LPNHE-2004-
30 août 2004

Rapport de Stage de Licence

Christophe Becavin

Université Paris VII

Magistère de Physique

Identification d'électrons et réjection de jets dans Atlas

LPNHE-Paris, Universités Paris 6-7, et IN2P3-CNRS

Abstract

Découvert au Fermilab en 1995, le top est le quark du modèle standard de la physique des particules le moins bien connu. Sa masse élevée le rend difficilement accessible par les accélérateurs de particules actuels. Le LHC (Large Hadron Collider) est un collisionneur proton-proton en construction actuellement au CERN près de Genève. Les collisions pouront produire des paires $t\bar{t}$ (quark top et antitop). Le groupe ATLAS du LPHNE prépare des analyses visant à le rechercher lorsque la paire $t\bar{t}$ produit deux leptons dans l'état final, dont au moins un électron. Le bruit de fond dans ces analyses provient en partie de jets (ensemble de particules formant un faisceau collimé), qui vont être pris pour des électrons dans le détecteur. Mon stage a porté sur la mise au point de critères de sélection identifiant efficacement les électrons et rejetant les jets.

ATLAS
LPNHE Paris 6-7

Table des matières

1 Introduction

1.1 La physique du Top

1.1.1 Le modèle standard

La physique des particules s'attache à développer notre connaissance des quarks (qui sont notamment les constituants des neutrons et des protons), des leptons (comme l'électron) et des interactions qui les gouvernent. Ces particules sont considérées, à ce jour, comme les constituants élémentaires de la matière. Les avancées, étroitement liées, de la théorie et des observations expérimentales, ont conduit à ce que l'on appelle le "Modèle Standard". A l'aide de douze constituants élémentaires (cf tableau 1) ce modèle permet de décrire les mécanismes internes de la matière, notamment les différentes interactions fondamentales, à l'exception de la gravitation. Ces interactions sont :

- l'interaction électrofaible qui résulte de l'unification de l'interaction électromagnétique et de l'interaction faible, véhiculées respectivement par le photon et les bosons vecteurs $W^{\pm}$ et Z ;
- l'interaction forte, véhiculée par les gluons, qui est responsable de la cohésion des noyaux.

Fermions	Quarks	$\begin{pmatrix}\text{haut } u \\ \text{bas } d\end{pmatrix}$	$\begin{pmatrix}\text{charmé } c \\ \text{étrange } s\end{pmatrix}$	$\begin{pmatrix}\text{top } t \\ \text{beau } b\end{pmatrix}$
	Leptons	$\begin{pmatrix}\text{neutrino } \nu_e \\ \text{électron } e^-\end{pmatrix}$	$\begin{pmatrix}\text{neutrino } \nu_\mu \\ \text{mu } \mu\end{pmatrix}$	$\begin{pmatrix}\text{neutrino } \nu_\tau \\ \text{tau } \tau\end{pmatrix}$
Bosons	Bosons de jauge	photon γ - interaction électromagnétique		
		bosons vecteurs $W^{\pm}, Z^0$ - interaction faible		
		gluons g - interaction forte		
	boson de Higgs H ?			

TAB. 1 – *Les particules du modèle standard.*

Ce modèle n'a pas été mis en défaut par les observations expérimentales. Cependant, nous savons qu'il ne constitue pas, à lui seul, la théorie ultime de la physique des particules puisque, dans sa forme actuelle, certaines questions importantes restent sans réponse :

- comment pouvons-nous unifier l'interaction électrofaible avec les interactions forte et gravitationelle ?
- comment expliquer la disparition de l'antimatière alors qu'à sa naissance l'univers contenait autant de matière que d'antimatière ? Ce mystère pourrait trouver sa solution grâce à l'étude des violations des symétries discrètes CP (C pour conjugaison de charge, P pour parité) ;
- pourquoi - et comment - les particules élémentaires ont-elles une masse ? Et pourquoi ces masses sont-elles si différentes (de quelques eV à près de 170 GeV pour le quark top) ? Le mécanisme de Higgs par brisure de la symétrie électrofaible pourrait répondre à ces deux questions. Une conséquence remarquable est qu'il prédit l'existence d'une particule scalaire neutre : le boson de Higgs. Cette particule est à ce jour le dernier élément manquant du Modèle Standard.

L'accélérateur de particules LHC (Large Hadron Collider), dont la construction au CERN devrait être finie pour 2007, permettra d'apporter des réponses à ces différentes questions. Ce sera un collisioneur proton-proton qui fournira une énergie de 14 TeV dans le centre de masse. On prévoit une "basse" luminosité de $2.10^{33} cm^{-2}s^{-1}$ lors des premières années et une "haute" luminosité de $10^{34} cm^{-2}s^{-1}$ après. Le LHC fournira une énergie suffisante pour sonder la matière dans le secteur du Higgs, et permettre éventuellement de faire apparaitre de nouveaux phénomènes situés "au delà" du modèle standard...

1.1.2 Le Quark Top

Le top, le dernier des six quarks du modèle standard, à été découvert en 1995 au Fermilab, près de Chicago. Le faible nombre d'événements observés n'a pas permis d'explorer finement ses propriétés. Sa masse élevée ($\sim 175 GeV/c^2$) par rapport aux autres particules lui donne une place à part dans le Modèle Standard. Une meilleure étude de celui-ci permettra d'affiner les connaissances sur le modèle standard, mais aussi d'apporter des contraintes significatives sur la masse du Higgs.

De par sa très haute luminosité et sa grande énergie, le LHC sera une "usine à quarks top", avec environ 60000 paires $t\bar{t}$ produites par jour. Ce mode de production lors des collisions proton-proton sera dominant, la création d'un seul top contribuant pour environ 30%. La figure 1 montre la simulation d'un événement $t\bar{t}$ dans le détecteur ATLAS, que nous décrirons par la suite.

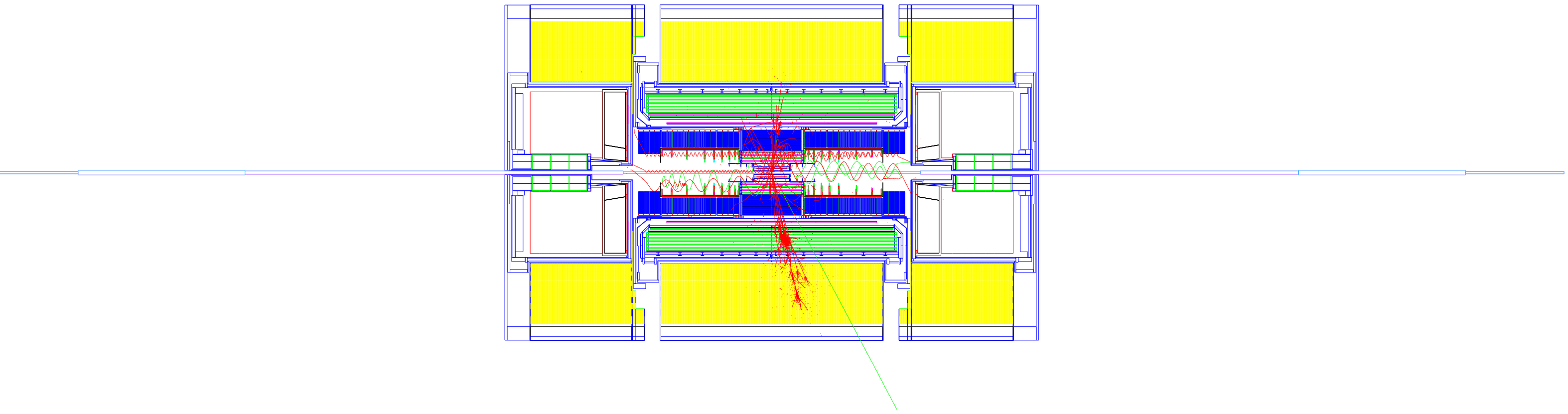

FIG. 1 – *Simulation d'un événement $t\bar{t}$ dans le détecteur interne et le calorimètre. Dans la partie interne (trajectographe) du détecteur on voit un grand nombre de traces. Dans le calorimètre électromagnétique (partie hachurée interne) et hadronique (partie hachurée externe) on voit deux jets se propager.*

Du fait de sa masse élevée, le top a une durée de vie très courte. Celui-ci va se désintégrer pour former un état intermédiaire où les particules vont initier des jets (on dit qu'elles s'hadronisent) ou des particules individuelles. Le top ne pourra donc être observé que via ses produits de désintégration.

