%, dues aux erreurs statistiques et aux erreurs sur l'échelle d'énergie hadronique. La distribution de $\Sigma E_T^{j-j-j-j}$ (distribution de gauche de la figure 5.41) montre un désaccord important entre les simulations Monte-Carlo (7,3 $\pm$ 4,6 événements prédits) et les données (15 événements observés). La raison la plus probable implique les corrections d'ordre supérieur qui ne sont pas incluses dans les générateurs actuels pour le processus de photoproduction. Dans le canal à cinq corps j-j-j-j-j, nous observons un événement pour une prédiction nulle. Ici aussi, l'absence de radiation QCD aux ordres suffisants rend la prédiction peu fiable. Un événement typique du canal j-j-j-j et l'événement du canal j-j-j-j-j sont montrés dans la figure 5.42.

En résumé, nous avons montré que les canaux j-j et j-j-j dominés par la photoproduction sont bien décrits par les simulations Monte-Carlo compte tenu des erreurs élevées jusqu'à 33%. Par contre le canal j-j-j-j présente un écart notable de 15 événements prédits pour 7,3 $\pm$ 4,6 prédits : cet écart montre les limitations actuelles des générateurs et de notre compréhension des processus de QCD aux ordres supérieurs. Dans le canal j-j-j-j-j, un événement est observé pour une prédiction nulle.

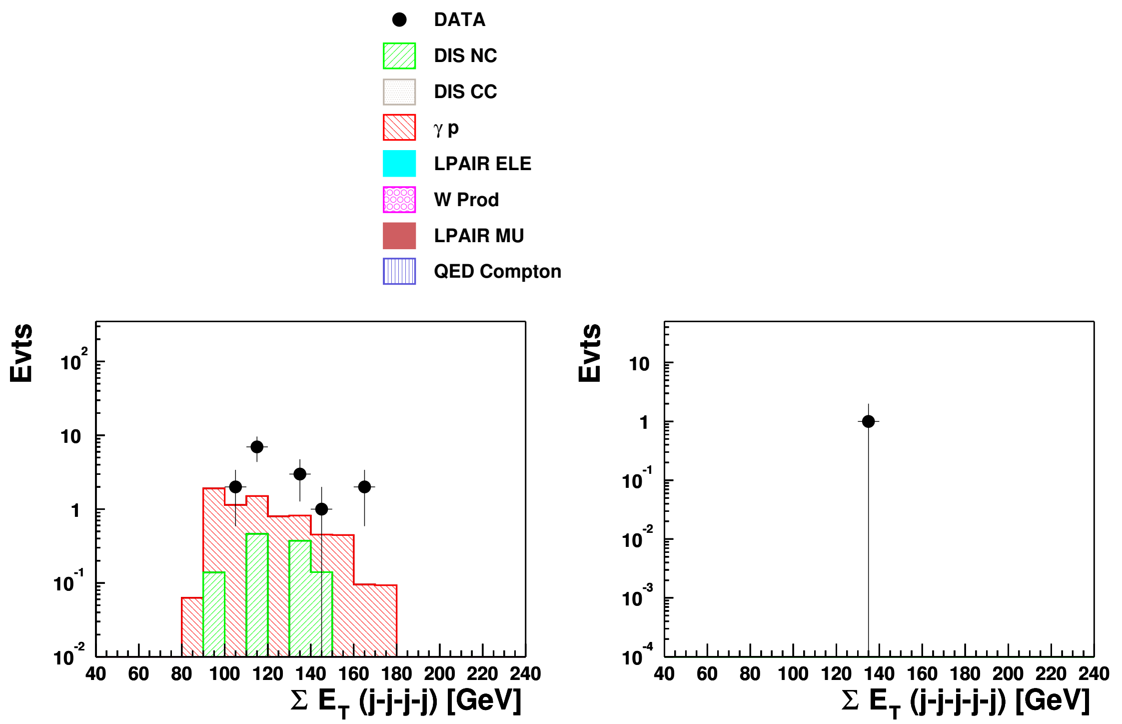


FIG. 5.41 – Distributions de $\Sigma E_T^{j-j-j-j}$ et $\Sigma E_T^{j-j-j-j-j}$.

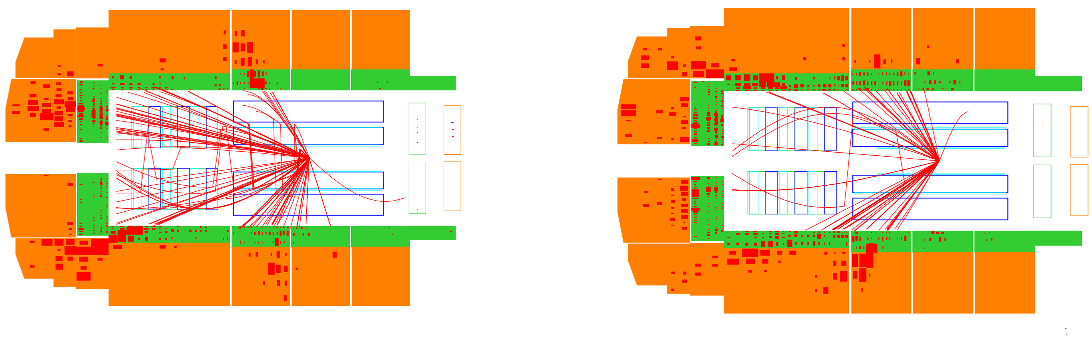


FIG. 5.42 – Vue longitudinale d'un événement du canal j-j-j-j (à gauche) et j-j-j-j-j (à droite).

5.5.8 Echantillon réduit

Dans les paragraphes précédents, nous avons discuté séparément les résultats obtenus dans chacun des 24 canaux. Les données sont bien décrites par les prédictions, mais, notamment dans le canal leptonique e-e et le canal à trois corps μ-j-ν, des déviations à grande masse subsistent. Il est plus facile de trouver des topologies atypiques dans ce type de canal à faible taux d'événements que dans les canaux à jets comme e-j ou j-j.

Dans cette partie, nous allons présenter une approche visant à discerner les topologies atypiques, pouvant provenir de désintégrations de particules lourdes, qui sera spécialement utile dans les canaux à grand nombre d'événements, comme par exemple e-j ou j-j. Contrairement aux analyses visant à tester un modèle précis, où la topologie attendue du signal est utilisée pour améliorer l'efficacité de détection, nous ne connaissons pas le signal. L'approche suivie dans notre cas consiste à se placer, en faisant le moins d'hypothèses possible sur les caractéristiques du signal attendu, dans une région de l'espace de phase dans laquelle les contributions du Modèle Standard sont réduites. L'effet de coupures additionnelles utilisant des variables topologiques a été étudié en vue de caractériser plus en détails la topologie des événements mesurés. Cette approche a déjà été utilisée dans la recherche de leptoquarks [68], en étudiant par exemple les corrélations entre l'angle polaire de l'électron dans le centre de masse du système électron-quark et sa masse invariante. Nous avons utilisé les variables suivantes pour introduire des coupures visant à sélectionner des topologies atypiques par rapport au Modèle Standard :

- $\mathbf{cos}(\theta^*_{part\ i,i-j})$:
  Cette variable est définie à partir de l'angle polaire de la particule i dans le centre du masse du système i-j.
- $\mathbf{X}_3$:
  Cette variable est définie comme :

$$X_3 = \frac{2E'_{part1}}{E'_{part1} + E'_{part2} + E'_{part3}} \tag{5.8}$$

  avec E'_{parti} l'énergie de la particule i dans le centre de masse des trois particules qui sont ordonnées en énergie décroissante. Il s'agit d'une variable de Dalitz décrite dans [69], définie uniquement pour les systèmes à trois corps ou plus. Elle vérifie $2/3 < X_3 < 1$.

Nous allons étudier l'impact de ces coupures sur la distribution en masse invariante des corps, dans les canaux à grande statistique e-j, j-j, j-ν, e-j-j et j-j-j. Les distributions des varaibles considérées sont montrées sur la figure 5.43 pour ces 5 canaux.

Dans le canal e-j, la topologie typique des courants neutres donne un électron diffusé dans la partie arrière du calorimètre et un jet projeté vers l'avant : la majeure partie des événements est concentrée à une valeur de $cos(\theta^*_{ele}) \approx$ -1. La distribution typique de $cos(\theta^*_{\nu})$ dans le canal j-ν suit une distribution similaire au canal e-j : le jet est projeté à l'avant et par conséquent le neutrino est reconstruit dans la partie arrière du détecteur. La topologie des événements du canal j-j est très différente :

les jets sont tous les deux projetés à l'avant, d'où la distribution symétrique de $cos(\theta^*_{jet1})$. Il est intéressant de noter que le domaine le plus peuplé est distribué uniformément autour de 0. Dans les canaux e-j-j et j-j-j, nous avons considéré la variable X_3. Comme on peut facilement s'en convaincre, X_3 vaut $2/3$ pour une topologie où les jets sont associés en forme d'étoile dans le plan transverse avec un angle azimutal respectif de 120^o. La figure 5.43 montre sa distribution pour les deux canaux e-j-j et j-j-j.

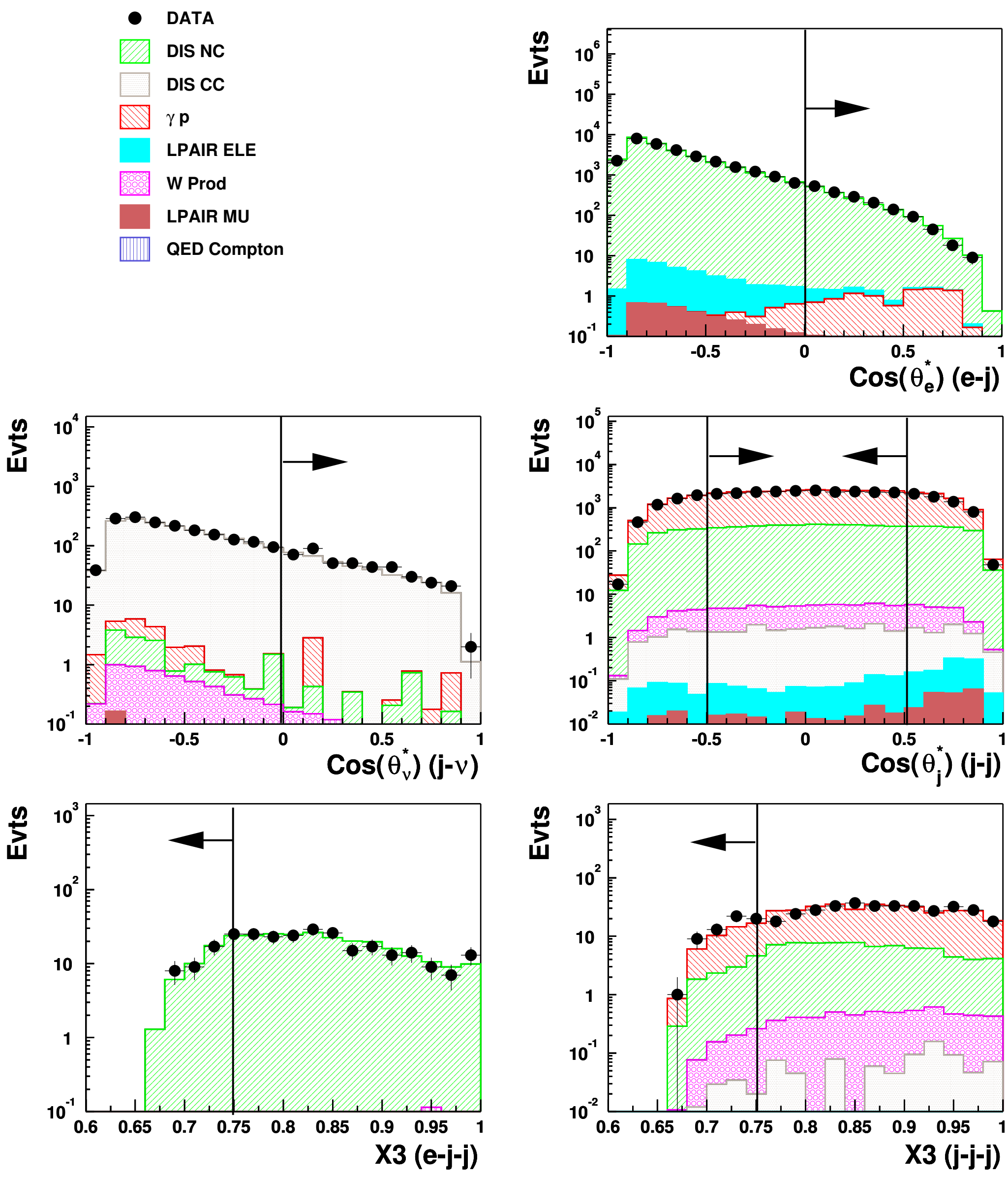


FIG. 5.43 – Distributions de $cos(\theta^*_{ele})$, $cos(\theta^*_{\nu})$, $cos(\theta^*_{j})$, X_3^{e-j-j} et X_3^{j-j-j}.

A partir de ces distributions, nous sélectionnons les régions dans l'espace de phase, indiquées par les flèches sur la figure 5.43, pour lesquelles la contribution du

Modèle Standard est réduite :

- Dans les canaux *e-j* et *j-ν* , nous demandons $cos(\theta^*) > 0$ pour sélectionner les topologies ou l'électron (le neutrino) est projeté à l'avant.
- Dans le canal *j-j*, il faut que $|cos(\theta^*_{jet1})| < 0,5$ afin d'éliminer les topologies où les deux jets sont émis à l'avant et à l'arrière respectivement, ce qui peut donner des masses arbitrairement grandes.
- Dans les canaux *e-j-j* et *j-j-j*, la variable X_3 doit être inférieure à 0,75 afin de favoriser les topologies rares où les corps sont associés en forme d'étoile.

Ces coupures sont résumées dans le tableau 5.18.

Canal	Coupures anti-MS
e-j	$cos(\theta^*_{ele}) > 0$
j-ν	$cos(\theta^*_{\nu}) > 0$
j-j	$\|cos(\theta^*_{jet1})\| < 0,5$
e-j-j	$X_3 < 0,75$
j-j-j	$X_3 < 0,75$

TAB. 5.18 – Coupures supplémentaires appliquées pour réduire la contribution des processus du Modèle Standard.

Les distributions de masse invariante avant et après les coupures sont montrées sur les figures 5.44 et 5.45.

Dans le canal *e-j*, la coupure appliquée réduit le taux d'événements d'un facteur 20. Les distributions avant et après application sont bien décrites par les simulations Monte-Carlo, avec des fluctuations statistiques importantes dans l'échantillon réduit. Parmi cinq événements reconstruits avec une masse invariante au-dessus de 250 GeV, seul un événement présente une topologie atypique, c'est-à-dire avec $cos(\theta^*) > 0$.

Dans le canal *j-ν*, le taux d'événements est réduit d'un facteur 5. Les distributions de la masse invariante sont également en bon accord avec les prédictions du Modèle Standard. Deux événements de masse invariante supérieure à 240 GeV sont observés pour 0,9 ± 0,05 prédits.

Dans le canal *j-j*, l'échantillon final est constitué de la moitié des événements initiaux. Les données sont en bon accord avec les simulations Monte-Carlo : aucune déviation significative n'est visible.

Dans les canaux *e-j-j* et *j-j-j*, la sélection restrictive réduit considérablement l'échantillon. Des fluctuations importantes dans les échantillons finaux sont visibles. Dans le canal *e-j-j*, aucun événement n'est reconstruit avec une masse au-delà de 175 GeV. L'échantillon est encore plus réduit dans le canal *j-j-j*, dans lequel aucun événement de masse supérieure à 130 GeV n'est reconstruit.

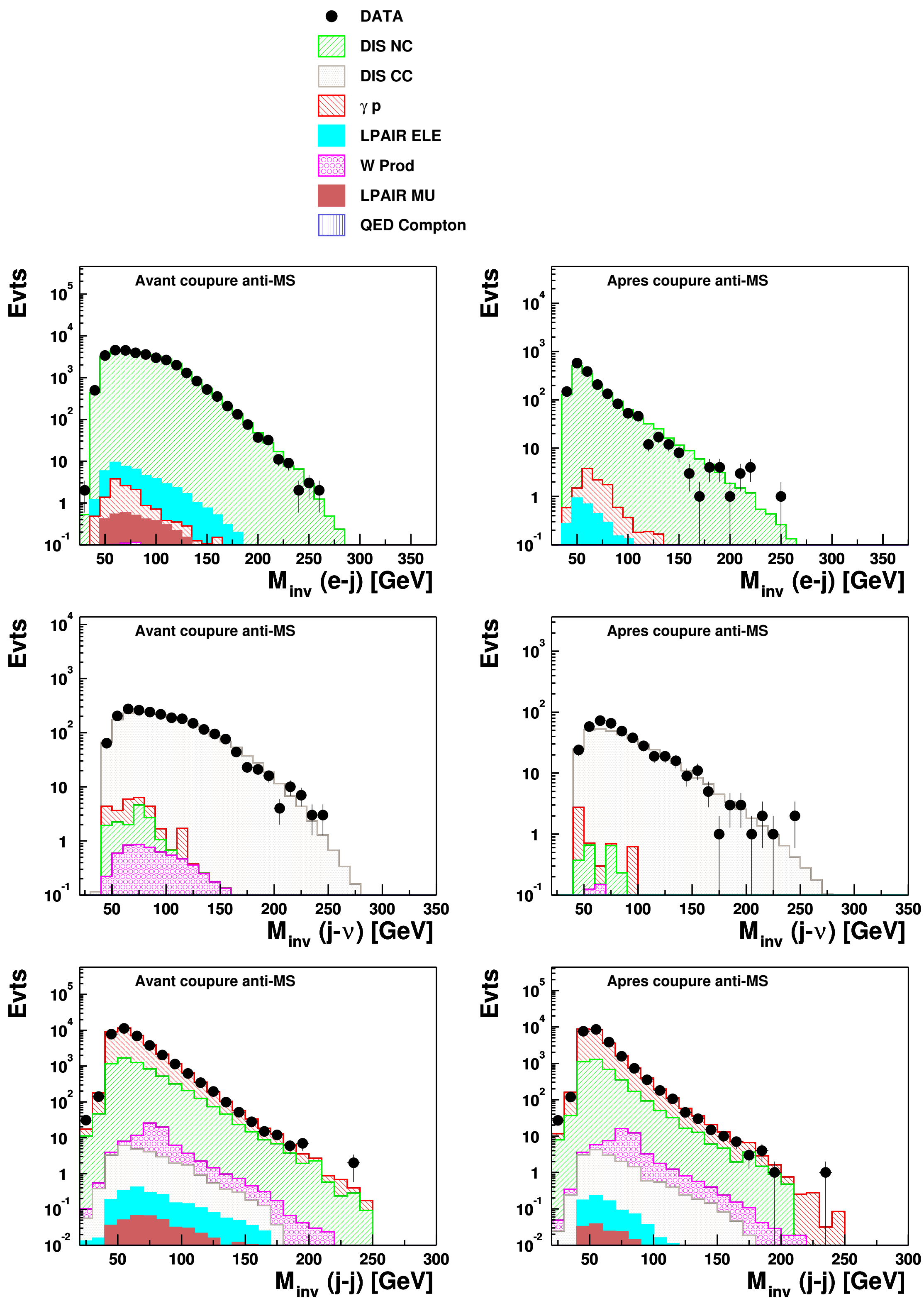


FIG. 5.44 – Distributions de la masse invariante, dans les canaux e-j,j-ν et j-j, avant (à gauche) et après (à droite) la coupure réduisant la contribution du Modèle Standard.

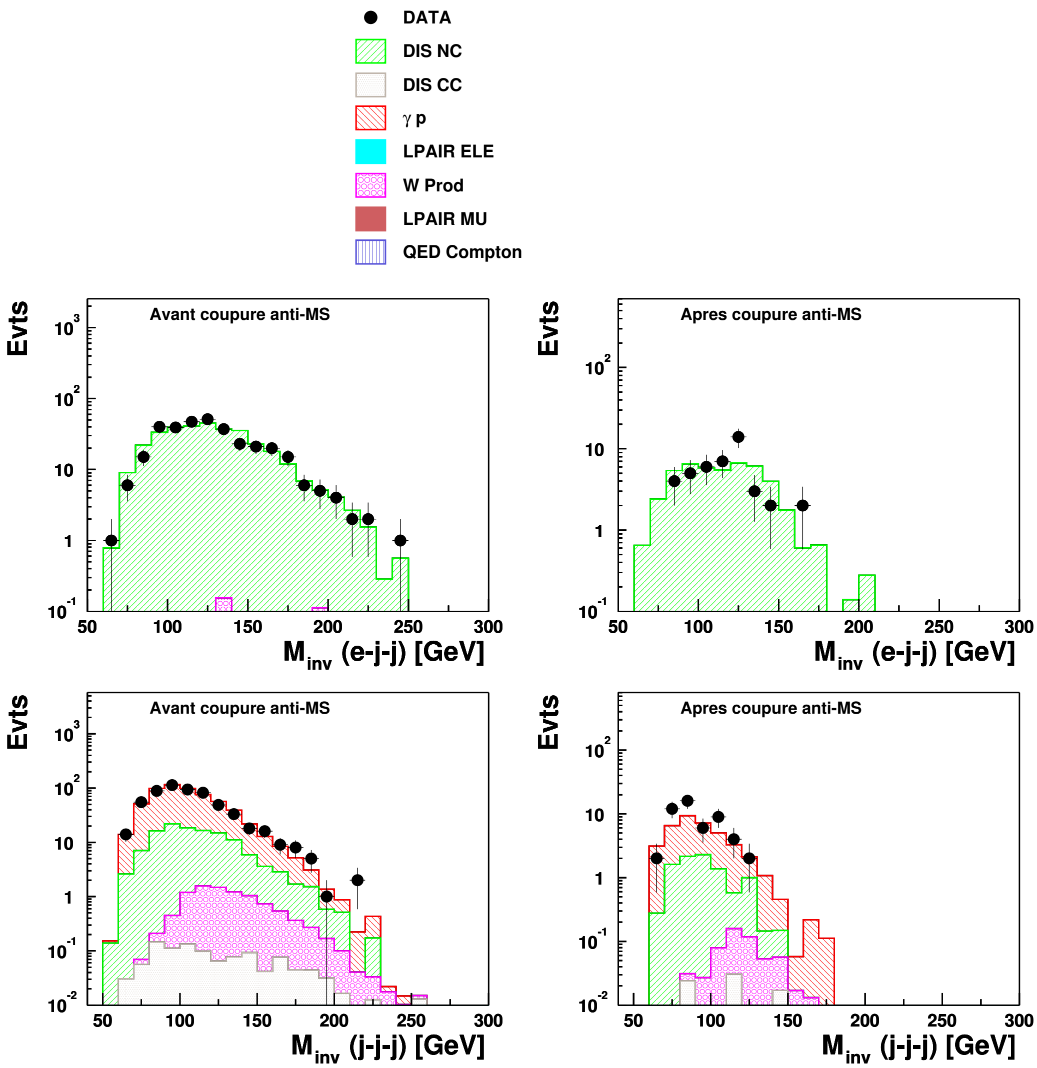


FIG. 5.45 – Distributions de la masse invariante, dans les canaux e-j-j et j-j-j, avant (à gauche) et après (à droite) la coupure réduisant la contribution du Modèle Standard.

En résumé, nous avons vu que les variables topologiques $cos(\theta^{*})$ et X_3 permettent de sélectionner des régions de l'espace de phase contenant des événements de topologie atypique dans le cadre du Modèle Standard. Les échantillons ainsi sélectionnés sont bien décrits par les prédictions du Modèle Standard et aucune déviation significative n'a été observée dans les distributions en masse invariante. Nous avons également étudié les distributions en ΣE_T de ces échantillons : nous observons aussi un bon accord entre les données et les simulations Monte-Carlo. Cette étude peut être étendue en utilisant d'une part d'autres variables topologiques, comme par exemple la sphéricité de l'événement, et en l'appliquant d'autre part à l'ensemble des canaux.

5.6 Applications aux données HERA II

Nous avons aussi analysé les toutes premières données de HERA II prises en 2002 et début 2003. Ce lot est traité à part en raison des modifications importantes entreprises au niveau de l'accélérateur et du détecteur. Il correspond à une luminosité intégrée de 2,4 pb^{-1}. Les lots de simulations Monte-Carlo utilisés correspondent aux conditions des détecteurs en 2000.

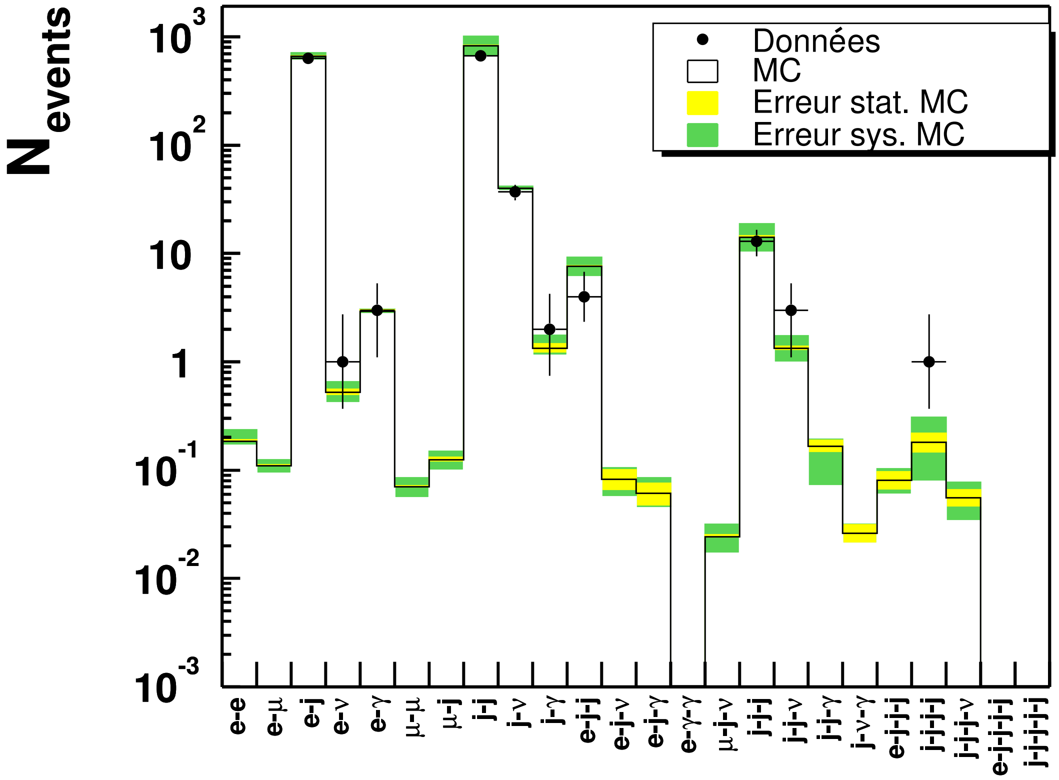


FIG. 5.46 – Taux d'événements par canal pour les données prises entre 2002 et 2003.

Le résultat pour les 24 canaux est montré sur la figure 5.46. Les taux de données sont bien décrits par les simulations Monte-Carlo. Un événement à quatre jets a été reconstruit. L'application aux données de la phase II de HERA permet un premier contrôle de la qualité des données et permet aussi de verifier le bon fonctionnement du système de déclenchement à grand P_T qui nous intéresse plus particulièrement.

Nous avons vu que les données sont, de manière générale, bien décrites par les prédictions et que la chaîne complète de l'analyse fonctionne bien avec les données de la phase II. Cependant, les lots de simulations Monte-Carlo utilisés correspondent aux conditions du détecteur en 2000 et les erreurs systèmatiques comme par exemple la calibration en énergie et les efficacités du système de déclenchement correspondant aux conditions en 2002 et 2003 ne sont donc pas prises en compte.

Dans l'avenir, l'analyse sera utilisée, entre autre, pour suivre en ligne les données et pour déceler des événements rares produits à grande énergie transverse.

5.7 Conclusion

Nous avons présenté la recherche de processus rares et de signaux de nouvelle physique dans une analyse modèle-indépendante sur l'ensemble des données prises entre 1994 et 2000 : 24 topologies différentes ont été étudiées à grande énergie transverse ΣE_T. Nous avons passé en revue les critères de sélection des cinq corps électron, muon, jet, neutrino et photon recherchés. Après application de la sélection brute, nous constatons que les canaux e-e, e-ν, j-ν, j-γ, μ-μ, μ-j et e-j-ν sont, soit contaminés par du bruit de fond non physique, comme par exemple des muons cosmiques, soit contaminés par du bruit de fond physique venant d'une mauvaise identification de l'un des corps. Des coupures supplémentaires tenant compte de la topologie ou de la cinématique permettent de réduire ces contributions.
Les taux globaux obtenus en appliquant les coupures finales montrent un bon accord entre les données et les simulations Monte-Carlo. Les incertitudes totales calculées à partir des erreurs statistiques des simulations Monte-Carlo et des erreurs systématiques sont faibles, de l'ordre de 4%, pour les canaux à deux corps mais augmentent jusqu'à dépasser 50% dans les canaux à trois corps ou plus. La source principale d'erreurs dans les canaux à plusieurs jets, où ces erreurs sont importantes, est l'incertitude sur l'échelle d'énergie. Mis à part les déviations connues des canaux e-e et μ-j-ν à grand ΣE_T, aucun écart significatif n'a été trouvé. L'étude des canaux e-γ-γ et des canaux à plusieurs jets, spécialement intéressants car aucune analyse dédiée n'en a été effectuée, montre les limites des générateurs actuels.

Afin de sélectionner des régions de l'espace de phase contenant des événements atypiques dans le cadre du Modèle Standard, des échantillons réduits ont été obtenus en appliquant des coupures additionnelles sur des variables topologiques. Nous observons également sur ces échantillons réduits un bon accord entre les données et les simulations Monte-Carlo. Aucune déviation significative n'a été constatée, ce qui montre la bonne description de nos données par les processus du Modèle Standard.

Enfin, l'application aux données de la phase II de HERA a permis de montrer que la chaîne d'analyse complète fonctionne bien et que les données sont bien décrites par les prédictions.

Chapitre 6

Interprétation statistique des résultats

What we call the beginning is often the end
And to make an end is to make a beginning.
The end is where we start from.

T S Eliot
Little Gidding

Dans le chapitre précédent, nous avons présenté une analyse qui étudie l'ensemble des topologies possibles impliquant des électrons, muons, jets, neutrinos et photons de grand P_T réparties en différents canaux. 24 canaux ont ainsi été analysés en détails et nous avons vu que les données observées y sont en général bien décrites par le Modèle Standard. Cependant, dans les canaux e-e et μ-j-ν, nous avons observé une déviation importante à grand $\Sigma\ E_T$ entre les données et les simulations Monte-Carlo. Dans ce chapitre, nous allons présenter une méthode destinée à repérer de manière automatique les régions d'écart maximum entre les données et les prédictions dans les différents canaux. Ensuite, nous allons quantifier la signification statistique de ces déviations.

Dans la suite, nous décrirons d'abord le principe de cette méthode, puis nous l'appliquerons aux résultats obtenus au chapitre précédent. Enfin, sa sensibilité à des modèles de physique nouvelle sera testée à partir d'un exemple de processus de production anormale du quark top.

6.1 Méthode d'interprétation statistique

Le principe est inspiré d'une méthode développée pour interpréter les résultats d'un grand nombre d'états finaux exclusifs, obtenus sur l'expérience DØ [25, 26, 70]. Le but est de quantifier les déviations observées afin de déterminer s'il s'agit d'une fluctuation statistique ou d'un signal réel. L'algorithme comprend les étapes suivantes :

- Nous définissons un espace de variable(s) discriminantes.
- Puis, nous divisons notre espace en différentes régions.
- Nous déterminons la région de plus grand intérêt, c'est-à-dire celle de déviation maximale.
- Finalement, nous calculons la signification statistique de l'écart observé dans cette région par rapport aux prédictions du Modèle Standard.

6.1.1 Choix d'un espace de variables

Dans un permier temps, nous devons déterminer les variables caractéristiques des événements observés, qui sont susceptibles d'être les plus discriminantes. Ainsi, la masse invariante M_{inv} et ΣE_T sont intéressantes car des signes de nouvelle physique sont plus susceptibles d'apparaître dans les régions à grande masse et à grand ΣE_T, où les prédictions du Modèle Standard sont faibles. Bien sûr, il est aussi possible d'étudier la combinaison de différentes variables afin de garder toutes les informations de corrélations possibles entre elles. L'algorithme utiliserait alors un espace à plusieurs dimensions. Cependant, pour des raisons de simplicité, mais aussi techniques liées à la précision statistique du Monte-Carlo, nous nous sommes pour l'instant limités ici à un espace unidimensionel. Les variables M_{inv} et ΣE_T sont ainsi retenues.

6.1.2 Localisation d'une déviation éventuelle

Afin de localiser une déviation éventuelle, nous divisons chaque distribution en différentes régions. Dans le cas d'une distribution unidimensionnelle, une région est définie à partir de sa borne inférieure et supérieure et est constituée de l'ensemble des points de données et de simulations Monte-Carlo entre ces bornes. De manière la plus générale, la totalité des régions possibles est obtenue en variant les bornes inférieures et supérieures selon toutes les combinaisons possibles. Nous appelons cette catégorie générale "régions à fenêtre glissante" : ainsi, à la fois la taille de la fenêtre et sa position peuvent varier. Ces régions sont spécialement intéressantes lorsque l'on recherche des résonances localisées.

La catégorie des régions à fenêtre glissante contient un sous-lot de régions, dont la borne supérieure est définie par la limite supérieure du spectre. Ce type de régions est appellé "région intégrée". Comme la nouvelle physique peut aussi se manifester par un excès d'événements, qui n'est pas forcément localisé, à grande masse ou à grand ΣE_T, ces régions sont également intéressantes, une préférence étant alors donnée aux queues de distributions.

Dans les deux cas, les bornes peuvent être déterminées en utilisant la distribution des données (appellée alors "région données") ou des simulations Monte-Carlo (appellée alors "région MC"). Nous allons limiter la taille minimale d'une région sélectionnée à deux fois la résolution de la mesure en masse ou en ΣE_T. Nous l'avons fixée de manière générale à 20 GeV.

Dans les régions peu peuplés des spectres des données, les bornes sont définies à partir de la médiane entre deux points de données. Lorsque les queues de spectre ne contiennent aucun point de données, la région potentielle est intégrée jusqu'à l'infini. Un exemple de région intégrée est présenté sur la figure 6.1 pour la distribution de masse invariante dans le canal j-γ. La borne inférieure est définie à partir de la médiane entre les points de données situés à 130 et 180 GeV. Par contre, pour les "régions MC", toutes les régions possibles, de taille supérieure à la résolution, sont autorisées. La méthode à "régions MC" a ainsi tendance à minimiser la contribution Monte-Carlo pour un même nombre d'événements de données : la seule contrainte vient de la limitation en résolution.

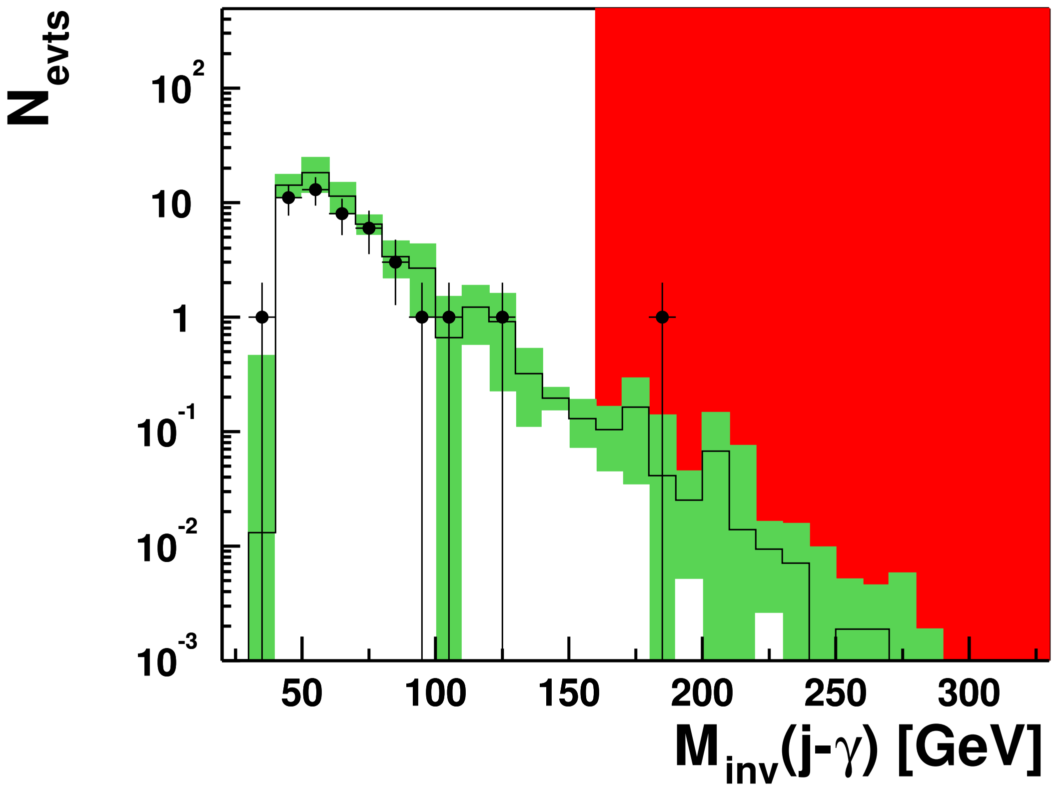


FIG. 6.1 – Distribution de la masse invariante dans le canal j-γ avec un exemple de région intégrée, dont la borne inférieure est définie à partir de la médiane entre les points de données situés à 130 et 180 GeV.

Pour chaque région R ainsi définie, nous déterminons la probabilité p_R que le nombre d'événements attendus (simulations Monte-Carlo) fluctue au-dessus du nombre d'événements observés (données). Cette probabilité est obtenue en utilisant la loi de Poisson :

$$p_R = \sum_{i=N_{obs}}^{\infty} \frac{e^{-b_R} b_R^i}{i!} \tag{6.1}$$

où b_R est le nombre d'événements attendus (Monte-Carlo) et N_{obs} le nombre d'événements observés dans une région R donnée.

Afin de tenir compte des erreurs statistiques et systématiques de la simulation Monte-Carlo, nous faisons l'hypothèse que celles-ci sont gaussiennes et symétriques. En convoluant alors la distribution poissonienne des événements attendus avec la distribution gaussienne des erreurs, nous obtenons la probabilité finale :

$$p_R^{data} = p_{N \geq N_{obs}^{data}} = \frac{\int_0^{+\infty} db \quad G(b; b_R; \delta b_R) \sum_{i=N_{obs}}^{\infty} \quad \frac{e^{-b} b^i}{i!}}{\int_0^{+\infty} db \quad G(b; b_R; \delta b_R)} \tag{6.2}$$

où $G(b; b_R; \delta b_R)$ est une distribution gaussienne de valeur moyenne b_R et de largeur δb_R. Nous calculons l'erreur totale δb_R à partir de la relation $\delta b_R = \sqrt{(\delta b_R^{sys}) + (\delta b_R^{stat})}$, où δb_R^{sys} est l'erreur systématique et δb_R^{stat} l'erreur statistique de la simulation Monte-Carlo. Une faible valeur de cette probabilité ($p_R^{data} \approx 0$) correspond alors à un excès de données dans la région R, alors qu'une grande valeur ($p_R^{data} \approx 1$) correspond à un déficit. La probabilité inverse $p_{N \leq N_{obs}^{data}}$ de la région de plus grand déficit est donnée par la relation suivante :

$$p_{N \leq N_{obs}^{data}} = \frac{\int_0^{+\infty} db \quad G(b; b_R; \delta b_R) \sum_{i=0}^{N_{obs}} \quad \frac{e^{-b} b^i}{i!}}{\int_0^{+\infty} db \quad G(b; b_R; \delta b_R)} \tag{6.3}$$

Cette formule peut être utilisée à la place de l'équation 6.2 pour la recherche de régions présentant un déficit de données. Nous nous intéresserons principalement dans la suite aux régions avec un excès de données. La région la plus intéressante est alors celle dont la probabilité p_R^{data} est la plus faible : celle-ci est notée p_{min}^{data}.

6.1.3 Signification statistique d'une région

Une fois que la région présentant l'écart maximal entre les données et les simulations Monte-Carlo est trouvée, nous voulons déterminer à quel point cette déviation est significative. Nous introduisons alors la notion d'expérience hypothétique similaire ("hse" en anglais) : N^{hse} événements sont tirés aléatoirement suivant la distribution des prédictions Monte-Carlo pour simuler la réalisation d'une expérience hypothétique. Les erreurs systématiques des distributions Monte-Carlo sont prises en compte, ainsi que les fluctuations poissoniennes du nombre d'événements observés dans les données. Les erreurs systématiques sont traitées de façon corrélée sur l'ensemble des spectres en ΣE_T et M_{inv}. Nous calculons alors la probabilité $p_R^{hse,i}$ pour chaque région R de chaque expérience hypothétique hse, i. Cette prescription nous permet de calculer le nombre d'expériences hypothétiques similaires, pour lesquelles nous observons $p_{min}^{hse,i} < p_{min}^{data}$, la probabilité $p_{min}^{hse,i}$ correspondant à la probabilité la plus faible observée lors de la réalisation d'une expérience i. Le taux S est donné par la relation :

$$S = \frac{1}{N_{hse}} \sum_{i=1}^{N_{hse}} \quad \Theta(p_{min}^{data} - p_{min}^{hse,i}) \tag{6.4}$$

avec $\Theta(x) = 0$ pour x < 0 et $\Theta(x) = 1$ pour x > 0. En d'autres termes, il s'agit en fait de compter le nombre d'expériences dans lequel nous aurions observé une déviation, située n'importe où dans les distributions, plus grande que celle vue dans les données. Ceci nous renseigne donc sur le caractère statistique de la déviation observée et nous dit si elle peut être expliquée par les prédictions Monte-Carlo et les erreurs associées. La variable S est appelée signification statistique. Dans la limite $S = 0$, un excès significatif a été observé car aucune expérience hypothétique similaire avec une probabilité associée $p_{min}^{hse,i}$ plus petite que dans les données n'a été trouvée. Dans la limite d'un grand nombre d'événements prédits, et quand la statistique gaussienne s'applique, S est distribuée uniformément entre 0 et 1. S peut être transformée en une variable, S_G, distribuée selon une loi gaussienne, en utilisant la relation suivante :

$$S = \frac{1}{\sqrt{2\pi}} \int_{S_G}^{\infty} e^{-t^2/2} dt \tag{6.5}$$

$S > 0,5$ correspond alors à $S_G < 0$ et $S < 0,5$ correspond à $S_G > 0$. Les excès les plus significatifs correspondent donc à un grand S_G. Plus précisément, la probabilité associée à une valeur de S_G donnée correspond à l'intégrale de la queue de distribution d'une gaussienne au-delà de S_G écarts types : une valeur $S_G = 2$ correspond par exemple à une probabilité de 2,3% que les déviations observées dans les données soit une fluctuation statistique.