Il existe trois désintégrations possibles du top, dont une domine : la désintégration $t \to Wb$ avec une probabilité de se produire de près de 99,9% (*c.f* Fig. 2). Le canal de production du top choisi est le $t\bar{t}$, on a donc production de deux bosons W^+ et W^-, et de deux quarks b qui en s'hadronisant vont donner deux jets. Pour les désintégrations des W, il y a trois canaux principaux de désintégration (*c.f* [ATLAS TDR Phys Vol 2, 1999]) :

- le canal à zéro lepton ($\simeq$ 44% des événements $t\bar{t}$) où les 2 W se désintègrent en quarks, qui vont former deux jets supplémentaires ;
- le canal à un lepton ($\simeq$ 30% des événements $t\bar{t}$, sans compter les τ) où l'un des W se désintègre en quark, l'autre en lepton ;

- le canal à deux leptons ($\simeq 5\%$ des événements $\mathrm{t\bar{t}}$, sans compter les τ) où les 2 W se désintègrent en leptons.

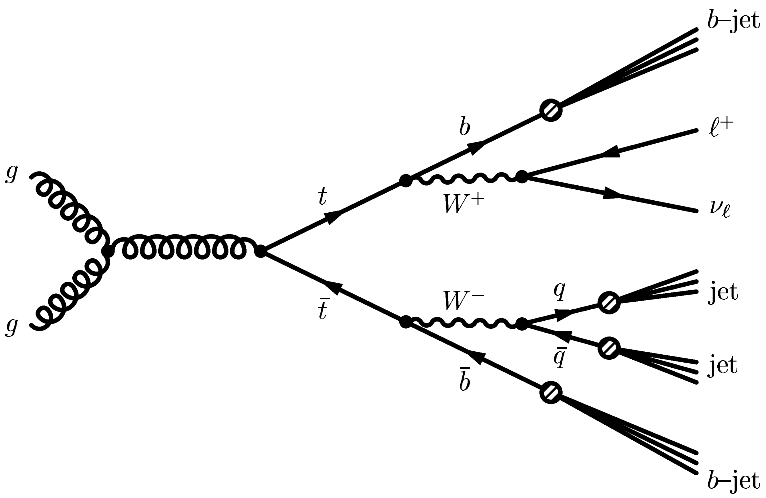

FIG. 2 – *Diagramme de Feynman du canal préférentiel de détection des paires $t\bar{t}$ dans ATLAS, avec des jets et des leptons présent dans l'état final.*

Lors de mon stage, je me suis penché sur ce dernier canal di-leptons avec au moins un électron présent dans l'état final. Le bilan final de la désintégrations de la paire $\mathrm{t\bar{t}}$ choisit est donc le suivant :

$$
\begin{aligned}
t\bar{t} &\rightarrow b\bar{b}W^+W^- \rightarrow e\nu\mu\nu + 2 \text{ jets ou} \\
t\bar{t} &\rightarrow b\bar{b}W^+W^- \rightarrow e\nu e\nu + 2 \text{ jets}
\end{aligned}
$$

Pour étudier le top il faut être capable de retrouver ces produits de désintégrations, donc être capable d'identifier les électrons. Or le problème d'un collisionneur comme le LHC est qu'il va produire une multitude de désintégrations, dont les produits auront la même signature expérimentale que les événements top recherchés. Par exemple on pourra avoir des jets qui seront identifiés comme étant des électrons. Tout celà formera ce qu'on appelle le bruit de fond instrumental. Le canal en di-leptons choisi est donc intéressant car il a une signature très claire et possède pour bruits de fond physique[1] des processus rares. De plus, la présence de deux leptons dans l'état final réduit le bruit de fond instrumental, car pour qu'un événement multijet simule un événement $\mathrm{t\bar{t}}$ à deux leptons, il faut que deux jets soient identifiés comme des leptons, ce qui est très peu probable.

Ainsi finalement pour l'étude du top, il faut pouvoir identifier efficacement les électrons tout en rejetant les jets, ce fut mon travail lors de mon stage. Cette problématique est importante, puisque pour atteindre des performances optimales dans la plupart des analyses de physique une réjection d'un facteur $\sim 10^5$ doit être atteint (*c.f* [ATLAS TDR Phys Vol 1, 1999]).

[1]Un bruit de fond physique est un processus qui possède les mêmes produits de désintégration.

1.2 Le détecteur Atlas

1.2.1 Présentation du détecteur

Le détecteur ATLAS est le plus gros des quatre détecteurs qui seront installés sur le faisceau du LHC. Il se présente sous la forme d'un cylindre d'un diamètre de 22 mètres, et d'une longueur de 44 mètres, pour un poids total de 7000 tonnes (cf. figure 3). On distingue deux parties, la partie tonneau qui est le corps du détecteur, et les deux bouchons situés à chaque extremité qui ferment le détecteur. On peut trouver une description complète dans [ATLAS TDR Phys Vol 1, 1999] et les références qui y sont incluses.

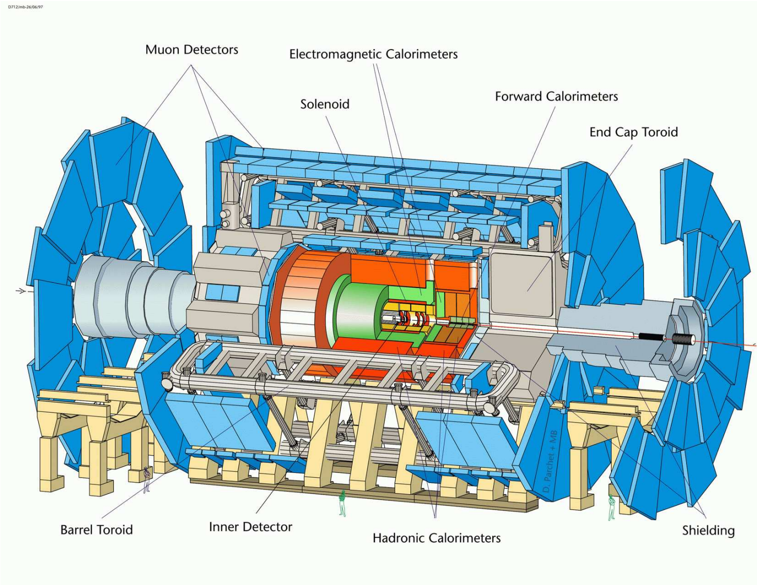

FIG. 3 – *Représentation du détecteur Atlas.*

En suivant la trajectoire d'une particule, depuis le point d'interaction situé au centre du détecteur, on trouve successivement :

- Le détecteur interne ou trajectographe (1,15 m de rayon, 7 m de long) : Les parties les plus proches du point d'interaction, cruciales pour la reconstruction des vertex (point de production de la particule), sont constituées de trois couches de pixels de silicium (généralement appelé simplement détecteur à pixels) et de détecteur à micropistes (SCT pour "semi-conductor tracking"), ayant une technologie similaire mais avec une granularité moins grande. Les couches les plus externes ont été réalisées à l'aide d'un détecteur à tubes à pailles (TRT, Transition Radiation Tracker) en raison du coût élevé de la technologie des détecteurs à silicium. Chaque trace chargée est échantillonnée sur 36 points par ce détecteur. De plus, et surtout en ce qui concerne ce stage, l'emploi de xénon en tant que milieu actif permet de détecter les photons de rayonnement de transition (i.e des photons X émis lors du passage d'électrons dans les radiateurs situés entre les pailles). Celà permet d'identifier les électrons et de rejeter les traces issues de hadrons, cette particularité sera primordiale dans la mise au point de l'algorithme d'identification des électrons.
- l'aimant solénoïdal : Il instaure dans le détecteur interne un champ magnétique qui courbe les trajectoires des particules et permet ainsi d'obtenir des information telles que le signe de la charge et l'impulsion.
- le calorimètre électromagnétique ou ECAL (2,25 m de rayon, 14 m de long) : il sert à

la mesure du dépôt en énergie des particules neutres et chargées formant des gerbes électromagnétiques (les photons et les électrons) ;
- le calorimètre hadronique ou HCAL (4,23 m de rayon, 12,20 m de long) : il sert à la mesure du dépôt en énergie des particules neutres et chargées formant des gerbes hadroniques (comme les pions, les kaons, etc...) ;
- le détecteur à muon (10 m de rayon, 44 m de long) : seuls les muons et les neutrinos arrivent à cette distance du point de collision. Les muons interagissent alors avec le détecteur tandis que les neutrinos s'en échappent.