6.1.4 Test de la méthode

Le choix de la région d'intérêt est le point fondamental de la prescription, et il faut montrer que la probabilité finale n'est pas biaisée par ce choix. Nous allons montrer dans un premier temps que la méthode basée sur les régions "à fenêtre glissante" est équivalente à celle utilisant les "régions intégrées". Nous utilisons les simulations Monte-Carlo et un total de 20 canaux. Afin de déterminer la signification statistique, nous remplaçons les données par des "fausses données" obtenues par tirage aléatoire suivant la distribution des simulations Monte-Carlo. Les régions sont définies à partir des médianes entre les points de "fausses données" (méthode "région données"). Le résultat est présenté sur la figure 6.2.

Nous remarquons que les deux méthodes donnent sensiblement le même résultat avec une signification statistique distribuée suivant une gaussienne. Cependant, sa valeur centrale vaut 0,13 ± 0,02 et l'écart type 0,92 ± 0,005. Ce biais provient du fait que la signification statistique montrée est calculée comme la somme des significations statistiques individuelles des canaux. Or, dans les canaux à faible statistique, le tirage poissonien asymétrique peut introduire des biais. Ce biais est donc propagé dans la signification statistique totale de tous les canaux.

Dans la suite, nous utilisons la méthode "à fenêtre glissante" qui est plus sensible aux déviations localisées. Un exemple des deux méthodes est montré pour le canal j-j-j dans la figure 6.3.

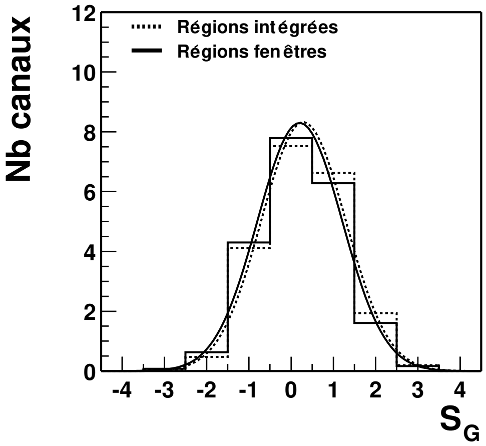


FIG. 6.2 – Distribution des canaux en fonction de la signification statistique pour les méthodes à “régions intégrées” (en pointillés) et “à fenêtre glissante” (en traits pleins). Les courbes en pointillés et en traits pleins sont le résultat d'un ajustement gaussien.

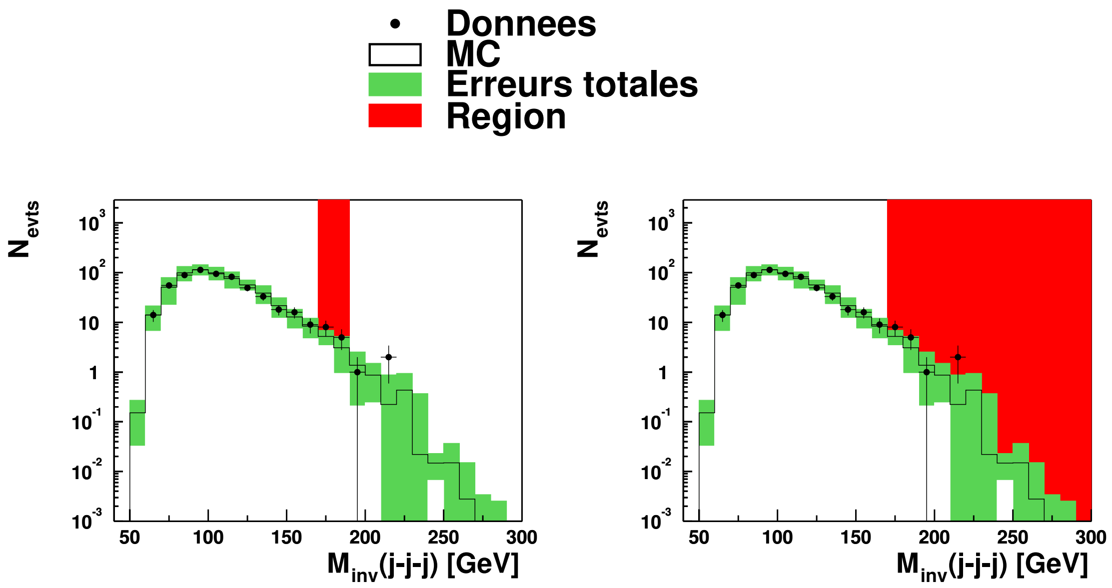


FIG. 6.3 – Canal j-j-j : région la plus intéressante “à fenêtre glissante” (à gauche) et “région intégrée” (à droite).

Dans un second temps, nous avons comparé les résultats utilisant des régions définies à partir de la médiane entre les points de données et celles définies à partir de la distribution Monte-Carlo. La distribution du nombre de canaux en fonction de la signification statistique est montrée sur la figure 6.4. A nouveau, les deux méthodes donnent des résultats sensiblement équivalents avec des distributions gaussiennes centrées sur 0,19 ± 0,004 et d'écart type 0,92 ± 0,003

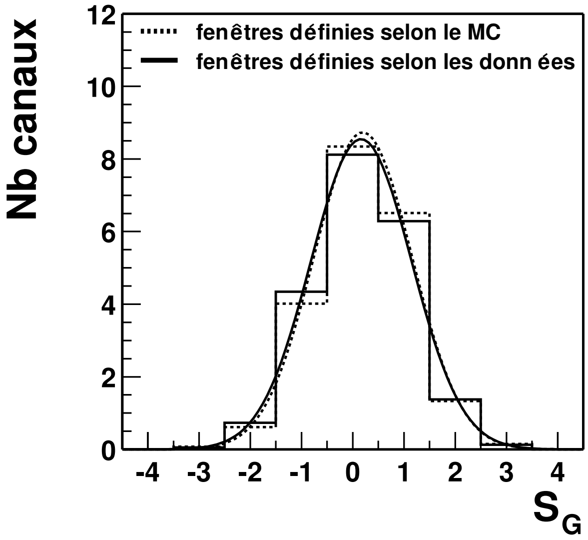


FIG. 6.4 – Distribution des canaux en fonction de la signification statistique pour les méthodes à "régions données" (en pointillés) et "régions MC" (en traits pleins).

6.2 Résultats

Dans le chapitre précédent, nous avons analysé 24 canaux en détails. Cependant, pour les canaux e-γ-γ, j-j-j-j, e-j-j-j-j et j-j-j-j-j, comme nous l'avons vu, le taux d'événements prédits n'est pas suffisament fiable pour être utilisé dans l'analyse statistique. Ces canaux seront écartés dans la suite. La méthode utilisant les régions "à fenêtre glissante" et "régions données" a été appliquée sur les distributions en masse invariante et ΣE_T des 20 canaux. La distribution Monte-Carlo a été obtenue, comme dans le paragraphe 6.1.4, en remplaçant les données par des "fausses données" obtenues par tirage aléatoire suivant la distribution des simulations Monte-Carlo et ceci pour chaque canal. L'ensemble de l'algorithme est ensuite appliqué sur ces fausses données. Les résultats en utilisant les distributions en masse invariante et de ΣE_T sont montrés respectivement dans la figure 6.5 et 6.6.

Nous observons que dans les deux cas, la distribution de la signification statistique des données est compatible avec celle des expériences Monte-Carlo. Des petites différences entre les résultats des deux distributions sont visibles. Cependant, elles sont comprises dans les erreurs statistiques. Dans deux canaux, la signification sta-

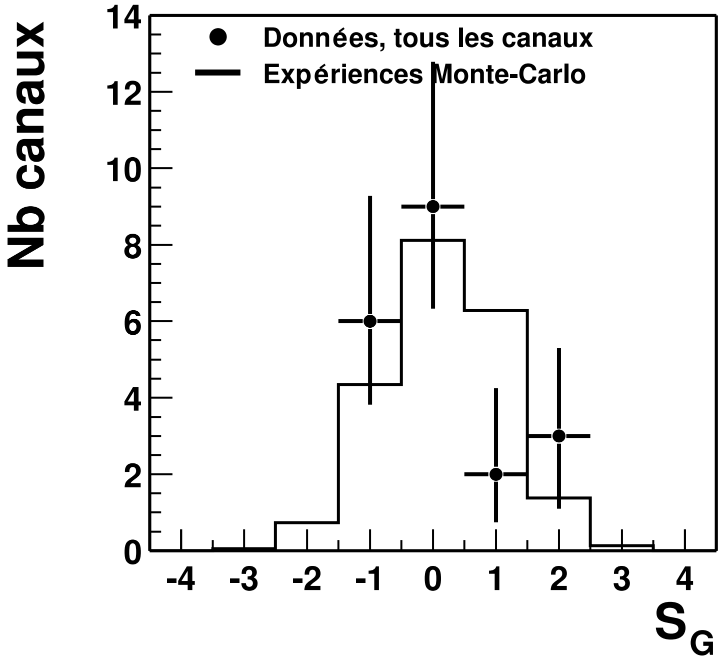


FIG. 6.5 – Distribution du nombre de canaux en fonction de la signification statistique pour les données et les prédictions avec la méthode “à fenêtre glissante” en utilisant la masse invariante.

tistique est supérieure à 2. Il s'agit des canaux e-e (2,1) et μ-j-ν (2,1). Les régions trouvées sont illustrées dans la figure 6.7. Dans des analyses dédiées [11] et [16], ces déviations ont déjà été observées. Le troisième canal présentant la déviation la plus significative est j-j-γ. Elle est de 1,5 pour les distributions en masse invariante et de 0,9 pour les distributions en ΣE_T. Ceci explique la fluctuation d'un événement observé entre les figures 6.5 et 6.6 dans le bin à 2. La probabilité d'observer parmi 20 canaux au moins 2 avec $S_G \geq 2$ est de 7,5%. Les distributions de masse invariante de tous les canaux sont présentées dans l'annexe C.

Enfin, afin de quantifier si la distribution en S_G des données est compatible avec la distribution des simulations Monte-Carlo, nous avons procédé à un test de compatibilité de χ^2. Nous définissons le χ^2 suivant la relation :

$$\chi^2 = \frac{\sum_{i=1}^{N_{canaux}} (S_i^{data} - S_i^{MC})^2}{S_i^{MC}} \tag{6.6}$$

Nous trouvons une valeur de $\chi^2/n.d.l. =$ 1,4 proche de la valeur optimale de 1,0. La probabilité associée d'observer ce résultat est de 0,85.

Le tableau 6.1 résume pour chaque canal la signification statistique S_G, la probabilité de fluctuation p_{min}^{data}, les bornes inférieures et supérieures de la région en masse choisie ainsi que les nombre d'événements observés et prédits pour la région d'excès la plus intéressante. Nous avons également étudié les régions de déficit. Le résultat est montré dans la table 6.2. Aucun déficit significatif n'est trouvé. L'absence

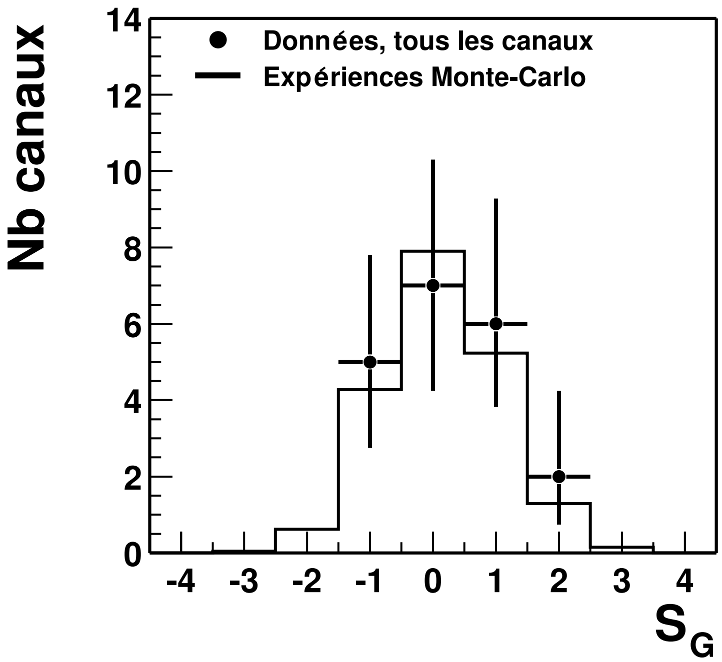


FIG. 6.6 – Distribution du nombre de canaux en fonction de la signification statistique pour les données et les prédictions avec la méthode "à fenêtre glissante" en utilisant la variable ΣE_T.

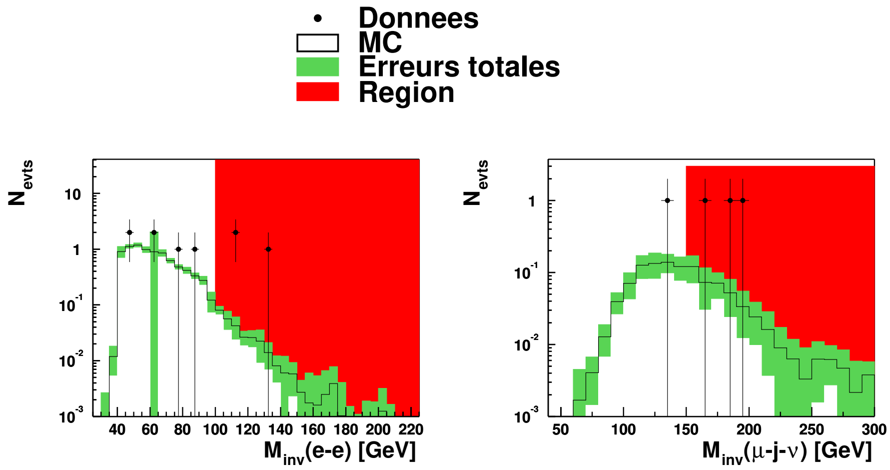


FIG. 6.7 – Région de plus grand intérêt dans les canaux e-e (à gauche) et μ-j-ν (à droite).

de données dans le canal j-ν-γ induit une probabilité nulle. Comme dans aucune expérience Monte-Carlo une probabilité inférieure à celle des données ne peut être générée, le canal j-ν-γ ne peut pas être considéré.

Canal	S_G	p_{min}^{data}	M_{inv}^{min} [GeV]	M_{inv}^{max} [GeV]	Données	MC
e-e	2.1	$4{\cdot}10^{-3}$	100	225	3	0.3 ± 0.1
e-μ	-0.5	0.86	40	250	3	4.9 ± 0.7
e-j	-0.6	0.55	205	235	52	52.1 ± 18.5
e-ν	-0.8	0.56	90	250	4	4.0 ± 0.9
e-γ	0.7	$4{\cdot}10^{-2}$	80	110	31	21.3 ± 1.5
μ-μ	$3{\cdot}10^{-2}$	0.37	60	180	2	1.3 ± 0.3
μ-j	-0.2	0.39	70	300	3	2.2 ± 0.5
j-j	-0.2	0.38	220	300	2	1.3 ± 0.8
j-ν	-0.4	0.21	210	230	17	11.1 ± 5.8
j-γ	-0.6	0.35	160	330	1	0.4 ± 0.1
e-j-j	-0.6	0.29	110	130	98	86.0 ± 19.3
e-j-ν	1.0	0.25	50	300	2	1.0 ± 0.3
e-j-γ	0.3	0.15	50	220	6	3.3 ± 0.9
μ-j-ν	2.1	10^{-2}	150	300	3	0.4 ± 0.2
j-j-j	0.1	0.17	170	190	13	8.3 ± 3.3
j-j-ν	-0.4	0.23	200	330	4	2.1 ± 1.3
j-j-γ	1.6	10^{-2}	100	120	5	1.1 ± 0.4
j-ν-γ	-0.7	1.0	50	999	0	1.6 ± 0.2
e-j-j-j	0.3	0.26	160	180	2	0.9 ± 0.5
j-j-j-ν	-0.3	0.89	50	320	1	2.6 ± 0.9

TAB. 6.1 – Résultats de l'algorithme de la méthode de régions à "fenêtre glissante" pour des distributions de masse invariante. La signification statistique S_G, la probabilité de fluctuation p_{min}^{data}, les bornes inférieure et supérieure de la région en masse choisie ainsi que le nombre d'événements observés et prédits pour la région d'excès la plus intéressante sont reportés.

Canal	S_G	p_{min}^{data}	M_{inv}^{min} [GeV]	M_{inv}^{max} [GeV]	Données	MC
e-e	0.5	0.25	50	85	3	5.4 ± 1.2
e-μ	1.0	0.29	40	250	3	4.9 ± 0.7
e-j	0.2	0.19	95	115	5610	6053.1 ± 499.8
e-ν	$5{\cdot}10^{-3}$	0.21	70	90	2	4.4 ± 0.9
e-γ	-1.9	0.47	60	80	39	40.1 ± 1.8
μ-μ	-0.6	0.63	50	70	1	1.3 ± 0.3
μ-j	0.1	0.30	50	70	1	2.5 ± 0.5
j-j	-0.4	0.36	150	170	43	51.7 ± 21.9
j-ν	0.1	0.10	160	180	67	91.4 ± 16.4
j-γ	-0.1	0.19	50	120	32	43.9 ± 11.0
e-j-j	0.2	0.10	70	90	21	31.1 ± 5.3
e-j-ν	-1.0	0.87	50	150	1	0.6 ± 0.2
e-j-γ	-2.0	0.69	130	150	1	1.1 ± 0.5
μ-j-ν	-2.1	0.86	50	150	1	0.6 ± 0.2
j-j-j	-0.4	0.33	120	150	100	117.5 ± 36.4
j-j-ν	-0.4	0.21	110	160	27	35.8 ± 7.9
j-j-γ	10^{-2}	0.26	130	300	1	3.2 ± 2.0
j-ν-γ	-	-	-	-	-	-
e-j-j-j	0.1	0.40	50	160	1	2.1 ± 0.8
j-j-j-ν	0.8	0.31	50	320	1	2.6 ± 0.9

TAB. 6.2 – Résultats de l'algorithme de la méthode de régions à "fenêtre glissante" pour des distributions de masse invariante. La signification statistique S_G, la probabilité de fluctuation p_{min}^{data}, les bornes inférieure et supérieure de la région en masse choisie ainsi que le nombre d'événements observés et prédits pour la région de déficit la plus intéressante sont reportés.

6.3 Sensibilité : exemple de la production anormale de quark top

Dans les recherches dédiées de nouvelle physique, les techniques de sélection sont adaptées aux caractéristiques des événements attendus pour un modèle particulier. L'analyse développée est alors spécialement sensible au signal recherché. Dans le cadre de notre analyse modèle-indépendante et de l'interprétation statistique, nous avons aussi étudié la sensibilité à un éventuel signal de nouvelle physique. Nous avons choisi l'exemple de production anormale de quark top car ses produits de désintégation peuvent être reconstruits dans plusieurs canaux leptoniques et de jets et qu'une étude dédiée est en cours de publication [64].

La section efficace de production de quark top dans le Modèle Standard à HERA étant très faible - nous attendons de l'ordre de 0,1 événement pour la totalité des données prises entre 1994 et 2000 - il existe certains modèles [71, 72] prédisant une section efficace plus importante dans des processus de changement de saveur par courant neutre. Dans ces modèles, le quark top serait produit par couplage anormal (essentiellement $tu\gamma$ à HERA). A partir des résultats obtenus dans la partie 5.5 et la méthode d'interprétation statistique développée, nous pouvons étudier notre sensibilité à ce type de nouvelle physique.
Pour ce faire, des événements de fausses données ont été créés en sommant aux prédictions Monte-Carlo une contribution d'un nombre d'événements variable, provenant de la désintégration d'un quark top.