Le détecteur étant cylindrique, on utilisera les coordonnés cylindriques (η, ϕ, z), où $\eta = -\ln\tan\theta/2$ est la pseudo-rapidité qui est une variable provenant de calcul de relativité restreinte. On l'utilise à la place de θ, car l'occupation du détecteur ($\sim$ nombre de particules par unité de surface) est alors plus uniforme. Dans ce qui suit, nous nous intéresserons plus particulièrement au sous-détecteur qui nous concerne directement, à savoir le calorimètre électromagnétique.

1.2.2 Le calorimètre électromagnétique

Il s'agit du détecteur sur lequel repose l'essentiel du travail du groupe Atlas du LPHNE. Le rôle du calorimètre électromagnétique est de mesurer avec précision l'énergie et la position des électrons, et des photons. On comprend donc qu'il va jouer un rôle très important pour la mise en place d'un algorithme d'identification des électrons. Comme ce détecteur fait partie de ATLAS on distingue là aussi deux parties, la partie centrale (tonneau) couvrant le domaine de pseudo rapidité $|\eta| < 1.45$, et les deux “bouchons” couvrant les domaines $1.45 < |\eta| < 3.2$.

Son fonctionnement repose sur deux éléments :
- un milieu absorbeur formé de plaques de plomb pliées en forme d' "accordéons" ($\sim$ 1.5 mm d´épaisseur). Lorsque des électrons, positrons ou photons traversent le plomb, ils produisent une gerbe de particules encore très énergétiques. Celles-ci produisent de nouvelles gerbes en cascades jusqu'à parvenir à des particules de trop basse énergie pour intéragir avec la matière.
- un milieu actif, l'argon liquide ($\sim$ 4 mm d'épaisseur). Les particules de basse énergie de la gerbe traversent l'argon liquide, elles intéragissent avec les atomes par ionisation (pour les particules chargées) et par effet Compton ou par effet photoélectrique (pour les photons). Les électrons issus de cette ionisation sont dirigés par le champ électrique sur les électrodes. L'intensité mesurée est alors proportionelle à l'énergie initiale.

Le calorimètre électromagnétique est divisé en trois compartiments (cf. Fig 4), utilisés par l'algorithme d'identification des électrons :
- le premier compartiment possède une granularité très fine (environ dix fois plus fine en pseudo-rapidité que les autres compartiments) de manière à mesurer très précisément l'angle et la position du début du développement de la gerbe.
- le deuxième compartiment, le plus grand, va contenir l'essentiel de la gerbe et va donc permettre la mesure de l'essentiel du dépôt en énergie.
- le troisième compartiment va permettre la mesure des gerbes de très grandes énergies qui pourraient ne pas être entièrement contenues dans les deux premiers compartiments. De plus, ce compartiment sera essentiel pour la séparation entre les gerbes issues d'électrons de basse énergie qui vont déposer peu d'énergie dans ce compartiment et les hadrons qui eux vont y déposer une part substantielle de leur énergie.

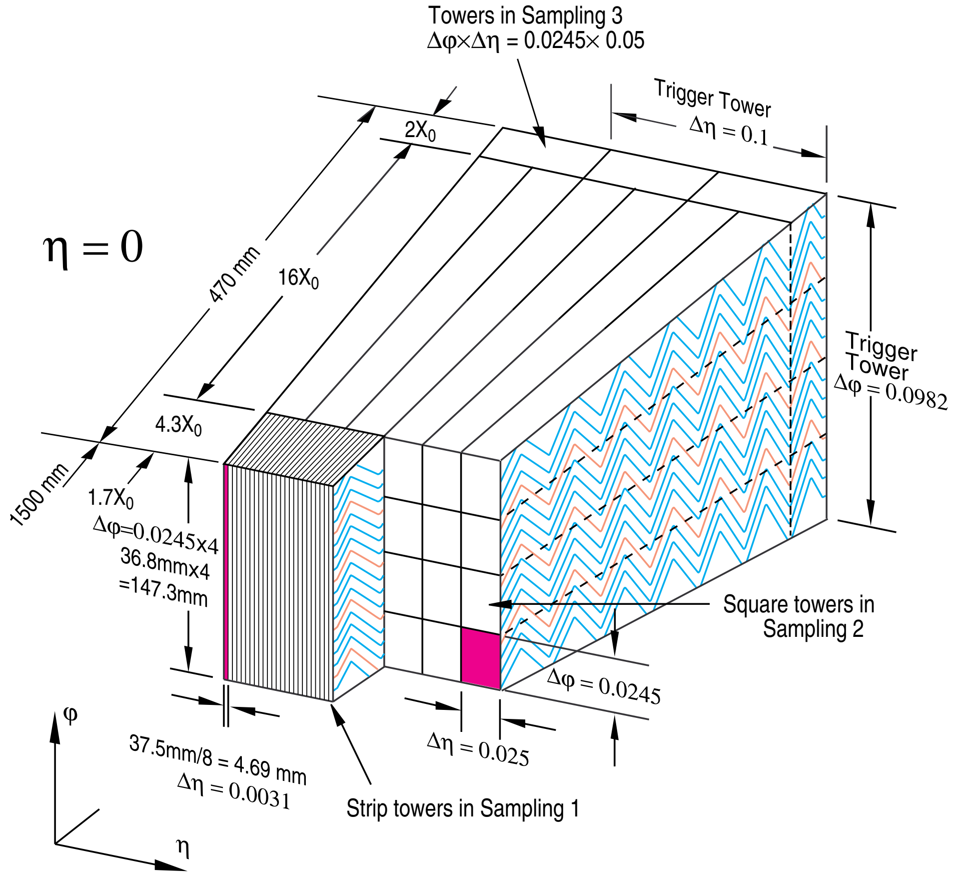

FIG. 4 – *La géométrie projective, la structure en accordéon, la structure en trois compartiments et la granularité du calorimètre électromagnétique.*

2 Préparation de l'analyse

Nous avons vu que pour observer le top il faut pouvoir isoler les particules qui proviennent des désintégrations successives. Notamment il faut isoler les électrons, provenant de la désintégration de W, des jets provenant du bruit de fond instrumental. Un tel algorithme d'identification des électrons ne peut être crée que si l'on dispose de données, nous donnant pour chaque type de particules, la réponse du détecteur à celle-ci. C'est le but de la simulation décrite ci-dessous.

2.1 Simulation

Afin de simuler les désintégrations voulues (comme $pp \to t\bar{t}$) on utilise le générateur Monte Carlo Pythia [T. Sjöstrand et al., 2002]. D'une collision de deux faisceaux de protons à 7 TeV celui-ci simule la création des particules et des processus qui s'en suivent. En sortie de ce logiciel on a un fichier contenant toutes les particules stables créees. Ce fichier est différent entre deux simulations car la création de telle ou telle particule est un processus aléatoire dépendant de la valeur de la section efficace de production de la particule.

Après avoir simulé les désintégrations qui peuvent avoir lieu dans Atlas, il faut un logiciel qui nous permet de simuler la réponse de l'appareillage au passage des particules créees. C'est le role de Geant [R. Brun et al., 1986].