Nous avons tout d'abord calculé, à partir d'une simulation Monte-Carlo de production de quark top, l'efficacité d'identification pour chacun des canaux. Celle-ci est déterminée en utilisant toutes les coupures de l'analyse décrites dans le paragraphe 5.1 et suivants. Nous n'avons pas considéré le canal de désintégration du $W^{\pm}$ en tau. En effet, les leptons tau ne sont pas reconstruits dans cette analyse et la proportion d'événements provenant de ce canal de désintégration du $W^{\pm}$ dans les canaux e-j-ν et μ-j-ν de notre analyse est négligeable. L'efficacité de notre sélection vaut, en incluant les rapports de branchement [73] du $W^{\pm}$, à peu près 2% dans les canaux leptoniques e-j-ν et μ-j-ν, et atteint 26% dans le canal j-j-j. Bien que la topologie du quark top soit à trois corps, l'un des corps peut ne pas être sélectionné. Par exemple, le canal j-j peut recueillir des événements top où l'un des jets ne satisfait pas les critères de sélection et le canal j-ν ceux où l'un des leptons n'est pas accepté. L'efficacité de ces deux canaux est respectivement 28% et 6%. Nous observons également des états finaux à quatre corps, dus à des radiations de gluons, comme par exemple j-j-j-j, où l'efficacité vaut environ 1%.

Ensuite, nous procédons à deux tirages pour simuler la présence d'un signal de quark top :

- Premièrement, nous déterminons, dans chaque canal, de “fausses données” en tirant des événements aléatoirement selon la distribution des prédictions Monte-Carlo des processus du Modèle Standard, en tenant compte des erreurs totales.

- Deuxièmement, nous tirons aléatoirement $N_{top,i}$ événements qui sont ajoutés, en tenant compte de l'efficacité de chaque canal i, aux "fausses données" du lot précédent. Les $N_{top,i}$ événements sont tirés selon la distribution de masse invariante reconstruite des corps.

Le nombre final d'événements de quark top ajoutés aux prédictions du Modèle Standard dans le canal i, $N_{top,i}$, est calculé, en tenant compte de l'efficacité du canal i et du nombre initial N_{top}^{ini}, suivant la relation :

$$N_{top,i} = N_{top}^{ini} \cdot \epsilon_i \tag{6.7}$$

La figure 6.8 montre un exemple d'un tirage aléatoire de la distribution en masse invariante reconstruite dans le canal j-j-j avec $N_{top}^{ini} = 120$ correspondant à une section efficace $\sigma(top) = 1\ pb$, donc $N_{top,jjj} = 29$ (points ouverts). Nous remarquons un excès de "fausses données" (points pleins) dans un intervalle de 30 GeV autour de la masse invariante du quark top. Nous avons donc constitué un lot de "fausses données" à partir de la distribution Monte-Carlo des processus du Modèle Standard et du signal de quark top. Ce lot peut être alors comparé aux prédictions et nous pouvons appliquer notre méthode statistique et en déterminer la sensibilité aux déviations observées dans les différents canaux.

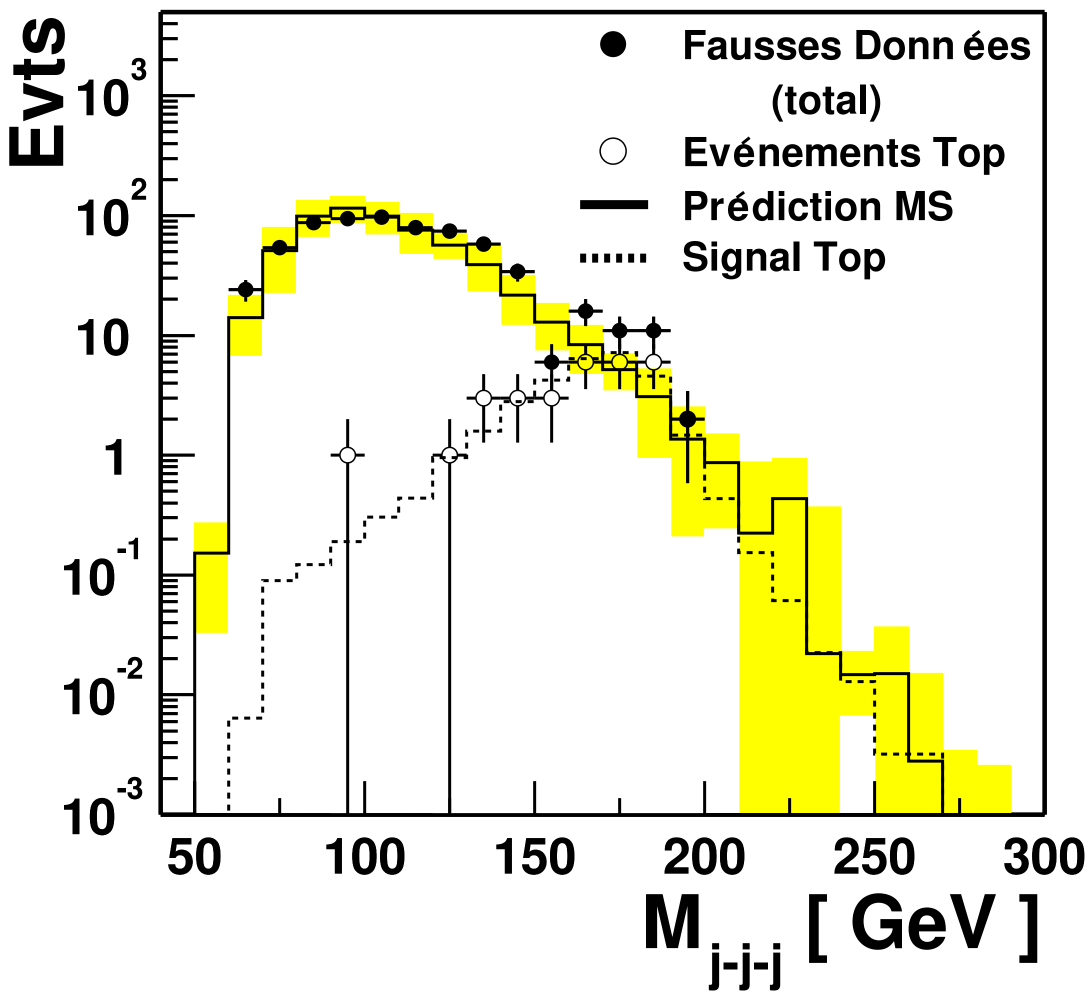


FIG. 6.8 – Distribution de la masse invariante du canal j-j-j à partir de "fausses données" (en points pleins) contenant du signal de quark top (points ouverts) et les processus du Modèle Standard (histogramme noir).

Le nombre d'événements N_{top}^{ini} peut être relié à la section efficace de production à travers la relation $\sigma(top) = \mathrm{N}_{top}^{ini} \; / \; \mathcal{L}_{int}$, où $\mathcal{L}_{int}$ est la luminosité intégrée, donc 118,3 pb^{-1} pour la période de prises de données entre 1994 et 2000.

La signification statistique S_G de la déviation observée en fonction de la section efficace de production $\sigma(top)$ alors obtenue est montrée sur la figure 6.9 pour les trois canaux e-j-ν, μ-j-ν et j-j-j. Elle augmente en fonction de $\sigma(top)$, jusqu'à atteindre 3,0 à $\sigma(top) = 1{,}4\ pb$ dans le canal μ-j-ν. Le nombre fini de tirages aléatoires de N_{top}^{ini}, et donc les imprécisions statistiques qui en résultent provoquent des fluctuations comme par exemple celles visibles dans le canal μ-j-ν pour $0,9 < \sigma(top) < 1,1\ pb$. Malgré le fait que le "bruit de fond" irréductible dans le canal j-j-j soit important, la sensibilité est du même ordre de grandeur que dans les canaux leptoniques :

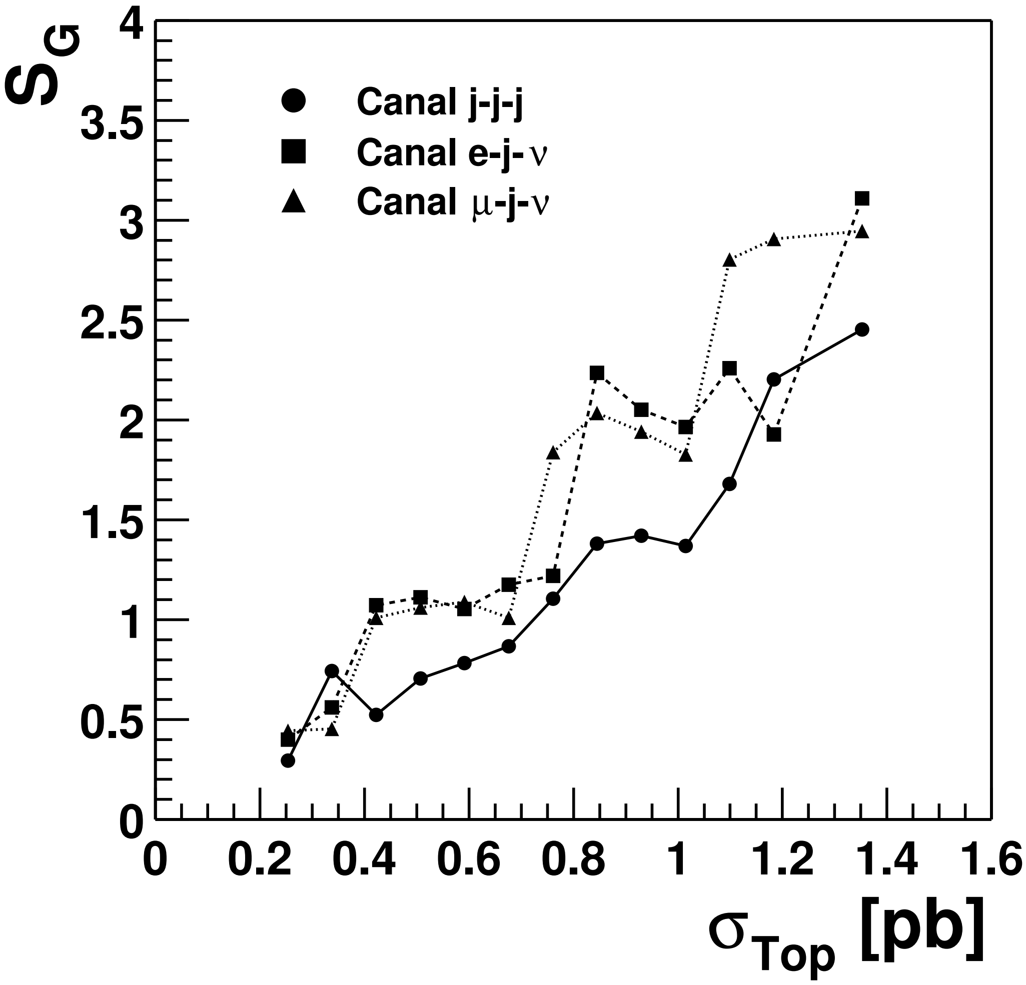


FIG. 6.9 – La signification statistique S_G en fonction de la section efficace du quark top dans les canaux e-j-ν, μ-j-ν et j-j-j.

l'efficacité plus importante (d'un facteur 10 environ) et la baisse importante de la contribution du Modèle Standard en fonction de la masse invariante permettent d'atteindre une sensibilité comparable aux canaux leptoniques.

En utilisant les limites d'exclusion, déduites dans les processus de désintégration radiative de quark top, établies à 95% de niveau de confiance par l'expérience CDF [74] et les limites associées du couplage $tu\gamma$ [72], la limite supérieure de production à $HERA$ serait environ 1 *pb*. Les résultats de la figure 6.9 montrent que nous pourrions observer, avec notre analyse et la méthode d'interprétation statistique, une déviation de S_G à 2 correspondant à une probabilité de 2,3% dans les canaux μ-j-ν et e-j-ν.

En combinant les canaux de désintégration leptonique et hadronique, des limites supérieures plus strictes (à 95% niveau de confiance) de production anormale de quark top ont été déduites dans les analyses dédiées de $ZEUS$ [18] et $H1$ [64] :

$$\sigma(ep \to etX) < 0,225\ pb \quad (ZEUS) \tag{6.8}$$
$$\sigma(ep \to etX) < 0,6\ pb \quad (H1) \tag{6.9}$$

Ainsi, l'ensemble de cette analyse est moins sensible qu'une analyse dédiée et optimisée pour la recherche d'un signal spécifique. Cependant, elle aurait pu permettre d'obtenir une indication sur la présence d'un signal top non exclu par les autres accélérateurs, ce qui est l'un des buts d'une telle analyse. Bien sûr, il serait intéressant par la suite de tester la sensibilité de notre analyse à d'autres modèles de processus au-delà du Modèle Standard, comme par exemple des modèles de $SUSY$ qui peuvent contribuer dans plusieurs canaux.

6.4 Conclusion et perspectives

En résumé, nous avons présenté une méthode d'interprétation statistique permettant de déterminer les régions d'écart maximal entre les données et les simulations Monte-Carlo. En utilisant les résultats de l'analyse présentée dans le chapitre précédent, il a été montré que les déviations maximales observées dans les différents canaux sont compatibles avec les fluctuations liées aux erreurs systématiques des prédictions Monte-Carlo. Nous avons montré que les déviations significatives des canaux e-e et μ-j-ν, observées dans les analyses dédiées [11] et [16], sont retrouvées et ont pu être quantifiées. Aucune déviation significative supplémentaire n'a été observée.

Dans le cadre d'un modèle particulier de production anormale de quark top par couplage $tu\gamma$, nous avons étudié la sensibilité de notre analyse et de la méthode d'interprétation statistique à des signaux de nouvelle physique. Une signification statistique correspondant à une probabilité de 0,14% dans le canal μ-j-ν serait atteinte pour une section efficace de $\sigma(top)$= 1,4 pb. La sensibilité de notre recherche générique est cependant inférieure à celle d'une analyse dédiée.

Afin de diminuer les erreurs statistiques et donc d'améliorer la fiabilité de la méthode à grande masse et à grand ΣE_T, nous devons générer des lots de simulations Monte-Carlo plus conséquents. Ensuite, les études de sensibilité peuvent être étendues à d'autres modèles particuliers, comme par exemple des modèles de Supersymétrie. Ceux-ci sont spécialement intéressants lorsque les particules supersymétriques se désintègrent en cascades, en contribuant à plusieurs états finaux.

Conclusion et perspectives

L'étude présentée dans cette thèse concerne l'exploration des limites à haute énergie de l'espace de phase accessible à HERA, par l'analyse de tous les événements enregistrés entre 1994 et 2000 contenant des électrons, des jets, des muons, des photons et des neutrinos de grande énergie transverse. Nous avons vérifié que la production de ces topologies est en général bien décrite par les processus du Modèle Standard et recherché un possible excès dans les configurations cinématiques peu probables pour ce dernier.

L'analyse a été effectuée dans le nouvel environnement d'analyse orienté objet de H1. Il s'agit de la première analyse de H1 utilisant ces nouveaux outils. Une contribution importante, lors de la thèse, a été la mise en place des structures C++ essentielles et la validation des identificateurs de particules. Dans ce cadre, un algorithme de reconstruction hadronique améliorant, par rapport à l'algorithme implémenté en fortran, la résolution en énergie de l'état hadronique final à bas P_T, a été élaboré. Des algorithmes d'identification de hadrons et de reconstruction de jets existant en fortran ont également été implémentés et intégrés en tant qu'algorithmes par défaut. Les logiciels de cet environnement orienté objet sont devenus les outils d'analyse standard aussi bien pour le contrôle de qualité que pour l'analyse physique des données de la phase II de HERA.

Notre analyse a porté sur l'étude des spectres d'énergie transverse dans 24 états exclusifs formés à partir de deux, trois, quatre ou cinq corps en utilisant l'ensemble des données prises entre 1994 et 2000 : une bonne description des données par les prédictions basées sur la simulation des processus du Modèle Standard est observée dans la plupart des canaux. Mis à part les déviations connues des canaux e-e et μ-j-ν à grande énergie transverse, aucun autre écart significatif n'a été trouvé. Les difficultés rencontrées lors de l'analyse des canaux e-γ-γ et des canaux contenant plus de trois jets montrent les limites des générateurs actuels. Ces canaux sont spécialement intéressants car, pour l'instant, aucune analyse dédiée n'en a été effectuée. Pour les canaux à grande statistique, l'échantillon réduit obtenu en utilisant des variables toplogiques, afin de sélectionner des régions de l'espace de phase contenant des événements atypiques dans le cadre du Modèle Standard, ne montre aucune déviation significative.

Afin de quantifier l'accord global au Modèle Standard, une méthode d'interprétation statistique tenant compte des erreurs systématiques et statistiques a été développée. En raison de l'incertitude sur les prédictions dans certains canaux, 20 canaux sont retenus dans l'analyse statistique finale. L'accord global entre les données et les prédictions Monte-Carlo a été estimé à partir d'un test de χ^2 : la valeur de χ^2 par degré de liberté est 1,4 et la probabilité associée 0,85. Seuls, deux canaux présentent des déviations significatives de probabilité individuelle inférieure à 2,3% : dans les canaux μ-j-ν et e-e, des événements atypiques à grande énergie transverse sont observés. Leur présence a déjà été constatée dans les analyses dédiées aux événements avec leptons isolés et grande impulsion transverse manquante [16], et la production de paires d'électrons [11]. La probabilité d'observer deux canaux ou plus, parmi vingt, présentant une telle déviation est de 7,5%. Nous avons aussi étudié la sensibilité de cette méthode à un modèle particulier de production anormale de quark top par couplage $tu\gamma$. Une signification statistique correspondant à une probabilité de 0,14% dans le canal μ-j-ν serait atteinte pour une section efficace de $\sigma(top)$= 1,4 *pb*. En considérant les limites actuelles, mises par les expériences du LEP et du TEVATRON, correspondant à environ 1 *pb* à HERA, cette analyse pourrait donc permettre d'obtenir une indication sur la présence d'un signal. Cependant, cette méthode est moins sensible que les recherches dédiées de quark top effectuées par H1 et ZEUS.

Depuis 2002, HERA accumule de nouvelles données. Pour cette phase II de HERA, l'analyse sera utilisée pour suivre en ligne les taux d'événements de grande énergie transverse. Dans l'optique d'une recherche systématique, il serait possible d'augmenter la sensibilité à la production de leptons, en abaissant par exemple le critère d'énergie transverse minimale à 10 GeV. Une extension possible de l'analyse est l'inclusion d'autres types de particules, comme par exemple les leptons taus ou les jets de quark *b*.

Bibliographie

[1] C. Adloff et al. Measurement of neutral and charged current cross sections in electron proton collisions at high Q**2. *Eur. Phys. J.*, C19 p269–288, (2001), hep-ex/0012052.

[2] B. Heinemann. *Measurement of charged current and neutral current cross sections in positron proton collisions at s**(1/2) approx. 300-GeV.* Thèse de Doctorat, Université de Hambourg, (1999). DESY-THESIS-1999-046.

[3] C. Adloff et al. Measurement of neutral and charged current cross-sections in positron proton collisions at large momentum transfer. *Eur. Phys. J.*, C13 p609–639, (2000), hep-ex/9908059.

[4] C. Adloff et al. Dijet production in charged and neutral current e+ p interactions at high Q**2. *Eur. Phys. J.*, C19 p429–440, (2001), hep-ex/0010016.

[5] C. Adloff et al. Measurement and QCD analysis of neutral and charged current cross sections at HERA. (2003), hep-ex/0304003.

[6] C. Adloff et al. Measurement of the inclusive di-jet cross section in photoproduction and determination of an effective parton distribution in the photon. *Eur. Phys. J.*, C1p97–107, (1998), hep-ex/9709004.

[7] P. Bate. *High Transverse Momentum 2-jet and 3-jet Cross-section Measurements in Photoproduction.* Thèse de Doctorat, Université de Manchester, (1999).

[8] C. Adloff et al. Measurement of dijet cross sections in photoproduction at HERA. *Eur. Phys. J.*, C25 p13–23, (2002), hep-ex/0201006.

[9] S. Kermiche. Study of Compton QED at HERA : Luminosity and calibration of the H1 detector. Note LAL-94-14.

[10] M. Hennecke. *Experimental Investigation of Wide Angle Bremsstrahlung in the H1-Detector.* Thèse de diploma, Université RWTH Aachen, (1999).

[11] A. Aktas et al. Multi-electron production at high transverse momenta in e p collisions at HERA. (2003), hep-ex/0307015 (accepté par Eur Phys J).

[12] B. Leissner. Muon pair production. *Acta Phys. Polon.*, B33 p3275–3280, (2002).

[13] S. Baffioni. *Identification du lepton tau a HERA.* Rapport de stage de DEA, Université de Grenoble, (2000).

[14] N. Arteaga-Romero, C. Carimalo, and P. Kessler. High P(t) lepton pair production at e p colliders : Comparison between various production mechanisms. *Z. Phys.*, C52p289–296, (1991).

[15] A. Schoening. *Untersuchung von Prozessen mit virtuellen und reellen W+- -Bosonen am H1-Detektor bei HERA*. Thèse de Doctorat. DESY F11/F22-96-02.

[16] V. Andreev et al. Isolated electrons and muons in events with missing transverse momentum at HERA. *Phys. Lett.*, B561 p241–257, (2003), hep-ex/0301030.

[17] M. Schneider. *Recherche d'événements contenant un lepton isolé et de l'impulsion transverse manquante sur le collisionneur HERA avec le détecteur H1*. Thèse de Doctorat, Université de Strasbourg, (2003).

[18] Search for single-top production in $e^{\pm}p$ collisions at HERA. *Phys. Lett.*, B559 p153–170, (2003), hep-ex/0302010.

[19] J. Breitweg et al. W production and the search for events with an isolated high-energy lepton and missing transverse momentum at HERA. *Phys. Lett.*, B471 p411–428, (2000), hep-ex/9907023.

[20] I. A. D'Souza and C. S. Kalman. *Preons : Models of leptons, quarks and gauge bosons as composite objects*. World Scientific, Singapore, (1992).

[21] J. Wess and B. Zumino. Supergauge transformations in four-dimensions. *Nucl. Phys.*, **B70** p39–50, (1974).

[22] Nima Arkani-Hamed, Savas Dimopoulos, and G. R. Dvali. The hierarchy problem and new dimensions at a millimeter. *Phys. Lett.*, **B429** p263–272, (1998), hep-ph/9803315.

[23] H. Georgi and S. L. Glashow. Unity of all elementary particle forces. *Phys. Rev. Lett.*, **32** p438–441, (1974).

[24] M. J. Duff. The world in eleven dimensions : A tribute to Oskar Klein. *Prépublication*, (2001), hep-th/0111237.

[25] B. Abbott et al. Search for new physics in e muon X data at D0 using Sherlock : A quasi model independent search strategy for new physics. *Phys. Rev.*, D62 p092004, (2000), hep-ex/0006011.

[26] B. Abbott et al. A quasi-model-independent search for new high p(T) physics at D0. *Phys. Rev. Lett.*, 86 p3712–3717, (2001), hep-ex/0011071.

[27] M. Krasny et al. Generic Analysis of Large E_T Processes in Electron-Proton Scattering at HERA. *Note interne H1*, **H1-IN-523** , (1997).

[28] P. Zini. Qualification du calorimètre SpaCal de H1 et étude des événements de très grande énergie transverse. (1998).

[29] J. Dingfelder. *Suche nach neuen Signaturen in der e-p-Streuung bei HERA*. Thèse de diploma, Université de Heidelberg, (2000).

[30] I. Abt et al. The H1 detector at HERA. *Nucl. Instrum. Meth.*, **A386** p310–347, (1997).

[31] T. Kuhr. *Rekonstruktion von V0s mit dem H1-Silizium-Detektor*. Thèse de diploma, Université de Hambourg, (1998).

[32] B. Andrieu et al. The H1 liquid argon calorimeter system. *Nucl. Instrum. Meth.*, **A336** p460–498, (1993).

[33] B. Andrieu et al. Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations. *Nucl. Instrum. Meth.*, **A336** p499–509, (1993).

[34] T. Abe. http ://root.awa.tohoku.ac.jp/ tabe/grape/.

[35] T. Sjostrand. High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4. *Comput. Phys. Commun.*, **82** p74–90, (1994).

[36] G. Ingelman, A. Edin, and J. Rathsman. LEPTO 6.5 - A Monte Carlo Generator for Deep Inelastic Lepton-Nucleon Scattering. *Comput. Phys. Commun.*, 101p108–134, (1997), hep-ph/9605286.

[37] U. Baur, J. A. M. Vermaseren, and D. Zeppenfeld. Electroweak vector boson production in high-energy e p collisions. *Nucl. Phys.*, **B375** p3–44, (1992).

[38] C. Berger and P. Kandel. A new generator for wide angle bremsstrahlung. In *Workshop on Monte Carlo Generators for* HERA, (1998). Physics (Plenary Starting Meeting), Hamburg, Germany, 27-30 Apr 1998.

[39] A. Kwiatkowski, H. Spiesberger, and H. J. Mohring. HERACLES : an event generator for *ep* interactions at HERA energies including radiative processes : version 1.0. *Comp. Phys. Commun.*, **69** p155–172, (1992).

[40] B. Laforge. *Etude des événements a di-Jets dans la diffusion profondément inélastique avec large intervalle de rapidité dans l'experience H1 auprès de l'accélérateur HERA.* Thèse de Doctorat, Université de Paris-Sud, (1997).

[41] G. A. Schuler and H. Spiesberger. DJANGO : The Interface for the event generators HERACLES and LEPTO. In *Hamburg 1991, Physics at* HERA*, vol 3.*, pages 1419–1432, (1991).

[42] H. Jung. Hard diffractive scattering in high-energy *ep* collisions and the Monte Carlo generation RAPGAP. *Comp. Phys. Commun.*, **86** p147–161, (1995).

[43] T. Sjostrand. PYTHIA 5.7 and JETSET 7.4 : Physics and manual. *Prépublication*, (1995), hep-ph/9508391.

[44] http ://root.cern.ch.

[45] R. Brun, F. Rademakers, and S. Panacek. ROOT, an object oriented data analysis framework. Prepared for CERN School of Computing (CSC 2000), Marathon, Greece, 17-30 Sep 2000.

[46] B. Bertrand. *Skieur Magazine*, 29S, Nov. 2000.

[47] The H1 OO Physics Analysis Project. *Document interne.*

[48] V. Blobel. The BOS System. *Document interne.* (2001).

[49] FPACK : a general stand-alone package of FORTRAN77 and C programs for the machine-independent input/output of data blocks. *Document interne.*

[50] http ://www.cvshome.org/docs/manual/index.html.

[51] J.-F. Laporte. *Diffusion profondément inélastique à HERA et calibration absolue de la mesure en énergie d'un électron dans le calorimètre à Argon liquide de l'expérience H1.* Thèse de Doctorat, Université de Paris-Sud, (1991).

[52] L. West. How to use the Heavy Flavour Working Group Track, Muon and Electron Selection Code. *Document interne.* (1997).

[53] B. Portheault. Noise suppression in OO, status and plans. *Document interne (2003).*

[54] B. Heinemann. Measurement of charged current and neutral current cross sections in positron proton collisions at s**(1/2) approx. 300-GeV. pages 51–54, (1999). DESY-THESIS-1999-046.

[55] J. Marks. HFS - a software package to cope with the Hadronic Final State *Document interne.*

[56] E. Chabert et al. QBGFMAR : An Updated Phan Package for Cosmic and Halo Muon Topological Rejection in High P_T Physics Analysis. *Note interne H1*, **H1-IN-556** , (1998).

[57] C. Veelken et al. H1NonepBgFinder - Rejection of Cosmic Muon and Beam-Halo Events in the H1OO Framework. *Document interne*, **H1-IN-603**, (2002).

[58] The H1 OO Physics Analysis Project home page. *https ://www-h1.desy.de/icas/oop/.*

[59] J. Marks. QESCAT - e identification software in H1PHAN. *Document interne.* (1996).

[60] P. Bruel. *Recherche d'interactions au-delà du Modèle Standard à* HERA. Thèse de Doctorat, Université de Paris-Sud, U.F.R. Scientifique d'Orsay, (1998).

[61] M. Wobisch. *Measurement and QCD Analysis of Jet Cross Sections in Deep-Inelastic Positron-Proton Collisions at sqrt(s) = 300 GeV.* Thèse de Doctorat, Université RWTH Aachen, (1999). DESY-THESIS-2000-049.

[62] S. Caron. *Jets in Photoproduction at HERA.* Thèse de Doctorat, Université RWTH Aachen, (2002). DESY-THESIS-2002-035.

[63] S. Ferron. *Mesure des sections efficaces inclusives de jets dans les collisions photon-proton à HERA.* Thèse de Doctorat, Ecole Polytechnique, (2001).

[64] J. Dingfelder. *Search for Anomalous Top Production of Single Top Quarks with the H1 Experiment at HERA.* Thèse de Doctorat, Université de Heidelberg, (2003).

[65] D. Waters. *A first estimate of* $\sigma(e^-p \to e^-\ W^\pm X)$. Thèse de Doctorat, Université d'Oxford, (1998).

[66] B. et al Olivier. Search for Doubly-Charged Higgs Production at HERA. (2002). Prepared for the 31st International Conference on High-Energy Physics (ICHEP 02), Amsterdam.

[67] The OPAL Collaboration. Search for Charged Higgs Bosons in e^+e^- Collisions at $sqrts$ = 189-209 GeV. *Preliminary OPAL results.*

[68] C. Adloff et al. A search for leptoquark bosons and lepton flavor violation in e+ p collisions at HERA. *Eur. Phys. J.*, C11 p447–471, (1999), hep-ex/9907002.

[69] M. Wobisch. First studies of three-jet production in deep-inelastic scattering. Prepared for 9th International Workshop on Deep Inelastic Scattering (DIS 2001), Bologna, Italy, 27 Apr - 1 May 2001.

[70] B. Knuteson. Search for new physics in e mu X data at the Tevatron using Sleuth : A quasi-model-independent search strategy for new physics. *Int. J. Mod. Phys.*, A16S1B p888–890, (2001).

[71] G. M. de Divitiis, R. Petronzio, and L. Silvestrini. Flavour changing top decays in supersymmetric extensions of the standard model. *Nucl. Phys.*, B504 p45–60, (1997), hep-ph/9704244.

[72] T. Han and J.A. L. Hewett. Top charm associated production in high energy e+ e- collisions. *Phys. Rev.*, D60 p074015, (1999), hep-ph/9811237.

[73] K. Hagiwara et al. Review of particle physics. *Phys. Rev.*, D66p010001, (2002).

[74] F. Abe et al. Search for flavor-changing neutral current decays of the top quark in p anti-p collisions at s**(1/2) = 1.8-TeV. *Phys. Rev. Lett.*, 80 p2525–2530, (1998).

Annexe A

Algorithme *HADROO* de reconstruction hadronique

L'algorithme de reconstruction hadronique *HADROO* sépare deux domaines en P_T de comportement différent décrits ci-dessous. Pour optimiser la mesure de l'énergie, nous avons ordonné les particules chargées par ordre croissant en P_T pour associer en premier les clusters aux traces bien mesurées de bas P_T. L'algorithme boucle donc sur les particules chargées (traces) de P_T croissant, et effectue pour chaque trace les actions suivantes (l'énergie calorimétrique correspond toujours à l'énergie résiduelle après traitements des traces précédentes) :

Domaine à bas P_T^{trace}

Il est défini par l'inégalité suivante :

$$\left(\frac{dP_T^{trace}}{P_T^{trace}}\right) < \alpha \frac{50\%}{\sqrt{E_{trace}}} \tag{A.1}$$

A bas P_T, l'algorithme se comporte exactement comme *FSCOMB* :

- L'énergie de chaque trace est comparée à l'énergie déposée dans un cylindre de 25 cm de rayon autour de l'impact de la trace dans le calorimètre électromagnétique et 50 cm de rayon dans le calorimètre hadronique. Dans tous les cas, la trace est gardée en tant qu'objet hadronique :
  - Si l'énergie de la trace est plus faible que l'énergie contenue dans le cylindre associé, l'énergie de la trace est soustraite aux clusters du cylindre.
  - Sinon, tous les clusters du cylindre associé sont supprimés.

La figure A.1 montre le comportement de l'algorithme à bas P_T. Trois clusters sont contenus dans un même cylindre autour de l'impact de la trace dans le calorimètre avec une énergie totale de 6 GeV. L'énergie de la trace correspondante est de 5 GeV. L'algorithme gardera alors la trace et créée un hadron avec une énergie de 5 GeV. L'énergie de celle-ci est soustraite aux clusters. Dans l'étape suivante, il ne reste donc plus que 1 GeV d'énergie sous forme de clusters.

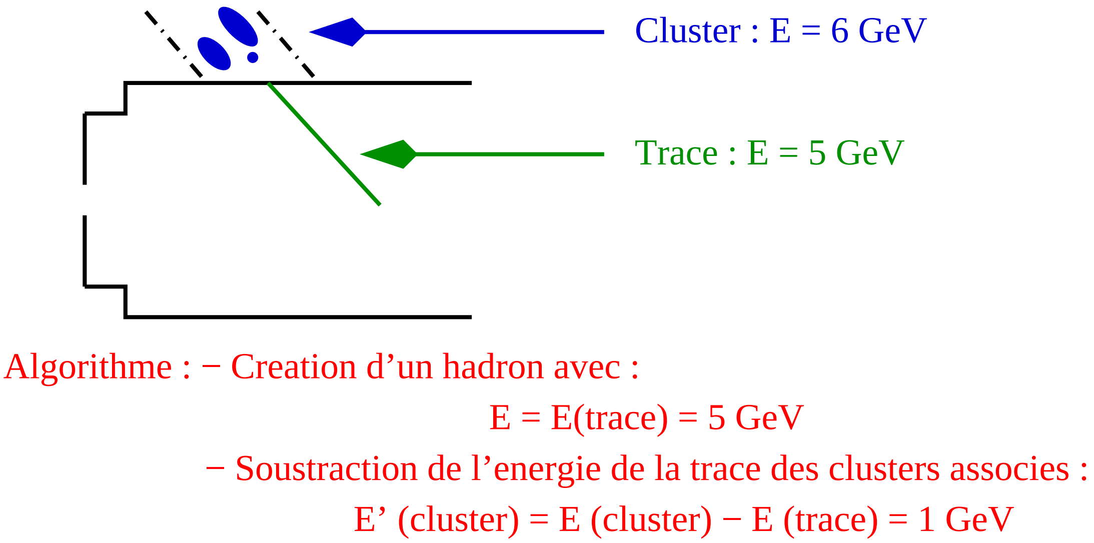


Fig. A.1 – Exemple illustrant le traitement de particules de basse impulsion transverse.

Domaine à grand P_T^{trace}

Il est défini par l'inégalité suivante :

$$\left(\frac{dP_T^{trace}}{P_T^{trace}}\right) \geq \alpha \frac{50\%}{\sqrt{E_{trace}}} \tag{A.2}$$

A grand P_T, l'algorithme compare les quantités E_{calo} et E_{trace} correspondant respectivement à l'énergie contenue dans le cylindre dans le calorimètre autour de la trace et à son énergie :

- Si E_{calo} et E_{trace} sont compatibles selon le critère dépendant du paramètre N_σ et défini dans le paragraphe *Paramètres* de la partie 4.4.3, la mesure calorimétrique est utilisée, et la trace est supprimée.
- Sinon :
  Si $E_{calo} \gg E_{trace}$: l'algorithme se comporte comme à bas P_T.
  Si $E_{calo} \ll E_{trace}$: la trace est supprimée et les clusters calorimétriques restent inchangés.

En effet, lorsque les énergies comparées sont compatibles, les hadrons sont bien mesurés mais la plupart du temps, la mesure du calorimètre est plus précise. Lorsque l'énergie de la trace est largement supérieure à l'énergie calorimétrique, il s'agit dans la majorité des cas d'une mauvaise mesure de la trace compte tenu que nous nous trouvons dans le domaine à grand P_T.

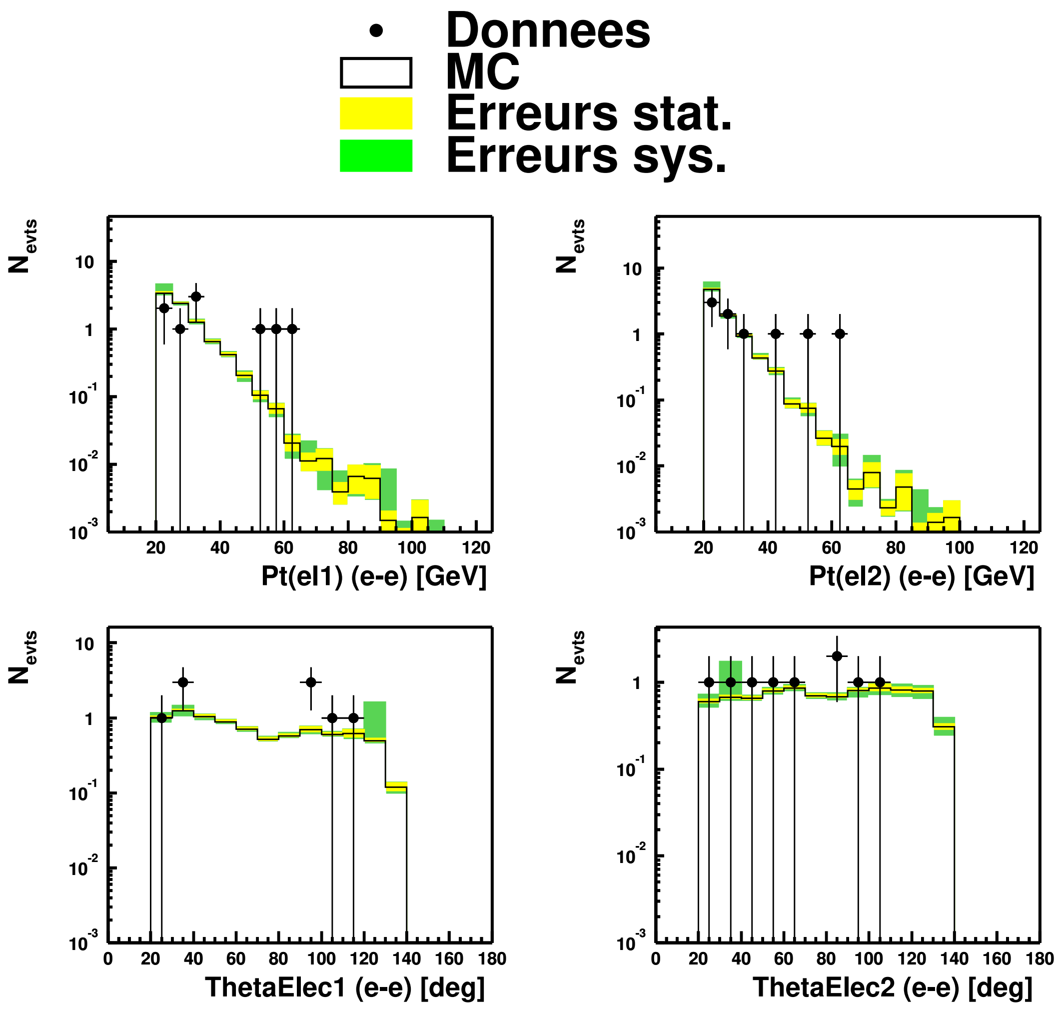
Traitement des clusters résiduels

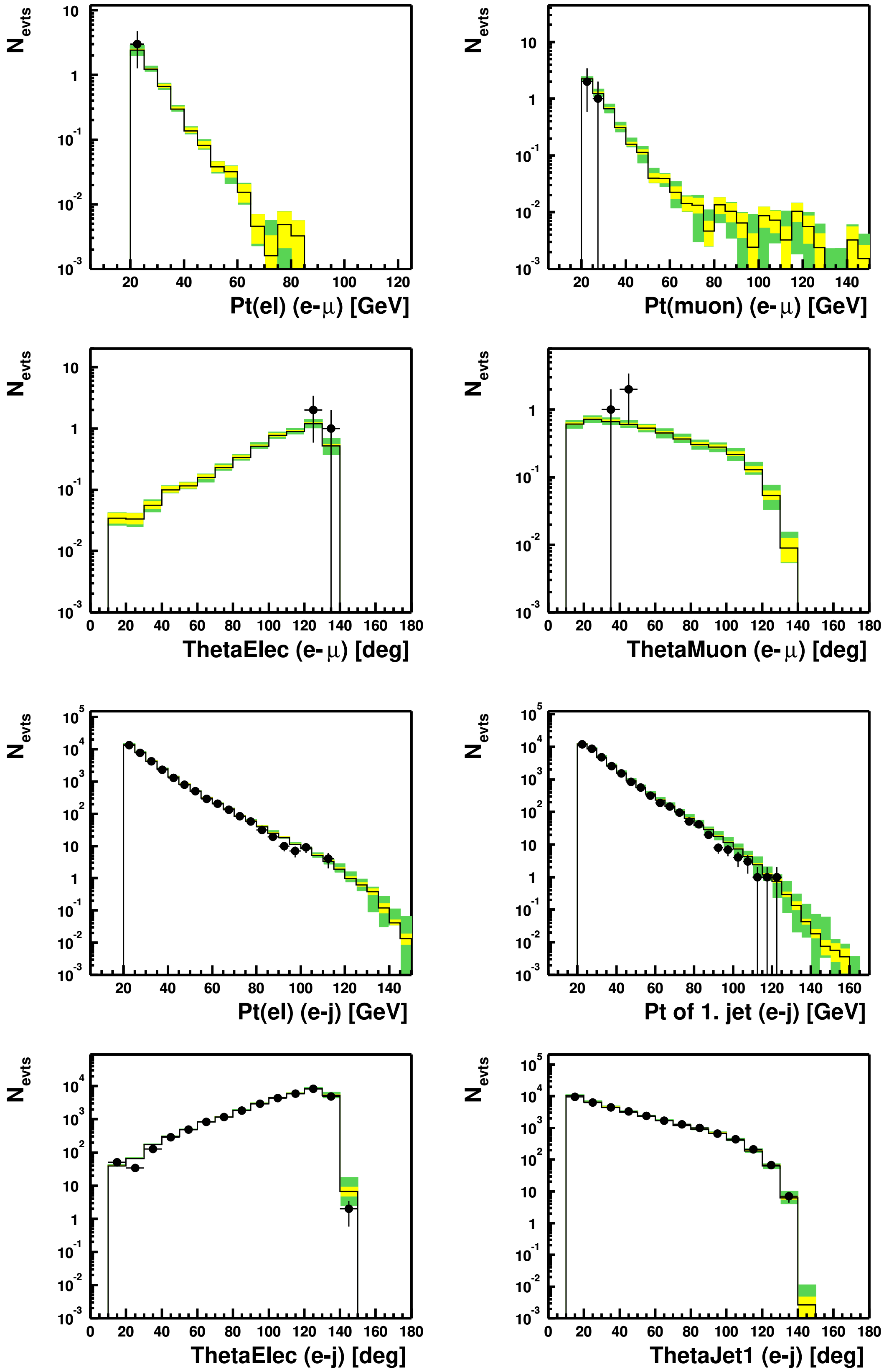
Lorsque la boucle sur les traces chargées est terminée, tous les clusters résiduels sont définis comme hadrons avec leur énergie calorimétrique. Ces hadrons correspondent en majorité aux particules neutres qui n'ont pas de trace associée, et aux particules chargées dont les traces étaient de mauvaise qualité.