2.2 Reconstruction

Geant nous donne des données brutes (des coups dans le détecteur interne, des énergies dans le calorimètre...), on applique un dernier logiciel nommé Athena pour avoir des informations exploitables.

Reconstruction dans le détecteur interne : Dans le détecteur interne les traces des particules sont reconstruites à partir des informations obtenues dans l'ensemble des détecteurs pixels+SCT (appelés "détecteurs de précision") et le TRT. Un algorithme appelé xKalman permet la reconstruction des traces dans le détecteur interne. Le principe est de rechercher des prototraces initiées dans les détecteurs de précision, puis de les extrapoler dans les détecteurs à pailles sachant que les trajectoires des particules chargées dans le champ magnétique de l'aimant solénoïdal sont décrites par des hélices, dont les différents paramètres sont evalués au point de plus petite approche du point d'interaction (A_0 le paramètre d'impact).

Reconstruction dans le calorimètre électromagnetique : La reconstruction dans le calorimètre s'effectue en plusieurs étapes :

- Les dépots d'énergie issus de Geant sont stockés dans des matrices, avec un nombre de cellules différents suivant les compartiments touchés. Les énergies dans les cellules sont calibrées en demandant que l'énergie totale déposée par un photon d'énergie transverse 50 GeV, soit égale à l'énergie incidente de cette particule ;
- Les cellules sont groupées en matrice de granularité $\Delta\eta \times \Delta\Phi = 0.025 \times 0.025$. Un algorithme par glissement de fenêtre (5×5) est alors appliqué. Un seuil est appliqué et les amas (clusters) situés au-dessus de ce seuil sont gardés ;
- Les paramètres de la gerbe sont mesurés dans cet amas : développement latéral, longitudinal, etc...

Le programme de reconstruction produit comme résultat un fichier contenant l'ensemble des données reconstruites pour chaque détecteur, ces données se présentent sous la forme d'un tableau (un Ntuple). Pour une collision donnée, on aura des informations sur les particules produites (vérité Monte Carlo), et sur la réponse du détecteur à ces particules. C'est en confrontant ces deux informations que l'on pourra ainsi écrire un algorithme d'identification des électrons et des jets.

2.3 Les données utilisées

Pour faire notre étude nous avons utilisé des données simulées dans le cadre du DC1 (Data Challenge 1) avec la disposition du détecteur final. On a travaillé - pour la basse luminosité - avec des lots d'électrons isolés de 25 GeV et des lots de dijets dits à bas $p_T > 17GeV/c$. On a travaillé - pour la haute luminosité - avec des lots d'électrons isolés de 30 GeV et des lots de dijets dits à haut $p_T > 25GeV/c$. Ces données sont utilisées pour étudier la séparation électron-jet. Celle-ci n'est pas étudiée directement sur des lots d'événements $t\bar{t}$. Une telle étude pourra être envisagée dans l'avenir mais dépasse le cadre du stage.

Les lots de dijets contiennent surtout des jets QCD (provenant de l'hadronisation de paires qq, qg ou gg, q pour quark et g pour gluon), mais d'autres processus physique commme la production des $W, Z, t\bar{t}, \gamma\gamma, \gamma q$ et γg peuvent s'ajouter, ils seront enlevés pour le calcul du rejet. De même un grand nombre d'électrons seront présents dans les jets. Une procédure pour les enlever sera décrite.

Un filtre est appliqué aux dijets pour rejeter les évenements qui ne pourraient jamais passer le système de déclenchement de niveau 1 (LV1). En moyenne $1/11,99$ des événements sont réellement présents dans les fichiers à basse luminosité et environ $1/9,8$ à haute luminosité. Ainsi chaque fichier contenant initialement 5000 événements n'en contiendra plus que ~ 400.

A ces événements sont ajoutés des événements dits d'empilement (pile-up). Ils correspondent à des interactions où les protons "échangent" une faible partie de leur impulsion. Les particules produitent ont alors une impulsion transverse faible. Elles vont déposer leurs faibles énergies dans les détecteurs et viendront perturber la bonne détection d'un événement intéressant. Le nombre de ces événements par croisement de faisceau suit une loi de Poisson dont la valeur moyenne est définie par la section efficace et la luminosité. Pour Atlas on s'attend à $4,6$ événements d'empilement à basse luminosité et ~ 23 événements d'empilement à haute luminosité.

Le maniement des données de jets simulés est malaisé. Un grand nombre d'événement a été engendré, près de 15 millions. Même si seuls les événements ayant passé le filtre ont été simulés et reconstruits, celà représente plus de 1 millions d'événements. De plus ceux-ci sont répartis sur près de 3000 fichiers. Les données sont présentes sous la forme de ntuples issus de la reconstruction des événements par Athena.

Des fichiers réduits (en nombre et en taille) permettent de travailler plus facilement. Dans cette présélection ne survivent que les événements pour lesquels au moins un cluster remplit toutes les conditions suivantes :

- l'énergie transverse est supérieure à 17 GeV (25 GeV à haute luminosité) ;
- la fuite d'énergie dans le HCAL est petite ;
- la forme de la gerbe dans la direction η est proche de celle d'un photon ;
- il n'y a pas de deuxième maximum dans les strips.

Ces critères sont similaires à ceux qui seront appliqués par la suite, mais il était important de décrire d'ores et déjà les données sur lesquelles nous allons travailler.

Ces données sont analysées en local sur les stations de travail situées à Paris. Pour ce faire un logiciel développé au CERN, couramment utilisé en physique des particules, s'appelant Paw est choisi. C'est un utilitaire graphique servant à traiter des données aussi bien experimentales que simulées. Il gère à la fois son langage propre et des routines Fortran, plus pratiques pour certaines tâches. Mon travail dans ce stage à été de modifier ou de créer ces routines, afin de travailler sur les données décrites ci-dessus, et ainsi mettre au point l'algorithme d'identification des électrons.

3 Séparation électrons-jets

3.1 Performance de l'analyse

A la fin d'une analyse pour vérifier les performances de l'algorithme on donne l'efficacité d'identification des électrons à différentes étapes. On définit l'efficacité d'identification comme le rapport entre le nombre d'électrons présents après l'analyse N^e_{res}, et le nombre présents au départ N^e_{tot}, ces deux variables étant calculés à l'aide de la vérité Monte Carlo contenu dans les données, c'est à dire le type des particules détectées :

$$\varepsilon = \frac{N^e_{res}}{N^e_{tot}} \tag{1}$$

Pour l'estimation de l'incertitude sur l'efficacité $\sigma(\varepsilon)$, uniquement l'erreur statistique sur le nombre de clusters initiaux et finaux est considérée. Elle vaut $\sigma(\varepsilon) = \sqrt{\frac{\varepsilon}{N^e_{tot}}(1-\varepsilon)}$.

Pour estimer le facteur de réjection des événements di-jets il est nécessaire de connaître leur nombre initial. Les fichiers avec lesquels nous travaillons ne sont pas complets puisque les événements ont passé un "event filter" ne gardant que les jets susceptibles de passer le LV1. La normalisation utilise les informations du groupe Atlfast[2] donnant les facteurs de normalisation à appliquer dans différents intervalles en η. La table 2 donne les facteurs de normalisation à appliquer à basse (resp. haute) luminosité issus de l'analyse de $N_{tot} = 2000000$ (resp. 600000) événements initiaux. Dans notre analyse nous étudions $N_f = 2650$ (resp. 1184) fichiers à basse (resp. haute) luminosité. Chacun contenait initialement (*i.e* avant l'"event filter") 5000 événements. Pour obtenir le nombre de jets initiaux il faudra donc effectuer le produit :

$$N_{\text{initiaux}} = \frac{N_f \times 5000 \times N_{\text{jets}}}{N_{tot}} \tag{2}$$

η	N_{jets} ($E_T > 17$ GeV)	N_{jets} ($E_T > 25$ GeV)
$\eta < 0.8$	407546	126680
$0.8 < \eta < 1.37$	272738	94506
$1.52 < \eta < 1.8$	123606	37070
$1.8 < \eta < 2.0$	88604	26212
$2.0 < \eta < 2.47$	630937	58918

TAB. 2 – *Nombre de jets initiaux issus de l'analyse de 2 millions de di-jets d'énergie* > 17 *GeV et de 600000 de di-jets d'énergie* > 25 *GeV avant l'event filter.*

On définit la réjection R par le rapport entre le nombre de jets initiaux N^i_{jet} sur le nombre de jets mal identifiés N^{res}_{jet} :

$$R = \frac{N^i_{jet}}{N^{res}_{jet}}. \tag{3}$$

L'incertitude statistique associée est donnée par $\sigma(R) = \frac{R}{\sqrt{N^{res}_{jet}}}$.