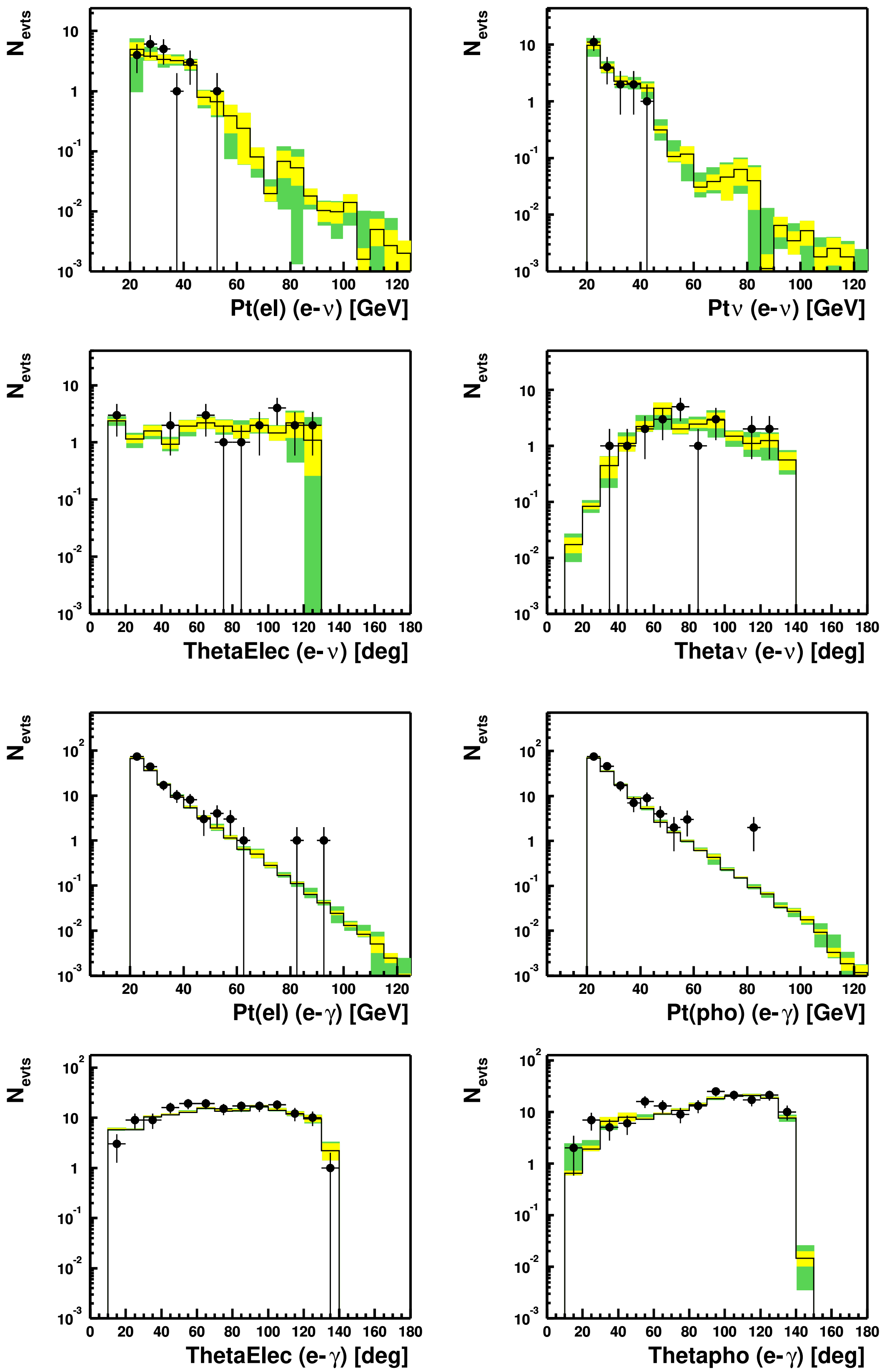
Annexe B

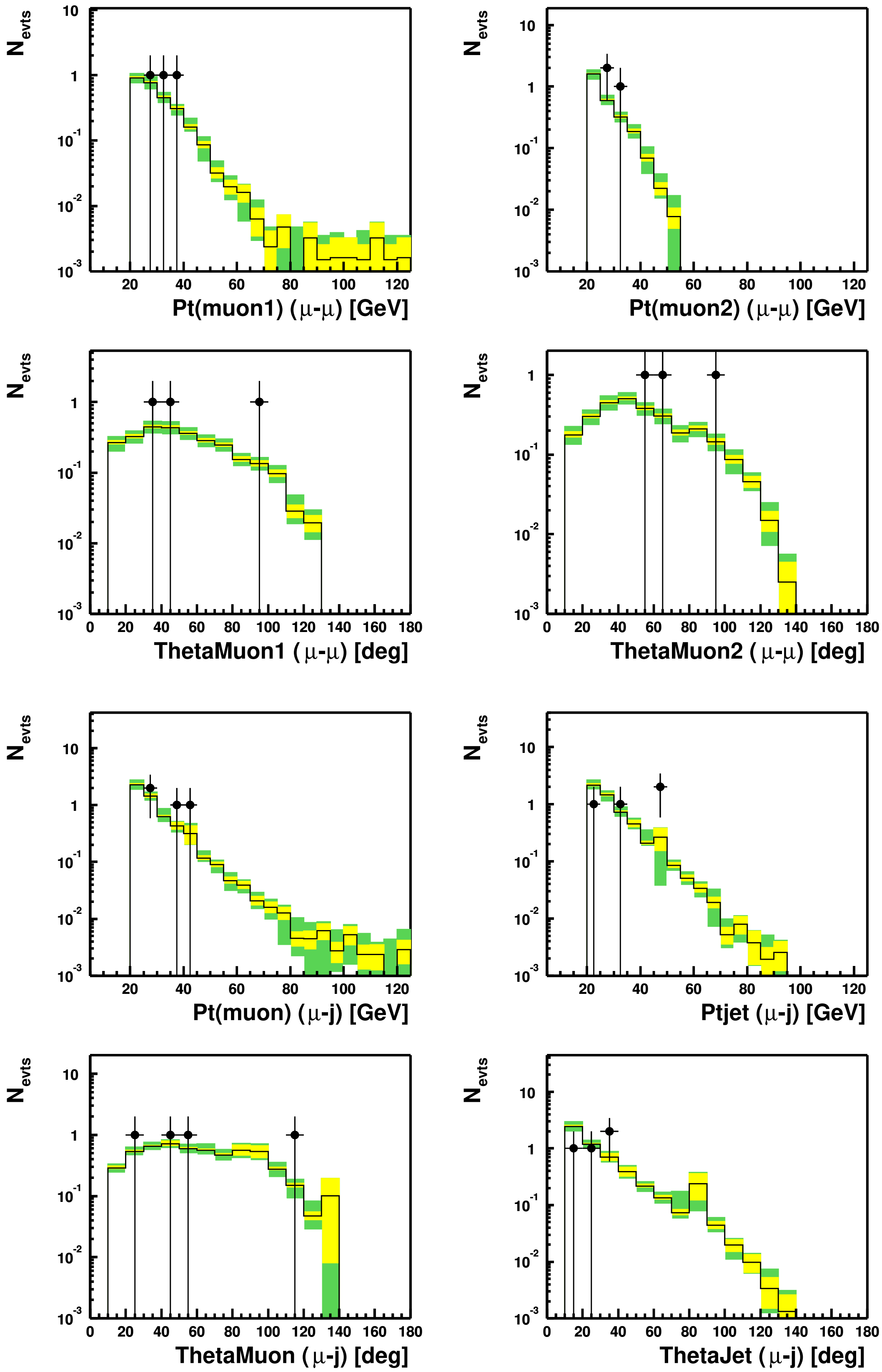
Distributions de P_T^{corps} et θ^{corps} par canal

Nous avons regroupé dans cette annexe les distributions des variables caractéristiques propres des corps présents dans chaque canal, soit le P_T et le θ de chaque corps. Nous pouvons constater ici le bon accord entre les données et les prédictions Monte-Carlo dans la plupart des canaux.

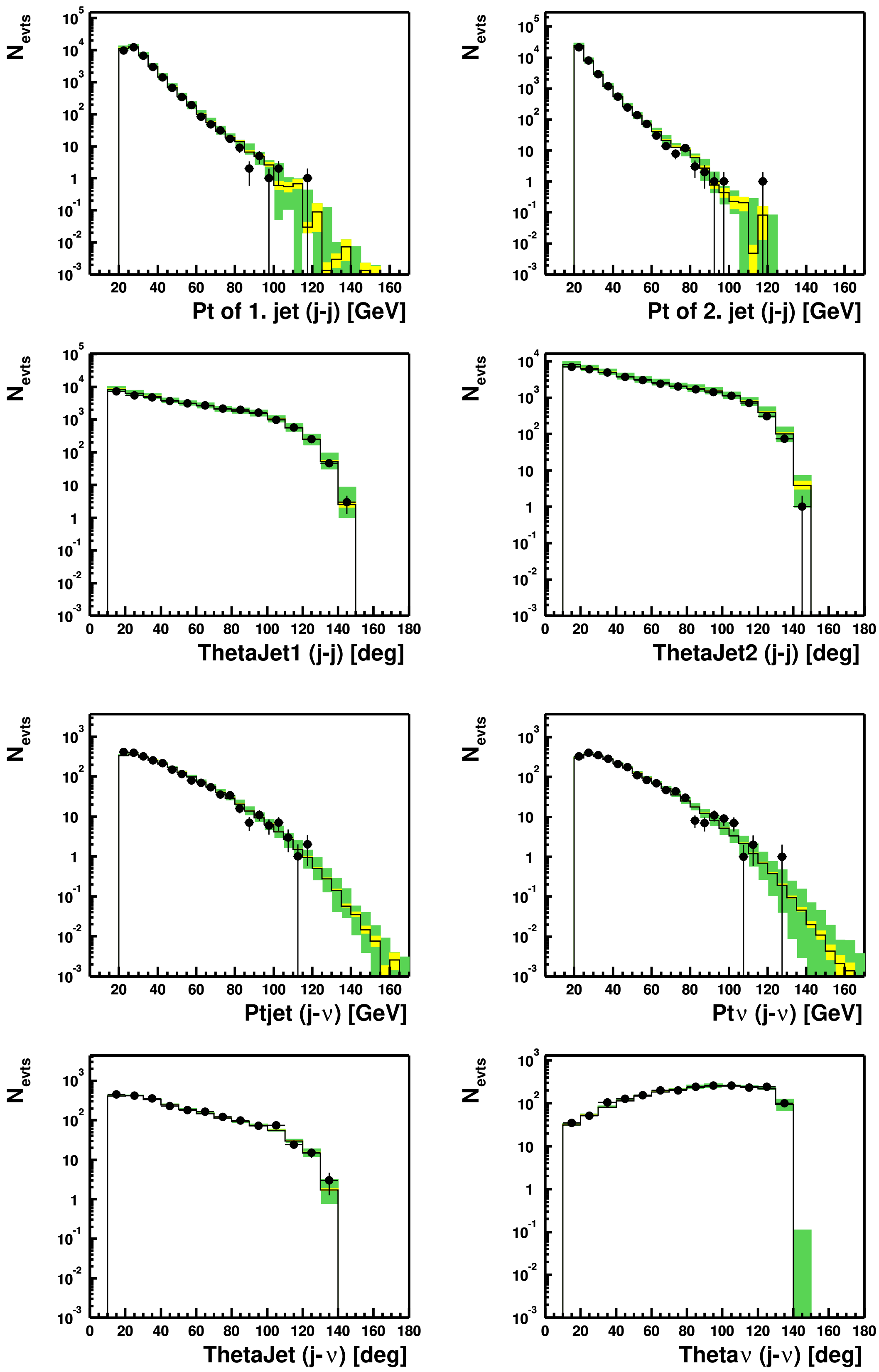



Nevts
Pt(el) (e-μ) [GeV]
Nevts
Pt(muon) (e-μ) [GeV]
Nevts
ThetaElec (e-μ) [deg]
Nevts
ThetaMuon (e-μ) [deg]
Nevts
Pt(el) (e-j) [GeV]
Nevts
Pt of 1. jet (e-j) [GeV]
Nevts
ThetaElec (e-j) [deg]
Nevts
ThetaJet1 (e-j) [deg]

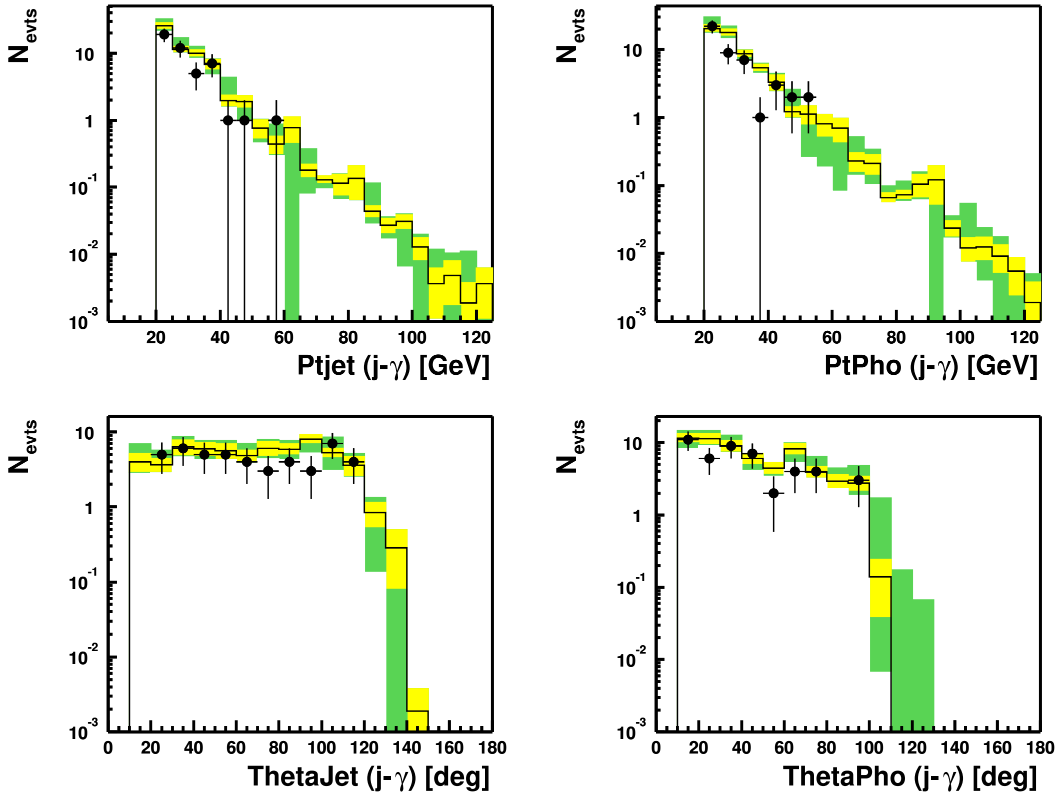

N_evts
Pt(el) (e-ν) [GeV]
Ptν (e-ν) [GeV]
ThetaElec (e-ν) [deg]
Thetaν (e-ν) [deg]
Pt(el) (e-γ) [GeV]
Pt(pho) (e-γ) [GeV]
ThetaElec (e-γ) [deg]
Thetapho (e-γ) [deg]

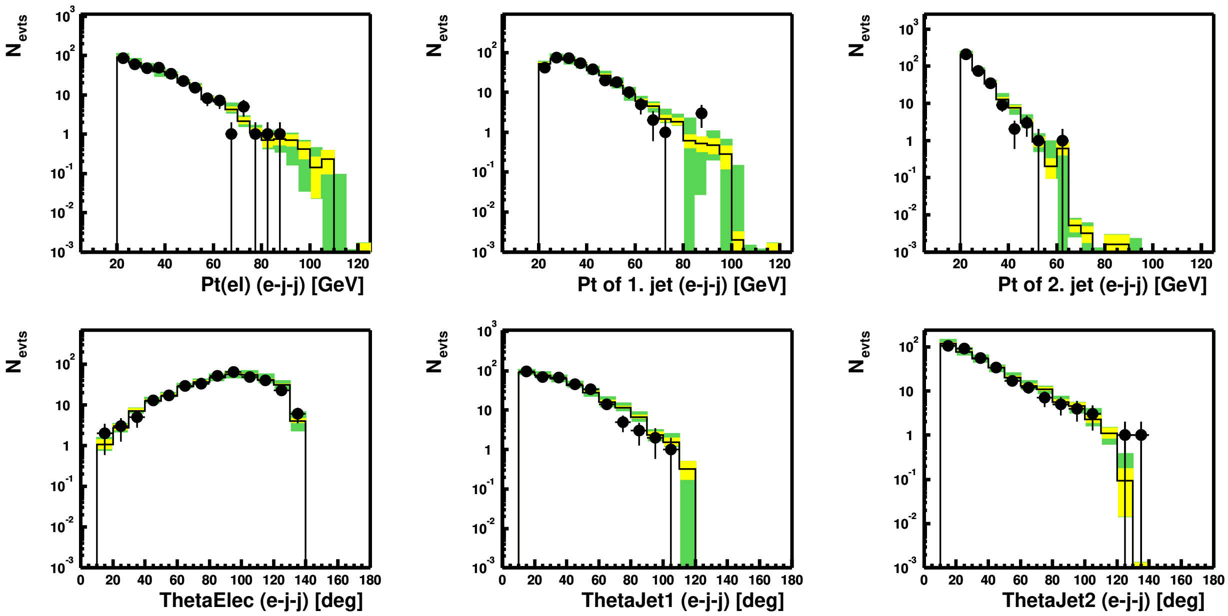


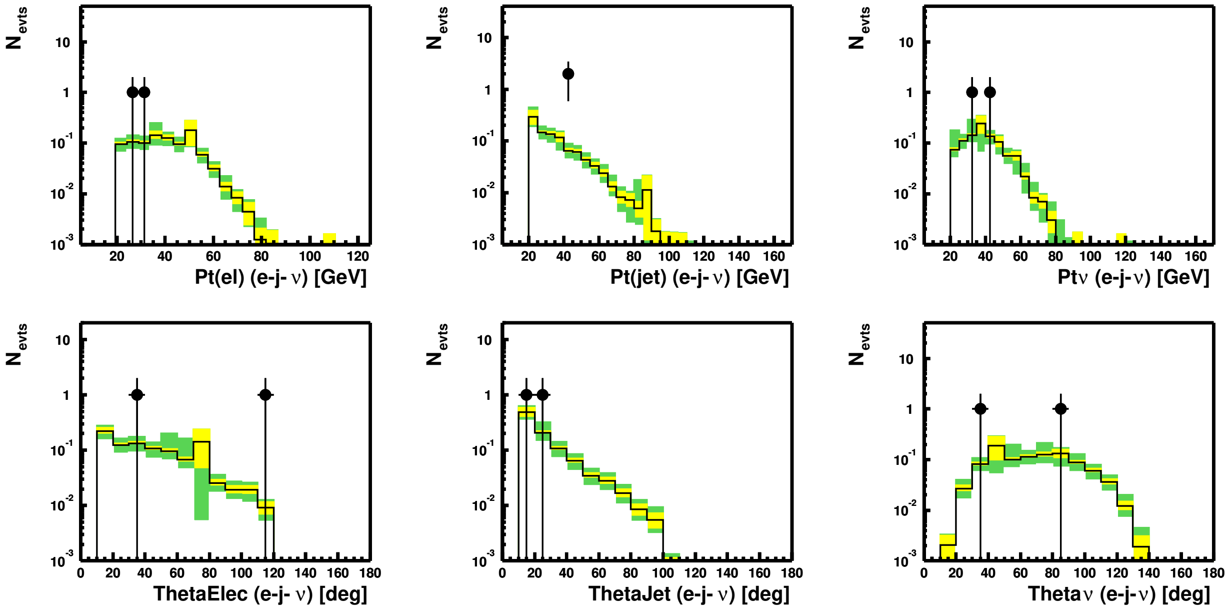
Pt(muon1) (μ-μ) [GeV]
Pt(muon2) (μ-μ) [GeV]
ThetaMuon1 (μ-μ) [deg]
ThetaMuon2 (μ-μ) [deg]
Pt(muon) (μ-j) [GeV]
Ptjet (μ-j) [GeV]
ThetaMuon (μ-j) [deg]
ThetaJet (μ-j) [deg]
N_evts