[2] http ://mwielers.home.cern.ch/mwielers/normalisation.html

3.2 Les différents critères utilisés

Pour séparer un signal de son bruit de fond (c'est à dire dans notre étude les électrons des jets), l'approche la plus simple consiste à appliquer une série de coupures sur des variables discriminantes. Mon travail a consisté à chercher les meilleures coupures possibles sur chaque variable considérée, c'est à dire de choisir celles qui permettront de rejeter la plupart des jets tout en gardant un maximum d'électrons. On distinguera trois types de coupures décrites dans les trois parties suivantes : la première provient du trigger LV1 dont on explique le rôle plus bas, la seconde concerne les variables du calorimètre et la dernière concerne les variables du détecteur interne. Pour ces deux dernières on va donner une description brève des variables qu'elles utilisent, les critères appliqués étant donnés dans l'annexe A.

Dans chaque sélection de coupures on s'est arrangé pour que l'efficcacité soit plate en η. Pour cela un ensemble de cinq zones en η ont été définies à partir de la géometrie du détecteur :

- $0 - 0.75$: l'épaisseur de plomb des absorbeurs change à $\eta = 0.8$. Celà a déjà une influence un peu avant ;
- $0.75 - 1.37$: le crack entre le tonneau et le bouchon, qui est exclu de l'analyse, commence à 1.37 ;
- $1.52 - 1.8$: le crack finit à 1.52. A 1.8 la granularité des strips change ;
- $1.8 - 2.0$: la granularité des strips change à 2.0 ;
- $2.0 - 2.47$: au-delà, la granularité des différents compartiments ne permet plus d'analyse précise.

3.2.1 Niveau 1 du système de déclenchement (LV1)

Atlas fournira pour une collision des quantités de données astronomique, rendant le stockage de toutes celles-ci impossible. Pour réduire ce flux de données, il y aura installé sur Atlas, un système de déclenchement (trigger), dont le but sera de réduire ces données à 100-200 Mo/s (stockage permanent). Le système de déclenchement possède trois niveaux, seul le premier est effectué au niveau de l'appareillage. Il accepte la fréquence totale du croisement des paquets de particules du LHC de 40 MHz. Ce qui ne passe pas le système de déclenchement sera perdu définitivement.

LV1	em25i	em30i
E_T (GeV)	> 19	> 20
ECAL isolation (GeV)	< 3	< 5
HCAL isolation (GeV)	< 2	< 3
HCAL coeur isolation (GeV)	< 2	< 2

TAB. 3 – *Coupures LV1 pour des électrons de $E_T = 25$ GeV et de $E_T = 30$ GeV.*

Pour la sélection des électrons on utilise un des menus pour des objets électromagnétiques de basse énergie : *single object trigger em25i* pour des électrons de $E_T = 25$ GeV et *single object trigger em30i* pour des électrons de $E_T = 30$ GeV. Il demande que le cluster ait une énergie suffisante, et une fuite d'énergie dans le HCAL assez faible. Les valeurs des coupures, décrites dans la table 3 et issues de [Diaz-Gomez M. et al., 2004], sont telles qu'environ 95% des électrons sont sélectionnés. En pratique toutefois on a trouvé une efficacité plus faible de 92% pour la sélection des électrons de $E_T = 25$ GeV.

3.2.2 Calorimètre électromagnétique

Utilisation du HCAL : Les électrons déposent pratiquement toute leur énergie dans le ECAL et très peu dans le HCAL. Pour mesurer les fuites dans le HCAL, on utilise comme variable le rapport de l'énergie transverse déposée dans le premier compartiment du HCAL, dans une fenêtre de $\Delta\eta \times \Delta\Phi = 0.2 \times 0.2$ autour de la cellule du cluster (η_c,ϕ_c), sur l'énergie transverse déposée dans le ECAL, on le note $Hadr$ (c.f. figure 5).

Utilisation du deuxième compartiment du ECAL : La majorité des gerbes dans le ECAL se développent dans le deuxième compartiment, cette partie du ECAL va donc être très importante pour la réjection des jets. En effet les gerbes électromagnétiques issues d'électrons sont en général moins larges que celles des jets. Pour l'isolation des électrons on va donc utiliser les variables donnant une mesure de la taille de ces gerbes dans le deuxième compartiment du ECAL, soit :

- la forme de la gerbe dans la direction η - $R_\eta(37)$ (c.f. figure 5) : Les électrons déposent presque toute leur énergie dans des unités de cellules de taille $\Delta\eta \times \Delta\Phi = 3 \times 7$ autour de la cellule du cluster (η_c,ϕ_c). Par conséquent, le rapport $R_\eta(37)$ de l'énergie déposée dans une fenêtre 3×7 (notée $E_2(3\times7)$) sur l'énergie déposée dans une fenêtre 7×7 (notée $E_2(7 \times 7)$) est important pour les électrons ;
- la largeur de la gerbe dans la direction η - ω_{η^2} (c.f. figure 5) : La largeur de la gerbe est calculée dans des unités de cellules 3×5 autour de la cellule du cluster (η_c,ϕ_c) : $\omega_{\eta^2} = \sqrt{\frac{\sum E_c \times \eta^2}{\sum E_c} - \left(\frac{\sum E_c \times \eta}{\sum E_c}\right)^2}$.

Utilisation du premier compartiment du ECAL Pour utiliser des critères sur le premier compartiment du ECAL, il faut disposer d'un minimum d'énergie (au moins 0,5% de l'énergie totale). Par ailleurs, la granularité doît être suffisante, d'où l'application des critères pour $|\eta| < 2,37$. En effet, au-delà de $|\eta| = 2,4$, la granularité du premier compartiment n'est plus suffisante pour une analyse précise.

Des études préalables ont montré que pour les jets, il y avait un ou plusieurs maxima, mais plutôt seulement un maximum pour les électrons. Nous allons donc rechercher un deuxième maximum. Pour celà on calcule le dépôt d'énergie dans une fenêtre $\Delta\eta \times \Delta\phi = 0,125 \times 0,2$ autour du cluster, deux cellules en ϕ sont sommées. Un deuxième maximum est recherché dans la direction η. S'il y a plusieurs maxima, seuls les deux plus grands sont considérés. Des critères sur les variables suivantes sont alors appliqués :

- la différence d'énergie entre le plus grand et le plus petit des deux maxima $\Delta E = E_{max2} - E_{min}$;
- l'énergie du deuxième maximum doît dépasser un seuil dépendant de manière linéaire de l'énergie transverse pour s'assurer de la stabilité de ce maximum par rapport à une fluctuation. On utilise donc la variable $\Delta E_{max2} = E_{max2}/(1. + 0.009 \times E_T)$ à basse luminosité et $\Delta E_{max2} = E_{max2}/(1. + 0.005 \times E_T)$ à haute ;
- la largeur totale de la gerbe $\omega_{total} = \sqrt{\frac{\sum E_i(i - i_{max})^2}{\sum E_i}}$ dans une fenêtre $\Delta\eta \times \Delta\phi = 0,0625 \times 0,2$ (c.f. figure 5).

Une fois ces critères appliqués, il ne reste que des gerbes étroites présentant un seul maximum. Il est encore possible de rejeter quelques jets en utilisant la forme des gerbes :

- la largeur de la gerbe dans trois cellules en $|\eta|$ $\omega_{3\ strips} = \sqrt{\frac{\sum E_i \cdot (i - i_{max})^2}{\sum E_i}}$;
- la fraction de l'énergie hors du coeur de la gerbe $F_{side} = \frac{E(\pm3) - E(\pm1)}{E(\pm1)}$.