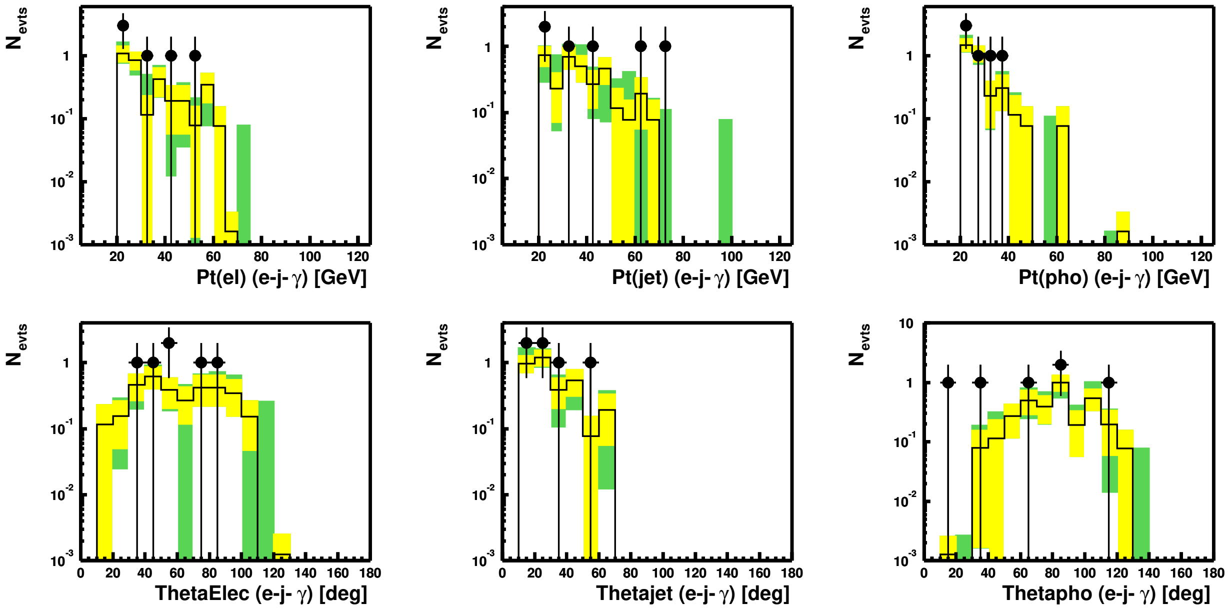


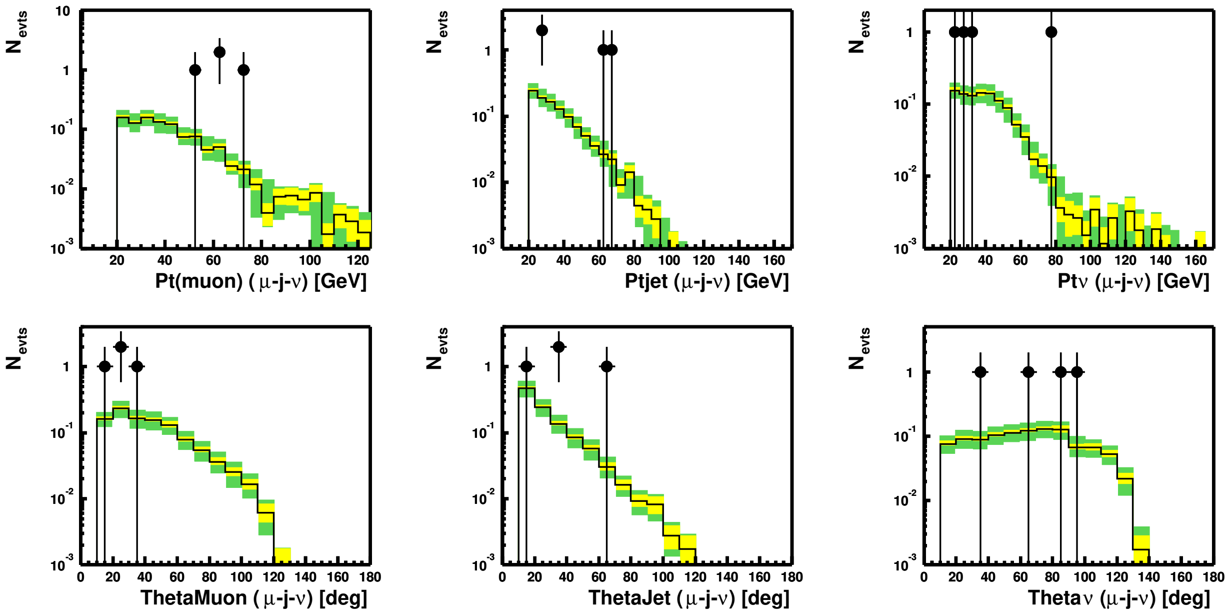
N_evts
Pt of 1. jet (j-j) [GeV]
Pt of 2. jet (j-j) [GeV]
ThetaJet1 (j-j) [deg]
ThetaJet2 (j-j) [deg]
Ptjet (j-ν) [GeV]
Ptν (j-ν) [GeV]
ThetaJet (j-ν) [deg]
Thetaν (j-ν) [deg]

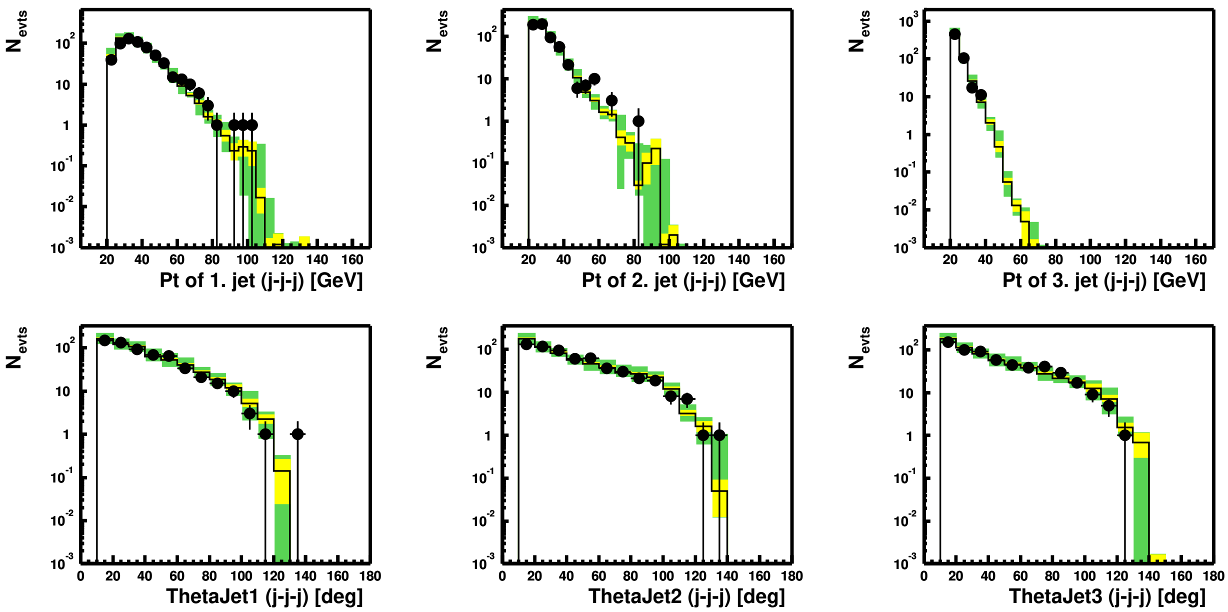

N evts
Ptjet (j-γ) [GeV]
PtPho (j-γ) [GeV]
ThetaJet (j-γ) [deg]
ThetaPho (j-γ) [deg]

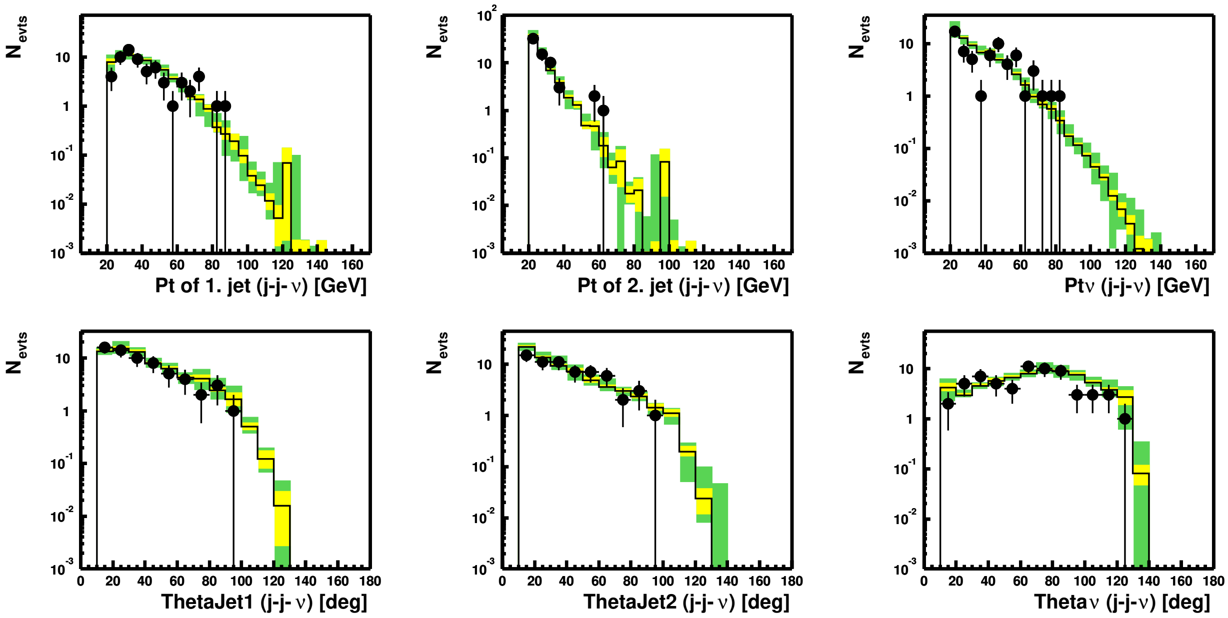

N evts
Pt(el) (e-j-j) [GeV]
Pt of 1. jet (e-j-j) [GeV]
Pt of 2. jet (e-j-j) [GeV]
ThetaElec (e-j-j) [deg]
ThetaJet1 (e-j-j) [deg]
ThetaJet2 (e-j-j) [deg]

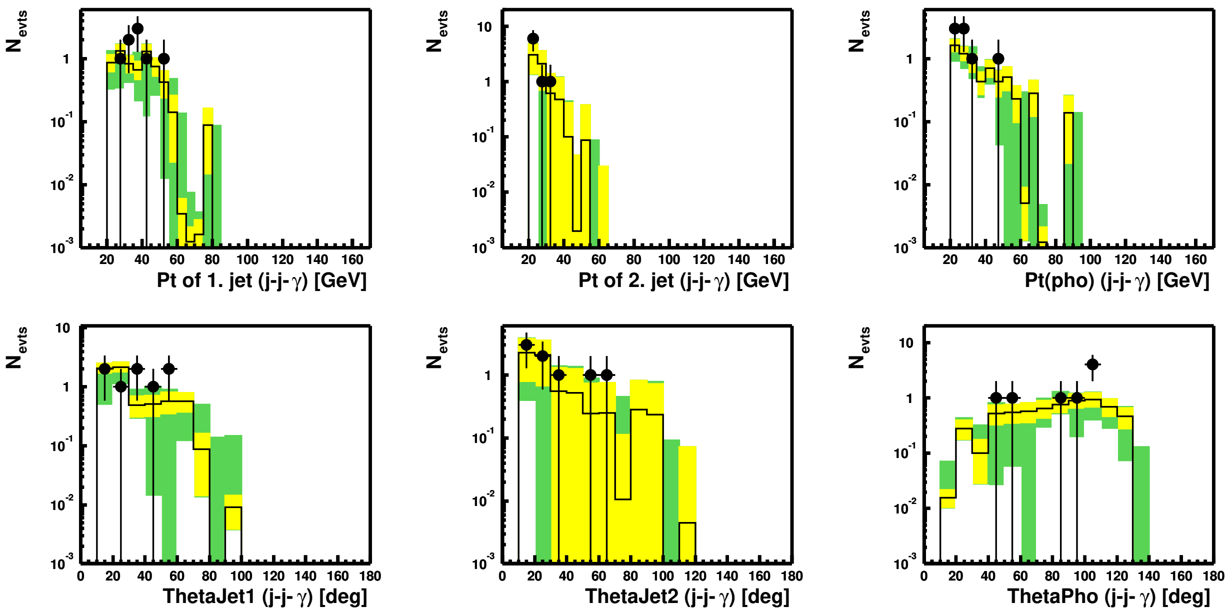

N_{evts}
Pt(el) (e-j- ν) [GeV]
Pt(jet) (e-j- ν) [GeV]
Ptν (e-j-ν) [GeV]
ThetaElec (e-j- ν) [deg]
ThetaJet (e-j- ν) [deg]
Thetaν (e-j-ν) [deg]

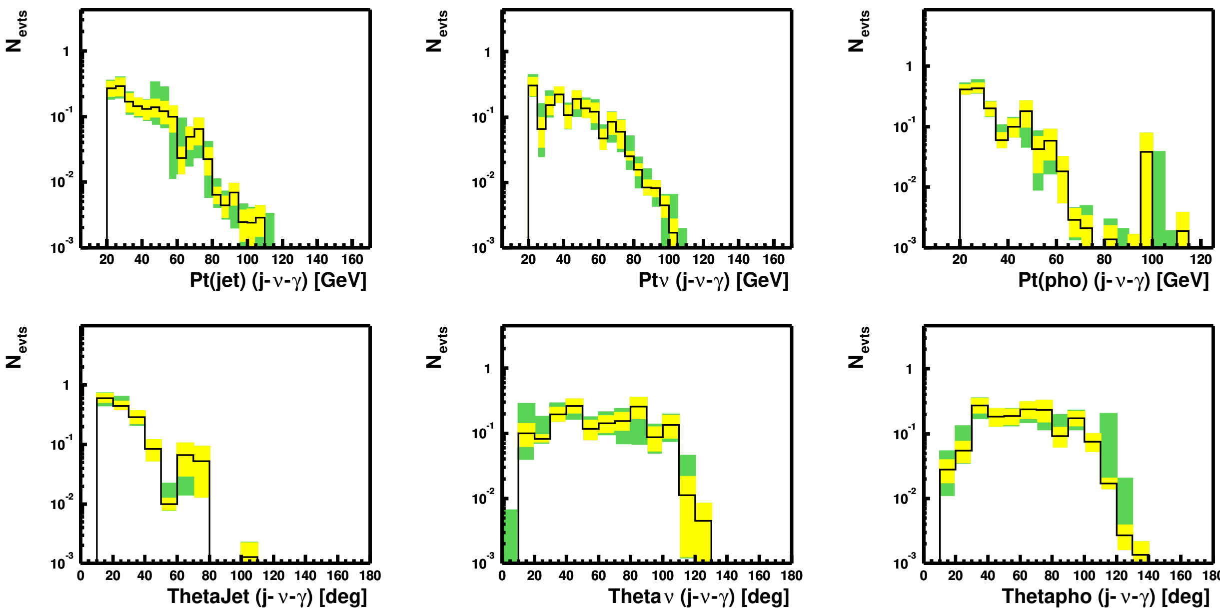

N_{evts}
Pt(el) (e-j- γ) [GeV]
Pt(jet) (e-j- γ) [GeV]
Pt(pho) (e-j- γ) [GeV]
ThetaElec (e-j- γ) [deg]
Thetajet (e-j- γ) [deg]
Thetapho (e-j- γ) [deg]

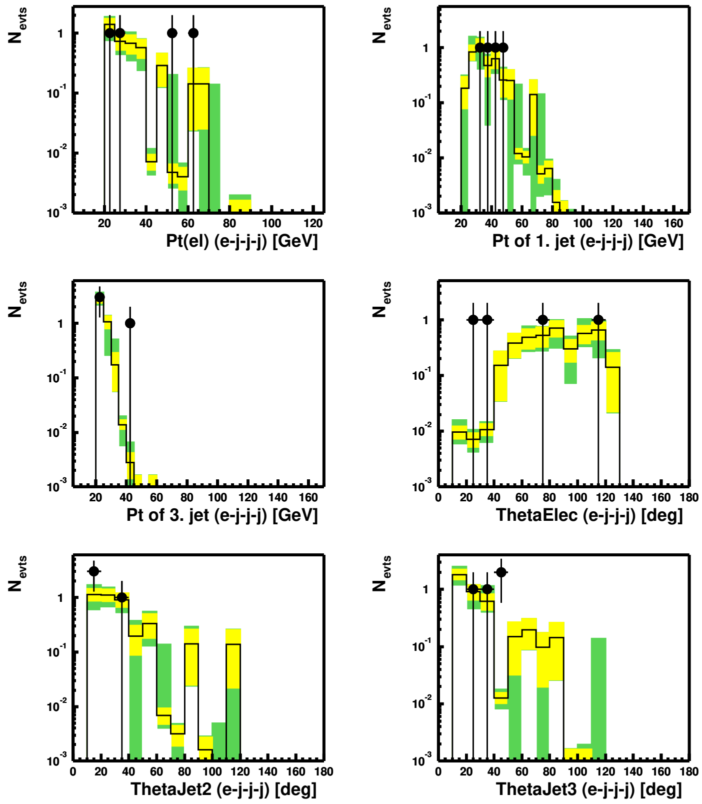

N_evts
Pt(muon) (μ-j-ν) [GeV]
Ptjet (μ-j-ν) [GeV]
Ptν (μ-j-ν) [GeV]
ThetaMuon (μ-j-ν) [deg]
ThetaJet (μ-j-ν) [deg]
Thetaν (μ-j-ν) [deg]

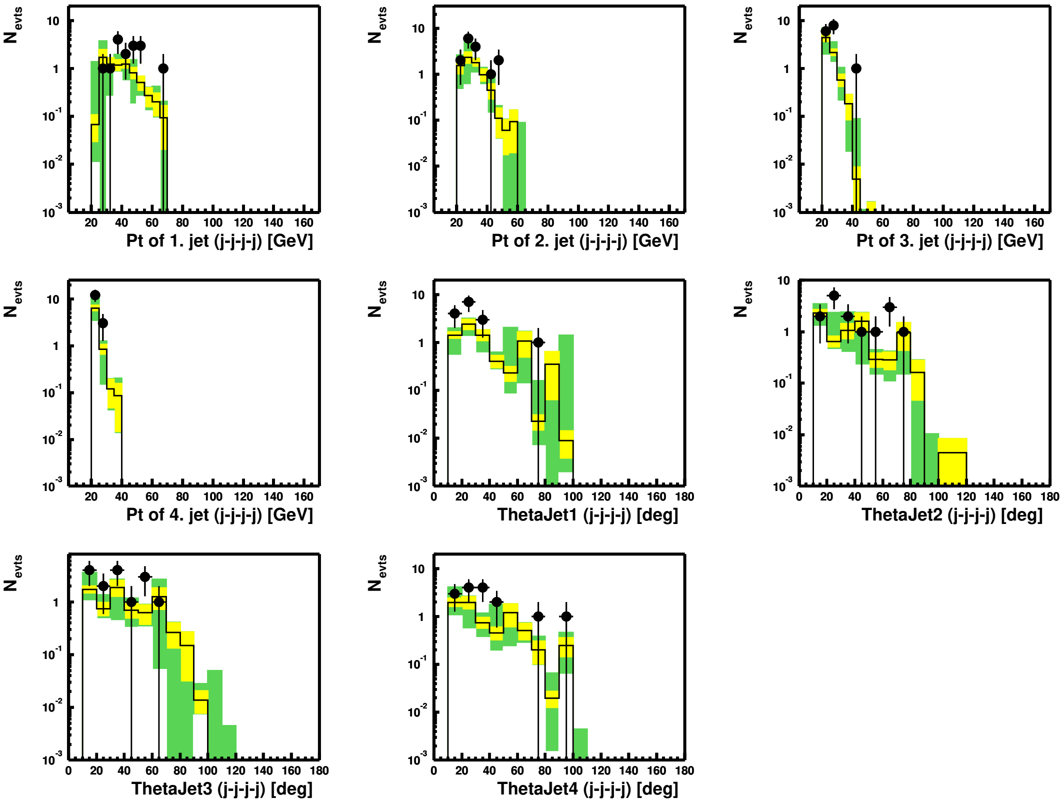

N_evts
Pt of 1. jet (j-j-j) [GeV]
Pt of 2. jet (j-j-j) [GeV]
Pt of 3. jet (j-j-j) [GeV]
ThetaJet1 (j-j-j) [deg]
ThetaJet2 (j-j-j) [deg]
ThetaJet3 (j-j-j) [deg]

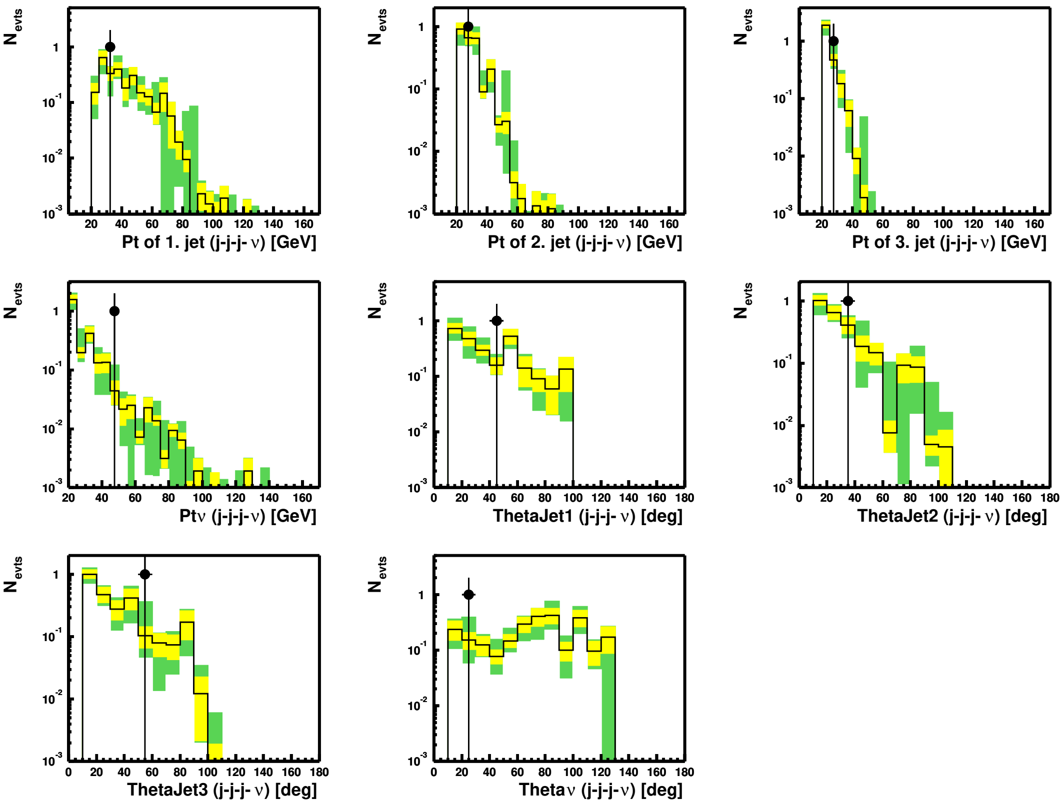

N_evts
Pt of 1. jet (j-j- ν) [GeV]
Pt of 2. jet (j-j- ν) [GeV]
Ptν (j-j-ν) [GeV]
ThetaJet1 (j-j- ν) [deg]
ThetaJet2 (j-j- ν) [deg]
Theta ν (j-j-ν) [deg]


N_evts
Pt of 1. jet (j-j-γ) [GeV]
Pt of 2. jet (j-j- γ) [GeV]
Pt(pho) (j-j- γ) [GeV]
ThetaJet1 (j-j- γ) [deg]
ThetaJet2 (j-j- γ) [deg]
ThetaPho (j-j- γ) [deg]


N_evts
Pt(jet) (j-ν-γ) [GeV]
Ptν (j-ν-γ) [GeV]
Pt(pho) (j-ν-γ) [GeV]
ThetaJet (j-ν-γ) [deg]
Theta ν (j-ν-γ) [deg]
Thetapho (j-ν-γ) [deg]


N_evts
Pt(el) (e-j-j-j) [GeV]
Pt of 1. jet (e-j-j-j) [GeV]
Pt of 2. jet (e-j-j-j) [GeV]
Pt of 3. jet (e-j-j-j) [GeV]
ThetaElec (e-j-j-j) [deg]
ThetaJet1 (e-j-j-j) [deg]
ThetaJet2 (e-j-j-j) [deg]
ThetaJet3 (e-j-j-j) [deg]


N_evts
Pt of 1. jet (j-j-j-j) [GeV]
Pt of 2. jet (j-j-j-j) [GeV]
Pt of 3. jet (j-j-j-j) [GeV]
Pt of 4. jet (j-j-j-j) [GeV]
ThetaJet1 (j-j-j-j) [deg]
ThetaJet2 (j-j-j-j) [deg]
ThetaJet3 (j-j-j-j) [deg]
ThetaJet4 (j-j-j-j) [deg]


N_evts
Pt of 1. jet (j-j-j- ν) [GeV]
Pt of 2. jet (j-j-j- ν) [GeV]
Pt of 3. jet (j-j-j- ν) [GeV]
Ptν (j-j-j-ν) [GeV]
ThetaJet1 (j-j-j- ν) [deg]
ThetaJet2 (j-j-j- ν) [deg]
ThetaJet3 (j-j-j- ν) [deg]
Thetaν (j-j-j-ν) [deg]

Annexe C

Distributions en masse invariante et régions de plus grand intérêt sélectionnées

Nous montrons dans cette annexe les distributions en masse invariantes dans les différents canaux en comparant les données aux simulations Monte-Carlo. Nous avons également reporté la région la plus intéressante sélectionnée par la méthode d'interprétation statistique.

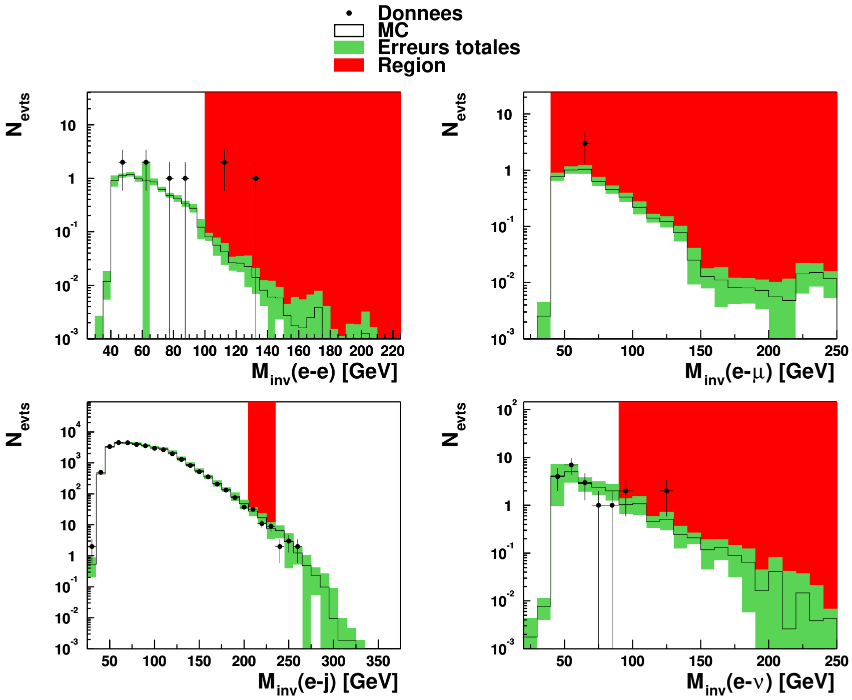


FIG. C.1 – Distributions de la masse invariante M_{inv} avec les régions de plus grand intérêt.

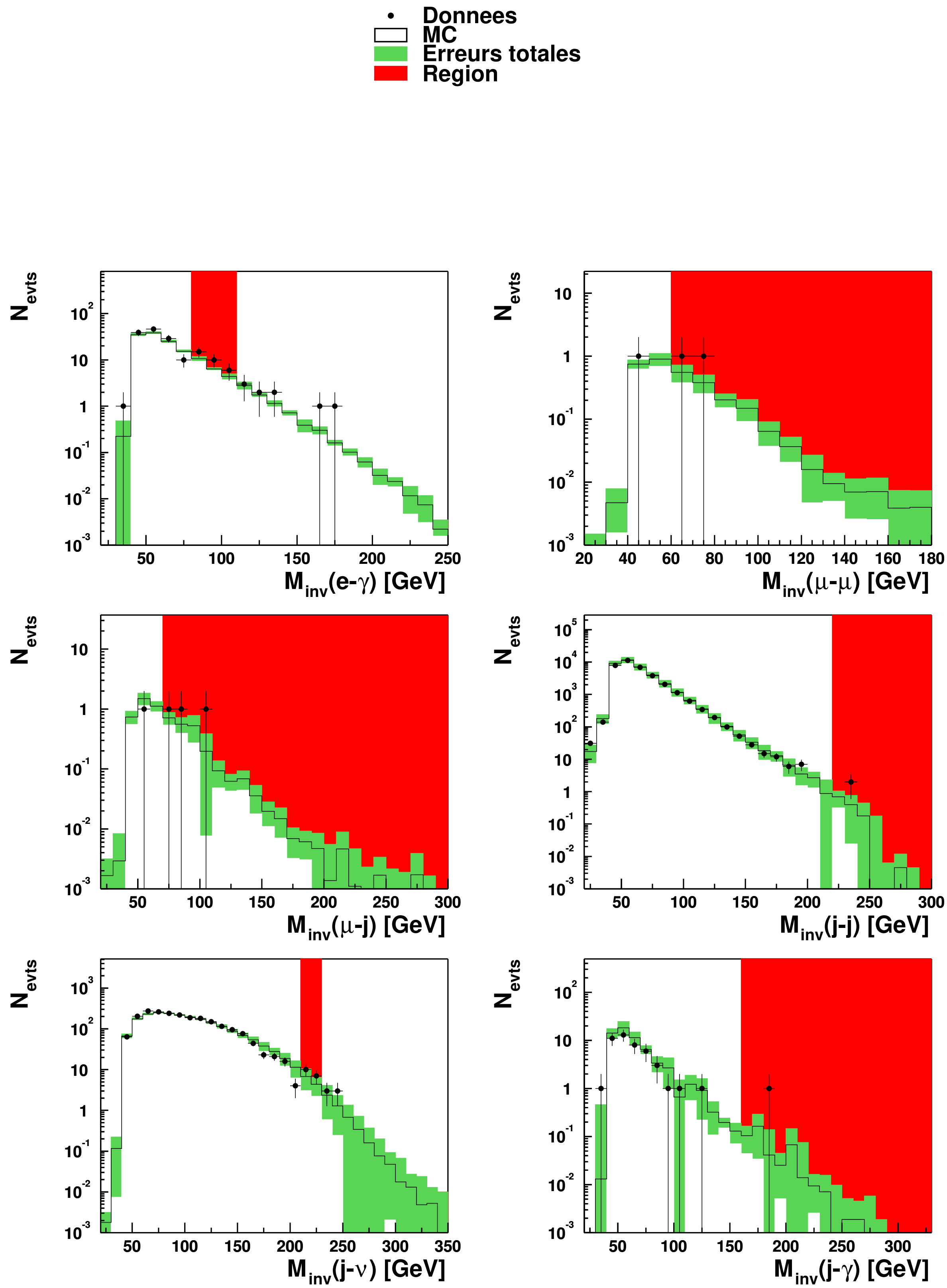


FIG. C.2 – Distributions de la masse invariante M_{inv} avec les régions de plus grand intérêt.

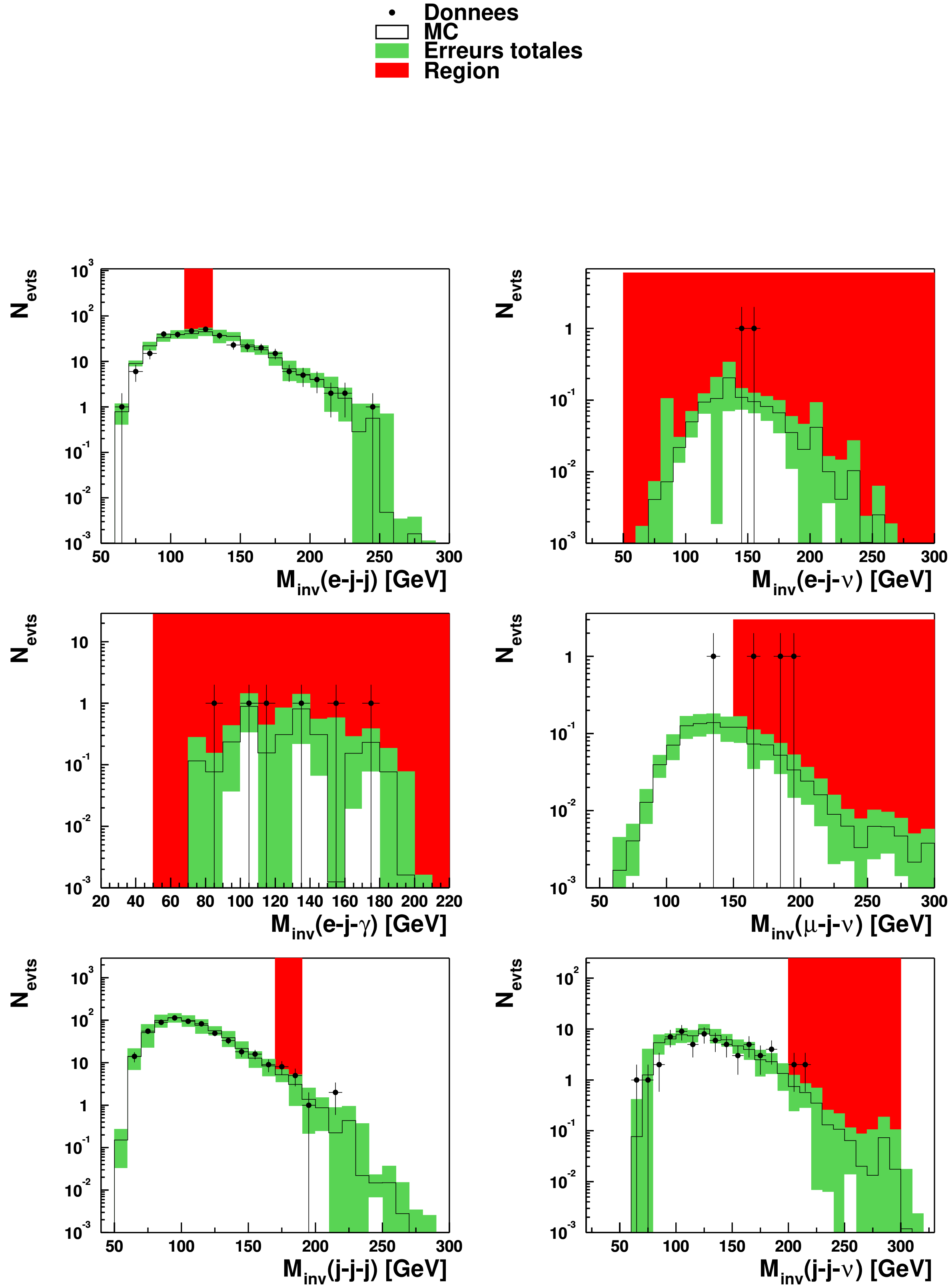


FIG. C.3 – Distributions de la masse invariante M_{inv} avec les régions de plus grand intérêt.

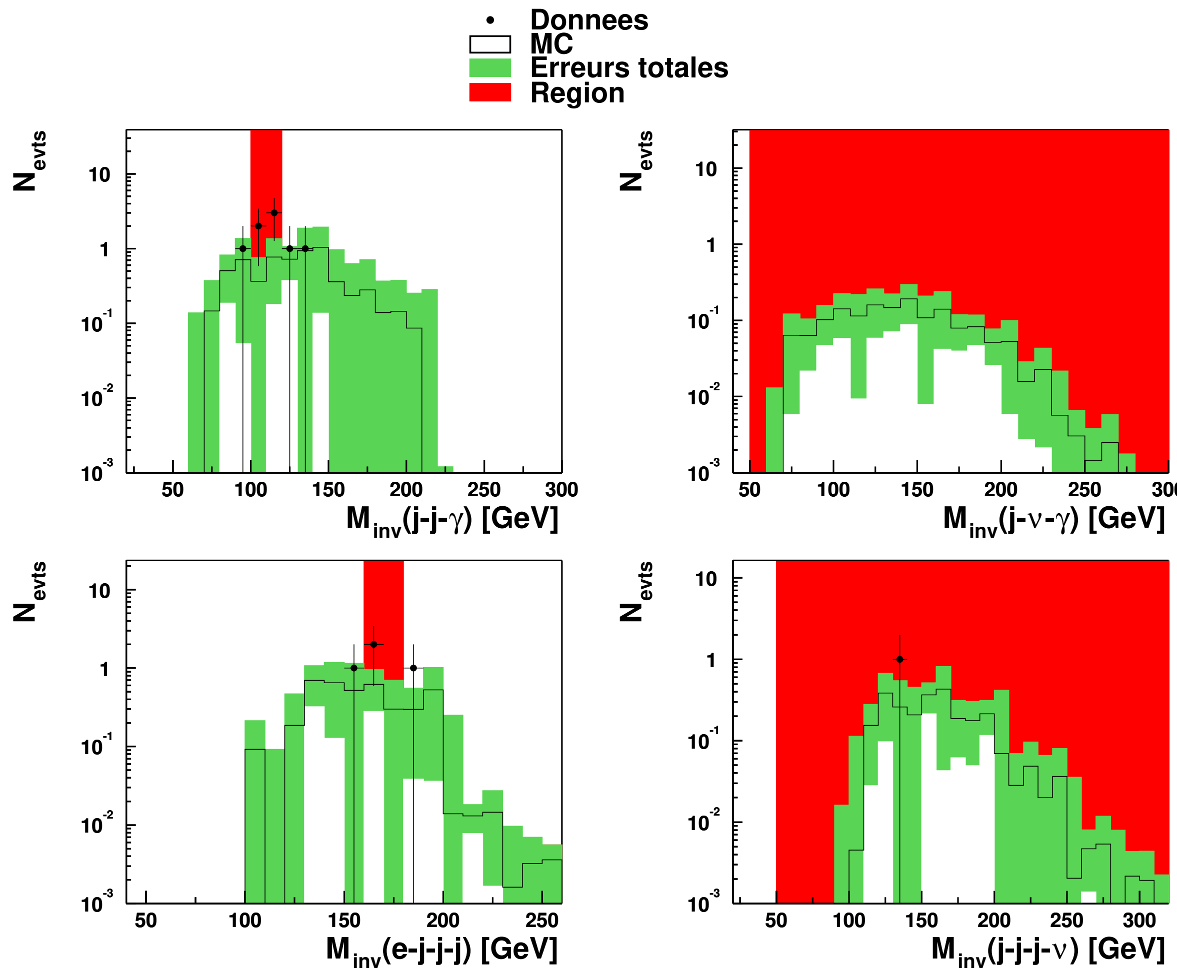


FIG. C.4 – Distributions de la masse invariante M_{inv} avec les régions de plus grand intérêt.

Recherche de déviations au Modèle Standard dans les processus de grande énergie transverse sur le collisionneur électron-proton HERA

Cette thèse présente une étude exhaustive des différentes topologies d'événements à grande énergie transverse observés sur le collisionneur électron-proton HERA. L'ensemble des données accumulées par l'expérience H1 entre 1994 et 2000 ont été utilisées, ce qui représente une luminosité intégrée de 118,3 pb^{-1}.

Nous avons tout d'abord participé au développement d'un nouvel environnement d'analyse orientée objet, en nous concentrant particulièrement sur la reconstruction des hadrons et des jets hadroniques. Cet environnement, qui est devenu l'outil d'analyse standard de H1 de la phase II de HERA, est mis à l'épreuve dans le cadre de notre sujet.

Les événements étudiés regroupent 24 topologies différentes pour lesquelles les spectres d'énergie transverse sont systématiquement étudiés. Un bon accord général au Modèle Standard est observé, mis à part dans quelques canaux à leptons isolés où plusieurs événements atypiques ont déjà été rapportés. La comparaison au Modèle Standard est quantifiée avec l'aide d'une méthode statistique spécifique développée pour l'occasion.

Mots-clés : Modèle Standard, HERA, H1, grande énergie transverse, générique, modèle indépendant, recherche, programmation orientée objet.

Search for deviations from the Standard Model in high transverse energy processes at the electron-proton collider HERA

In this thesis, an exhaustive analysis of various topologies in high transverse energy events observed at the electron-proton collider HERA is presented. The whole data set accumulated by the H1 experiment between 1994 and 2000, corresponding to an integrated luminosity of 118,3 pb^{-1}, has been used.

First of all, we contributed to the development of a new object oriented analysis framework, with main emphasis on the implementation of hadron and hadronic jet reconstruction algorithms. This environment, now a standard tool for the HERA II phase, has been used and evaluated within the scope of our subject.

The studied events gather 24 topologies for which transverse momenta distributions are systematically studied. A good overall agreement with the Standard Model is observed, except in some isolated lepton channels, where atypical events have already been reported. The comparison to the Standard Model is quantified by using a statistical method developped for this occasion.

Keywords : Standard Model, HERA, H1, high transverse energy, generic, model independent, search, object oriented programming.

Discipline : Physique des particules

Laboratoire :

Centre de Physique des Particules de Marseille
163, avenue de Luminy Case 907
13288 Marseille cedex 09
France