Les critères sur le calorimètre ont été ajustés pour chaque zone en η pour que le spectre de l'efficacité soit plat en η à environ 95% (pour les électrons qui passent le niveau LV1). Les tables 9 et 10 présentées dans l'annexe A, décrivent les seuils appliqués à basse et haute luminosité, les coupures de la table 9 sont similaires à celles de l'analyse [C. Serfon, 2004].

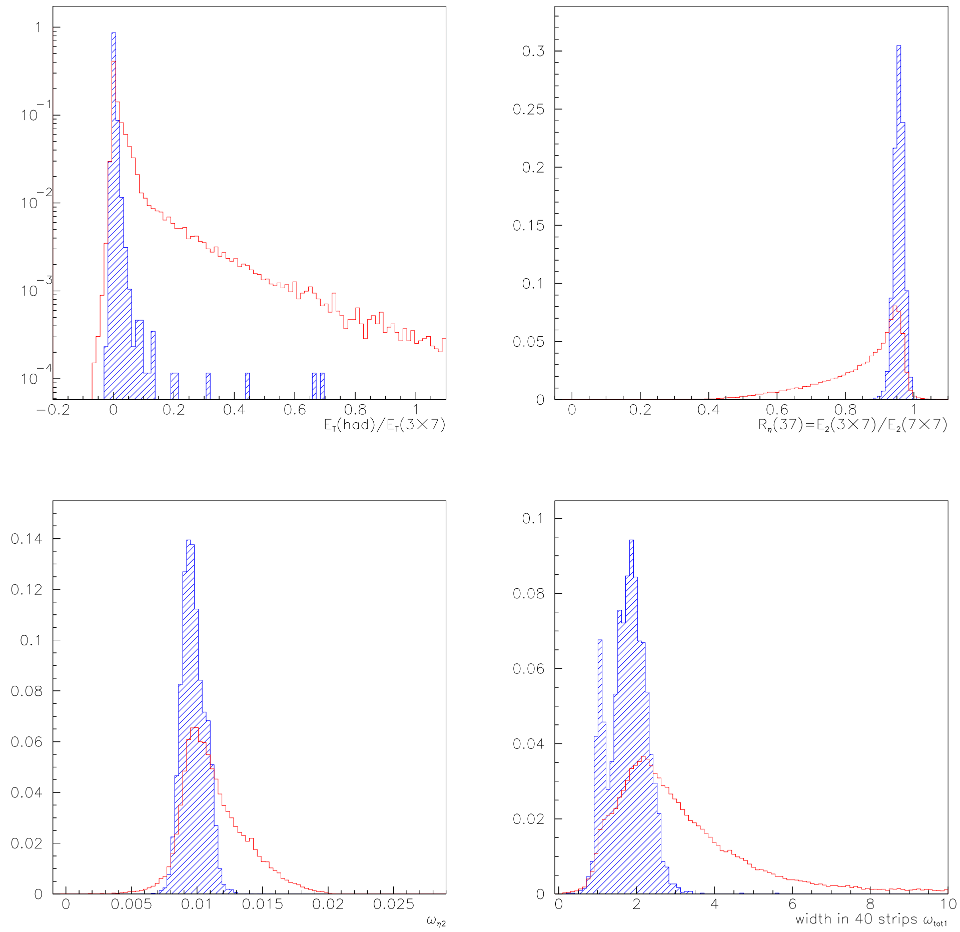

Fig. 5 – *Distributions de la fuite d'énergie dans le HCAL Hadr, de $R_\eta(37)$ la forme de la gerbe dans la direction η, de ω_{η^2} la largeur de la gerbe dans la direction η, et de ω_{total} la largeur totale dans le premier compartiment du ECAL. Les distributions sont montrées pour les jets et les électrons (histogrammes hachurés).*

3.2.3 Détecteur interne

Les variables du détecteur interne vont essentiellement nous donner des informations sur les trajectoires des particules.Ces variables peuvent être regroupées en 4 catégrories :

- Variables informant sur les qualités des traces laissées par les particules :
  - au moins neuf points de mesure dans l'ensemble Pixel+SCT ;
  - au moins un point de mesure dans la couche b des pixels ;
  - au moins deux points de mesure sont requis dans le détecteur à pixel ;
  - le paramètre d'impact transverse $|A_0| < 0,1$ cm.

- Correspondance trace-cluster : deux variables, $\Delta\eta$ et $\Delta\Phi$ résultent de la comparaison entre la valeur de η et de Φ à la sortie du détecteur interne, et celle à l'entrée du calorimètre. Ce sont donc des informations de correspondance, entre les traces de la première et de la deuxième partie de ATLAS.
- E/p, le rapport entre l'énergie E mesurée dans le calorimètre et l'impulsion p mesurée dans le trajectographe ;
- utilisation du TRT : on effectue le rapport entre le nombre de coups de rayonnement de transition N_{TRHits} et le nombre de coups total dans le TRT N_{straw}. Comme dit précédemment les électrons vont déposer des coups de "rayonnement de transition" (i.e. issus de rayons X lors du passage dans le détecteur) au contraire des pions. Ainsi plus le rapport est grand, plus il est probable que l'on ait affaire à un électron. Les coupures sur le TRT sont ajustées pour avoir une efficacité d'environ 90% sur les électrons en présence de cette seule coupure.

La table 11 présentée dans l'annexe A, montre les différentes coupures appliquées sur les variables du détecteur interne.

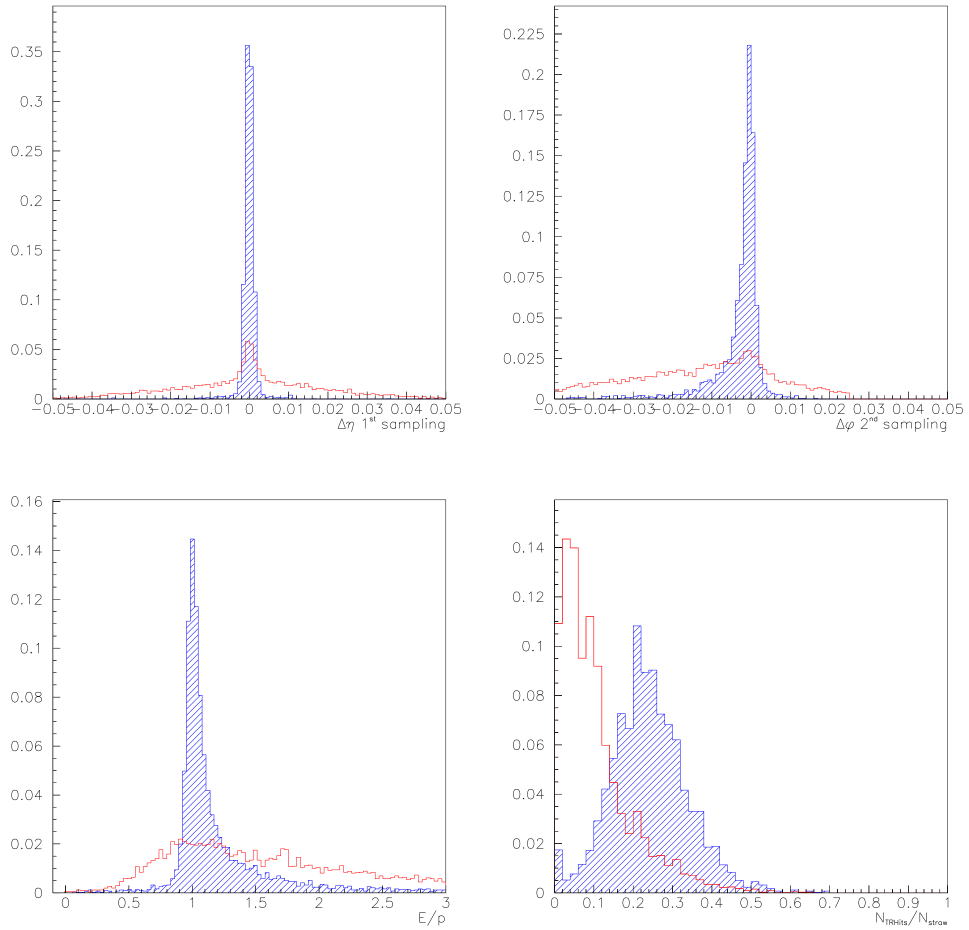

FIG. 6 – *Distributions de $\Delta\eta$, de $\Delta\Phi$, du rapport E/p et de N_{TRHits}/N_{straw}. Les distributions sont montrées pour les jets et les électrons (histogrammes hachurés).*

3.3 Corrélation entre les variables

Certaines méthodes d'analyse plus performantes que celle que nous utilisons, prennent en compte les corrélations entre les variables. En effet dans le cas de variables très corrélées, il peut y avoir des problèmes, car lorsqu'on va "couper" sur une de ces variables, cela va engendrer une coupure sur la distribution de la deuxième, ce que nous voulons éviter. Nous avons donc étudié ces corrélations ; non pas pour les utiliser pour une nouvelle méthode d'analyse, mais pour vérifier que les variables utilisées ne sont justement pas trop corrélées.

Pour cela, on calcule la matrice de corrélation entre deux variables μ et ν, dont les termes sont donnés par :

$$\rho_{\mu\nu} = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_{\mu i}-\bar{x}_{\mu})(x_{\nu i}-\bar{x}_{\nu})}{\sigma(x_{\mu})\sigma(x_{\nu})}, \tag{4}$$

avec $\bar{x_{\mu}}$ (respectivement $\bar{x_{\nu}}$) la moyenne de la distribution de la variable μ (respectivement ν), et $\sigma(x_{\mu})$ (respectivement $\sigma(x_{\nu})$) leur variance.

Au premier essai nous avons obtenu des résultats peu satisfaisants. La corrélation d'une variable par rapport à elle même n'était jamais égal à 1 (cela tournait autour de 0.950). Nous nous sommes alors rendu compte que cela était dû à un problème dans notre utilisation de Paw. En effet pour remplir un histogramme, nous le créons en donnant le nombre d'intervalles en abscisse. Ensuite nous remplissons cet histogramme avec les différentes valeurs de la variable étudiée. Or nous calculions nos moyennes et variances en utilisant ces histogrammes. Ces valeurs étaient donc dépendantes de la taille des intervalles des histogrammes. Par exemple pour le calcul de la moyenne de la fuite d'énergie dans le HCAL pour les électrons, la moyenne valait $< \text{Hadr} >= 0.00235$ pour un histogamme construit avec 100 intervalles et $< \text{Hadr} >= 0.00165$ pour 1000 intervalles. Finalement nous avons abandonné le calcul de la moyenne et de la variance à partir des histogrammes, pour le faire directement avec les valeurs des variables données par le Ntuple, suivant les formules :

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x \tag{5}$$

$$\sigma(x) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x^2 - \left(\frac{1}{N}\sum_{i=1}^{N} x\right)^2}. \tag{6}$$

La matrice de corrélation alors obtenue est présentée dans le tableau 4. Comme on s'y attendait les variables sont très peu correlées entre elles. Les seules qui font exception ($\rho > 0,2$) sont celles qui traduisent de manière similaire la forme de la gerbe. Par exemple, $R_{\eta}(37)$ qui donne le developpement latéral de la gerbe dans le deuxième compartiment est bien sur très corrélée ($\rho_{23} = -0,54$) à la variable donnant la largeur de la gerbe $\omega_{\eta 2c}$ dans ce même compartiment. Pour le premier compartiment F_{side} qui donne la fraction d'énergie déposée sur les bords de la gerbe, est elle aussi très corrélée ($\rho_{78} = 0.77$) avec la largeur totale de la gerbe. On remarque aussi que les variables provenant de parties différentes du calorimètre (HCAL, premier et deuxième compartiment du ECAL) sont très peu correlées ($\rho < 0.35$), à l'exception des variables donnant la largeur de la gerbe ($\omega_{\eta 2c}, \omega_{3\text{strips}}, \omega_{\text{total}}$). Les différentes parties du détecteur sont assez indépendantes. Nous représentons sur la figure 7 quelques unes des variables.

	Hadr	$R_\eta(37)$	$\omega_{\eta 2c}$	$\Delta E_{\max 2}$	ΔE	ω_{3strips}	ω_{total}	F_{side}
Hadr	1.000	0.035	−0.044	0.012	0.046	−0.027	0.057	0.047
$R_\eta(37)$	0.035	0.999	−0.540	−0.147	−0.124	0.148	0.216	0.083
$\omega_{\eta 2c}$	−0.044	−0.540	1.000	0.174	0.113	−0.217	−0.322	−0.145
$\Delta E_{\max 2}$	0.012	−0.147	0.174	1.000	0.553	−0.009	0.083	0.138
ΔE	0.046	−0.124	0.113	0.553	1.000	−0.117	−0.041	−0.041
ω_{3strips}	−0.027	0.148	−0.217	−0.009	−0.117	1.000	0.674	0.718
ω_{total}	0.057	0.216	−0.322	0.083	−0.041	0.674	1.000	0.773
F_{side}	0.047	0.083	−0.145	0.138	−0.041	0.718	0.773	1.000

TAB. 4 – *Matrice de corrélation pour le signal (électrons) dans les événements issus d'électrons de 25 GeV à basse luminosité.*

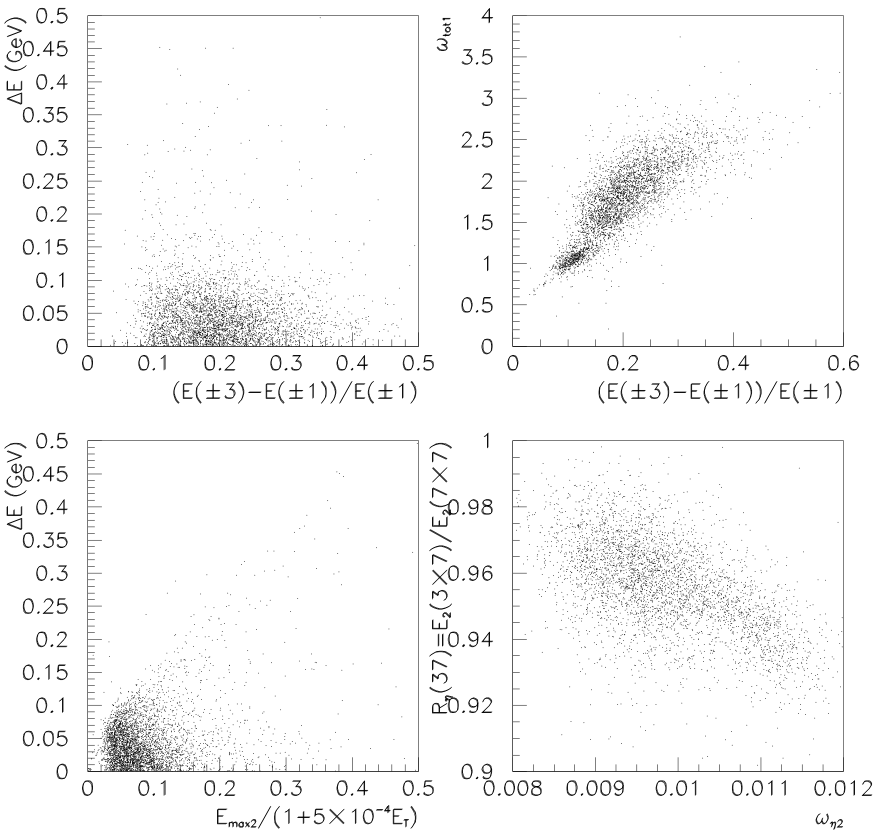

FIG. 7 – *Diagrammes montrant les distributions de ΔE en fonction de $Fside$ (en haut à gauche), ω_{tot} vs $Fside$ (en haut à droite), ΔE vs ΔE_{max2} (en bas à gauche) et $R_\eta(37)$ vs $\omega_{\eta 2c}$ (en bas à droite), pour les événements issus d'électrons de 25 GeV à basse luminosité.*

Les corrélations étant assez faibles, on ne s'attend donc pas à un gain important dans les performances avec des méthodes plus subtiles utilisant ces corrélations. Si de toute façon, une telle étude était envisagée, il faudrait commencer par essayer de gagner sur chaque compartiment indépendamment, et ne pas chercher à utiliser d'emblée toutes les variables.

4 Résultats

Nous allons dans cette partie présenter les résultats que nous avons obtenus. Nous les comparerons aux résultats de deux analyses : l'une faite en 1999 avec une version dite TDR de la géométrie du détecteur [P. Pralavorio, 1999], l'autre faite sur les mêmes données que celles que nous utilisons [C. Serfon, 2004].

4.1 Isolation des électrons

L'efficacité d'identification des électrons de 25 GeV, définie dans l'équation 1, est présentée en fonction de η sur la figure 4.1. Le trou que l'on peut observer est dû au crack entre le tonneau et les bouchons, qui n'est pas pris en compte dans l'analyse.

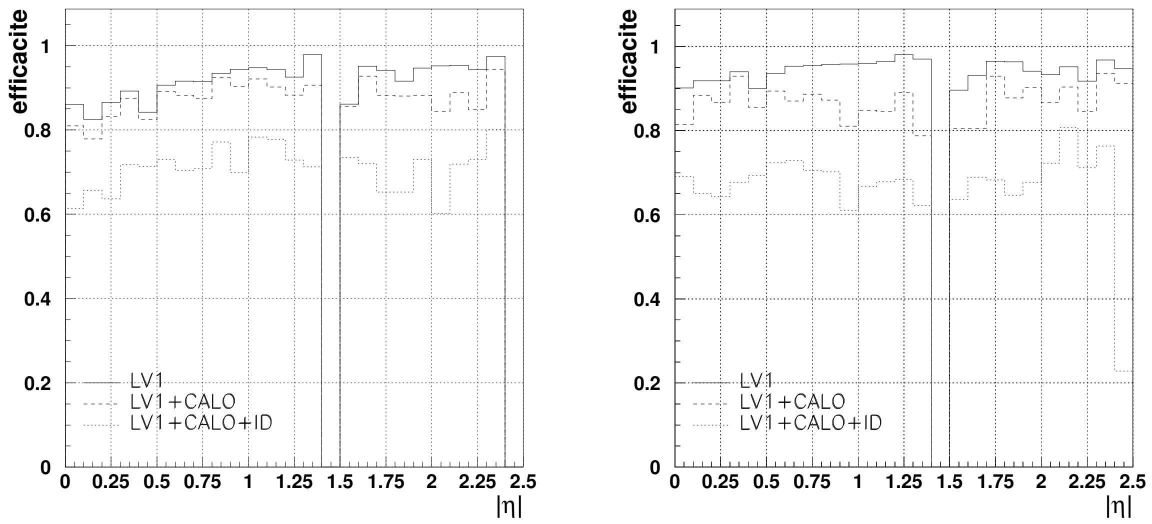

FIG. 8 – *Efficacité d'identification des électrons de 25 GeV (à gauche) de 30 GeV (à droite) en fonction de η. On différencie les résultats obtenus pour les électrons après le LV1, après le LV1+Calo et après toutes les coupures.*

La table 12 présentée dans l'annexe B donne, pour différents intervalles de η, le nombre de clusters initiaux, et l'efficacité d'identification correspondante, pour les différentes coupures appliquées.

L'efficacité totale est de $\sim 65\%$ pour les électrons de 25 GeV à basse luminosité, et de $\sim 67\%$ pour les électrons de 30 GeV à haute luminosité.

4.2 Rejet des jets

4.2.1 Composition des jets restants

Nous nous sommes aperçu qu'il restait un nombre de particules important donnant une réjection très faible. L'essentiel de ces particules sont des électrons (voir table 5). Or nous voulons rejeter les jets mais bien sur garder des électrons "intéressants". Ceux ci proviennent directement d'un quark (généralement un b). Il y a en revanche des électrons provenant de conversion que l'on veut rejeter.

Idéalement il existe dans le programme de reconstruction, un algorithme d'identification des conversions. Cependant il n'a pas été possible de l'utiliser dans cette analyse. Dans l'analyse [C. Serfon, 2004] la procédure pour différencier ces deux sources d'électrons est de considérer qu'un électron provenant directement d'un quark a une impulsion transverse $p_T > 16$ GeV, et qu'il ne provient pas d'une conversion (en utilisant la vérité Monte Carlo). Nous avons trouvé une méthode de séparation de la nature des électrons en utilisant le rayon de vertex (en utilisant la aussi la vérité Monte Carlo). En effet un électron issu directement d'un quark aura un rayon faible (on coupe à 10 cm). Une telle coupure permet d'éviter un réglage fin sur une coupure de l'impulsion transverse.

Dans la table 5 on voit qu'il reste 214 clusters à basse luminosité. Parmis eux, 145 sont des électrons provenant d'un quark, 26 sont issus de conversion, 34 sont des hadrons (ce hadron transporte plus de la moitié de l'énergie du cluster). Les 9 restants correspondent à des clusters auxquels sont associés plusieurs particules dont aucune n'est dominante. Ainsi le vrai nombre de clusters restant n'est pas de 214 mais $214 - 145 = 69$. De plus on remarque le bon travail du TRT qui rejette bien les hadrons et pas les électrons.

Composition des jets restants à basse luminosité					
	Clusters restants	e^-	conversion	hadron	Autres
Avant TRT	281	147	27	98	9
Apres TRT	214	145	26	34	9
Composition des jets restants à haute luminosité					
	Clusters restants	e^-	conversion	hadron	Autres
Avant TRT	149	88	19	42	0
Apres TRT	132	85	19	27	1

TAB. 5 – *FINAL : Type de particules restantes dans les clusters non rejetés par les coupures sur les jets à basse et haute luminosité.*

4.2.2 Réjections obtenues

La table 13 dans l'annexe B, donne les nombres de clusters finaux et les réjections après LV1 et après les coupures du calorimètre. On a une réjection d'environ 800 donc encore très loin des réjections attendues, qui sont de l'ordre de 10^5.

La table 14 dans l'annexe B, donne les nombres de clusters finaux et les réjections après les différentes coupures dans le trajectographe. On a à la fin une réjection de plus de 70000 à basse luminosité, donc assez proche de ce qu'on attend.

De plus on a noté dans la table 5 que sur les clusters restants, il y a encore un grand nombre d'électrons provenant de conversion : 26 sur un total de 69 à basse luminosité et 19 sur un total de 47 à haute luminosité. Ces électrons pourront être identifiés comme tels par un algorithme dédié, et une plus grande réjection pourra être atteinte.

Dans la table 6, on reporte les efficacités et réjections obtenues après chaque étape de l'analyse, ainsi que lorsqu'on enlève "à la main" les conversions. On peut alors presque doubler la réjection obtenue.

On peut déjà remarquer quelque chose : après les coupures sur le calorimètre on obtient pour les deux luminosités une efficacité autour de 87%, mais pour les réjections on obtient $R \sim 800$ pour la basse luminosité et $R \sim 3000$ pour la haute luminosité. Cette grosse différence entre les deux luminosités nous a parue bizarre[3].

[3]Les efficacités d'identification des électrons étant identiques à basse et haute luminosité, celà implique

	basse luminosité		haute luminosité	
Coupures	Efficacité (%)	Réjection	Efficacité (%)	Réjection
LV1	91.9 ± 0.06	4 ± 4.10^{-3}	94.2 ± 0.06	8 ± 8.10^{-3}
ECAL	87.9 ± 0.09	803 ± 10^1	86.9 ± 0.1	$3752 \pm 8, 1.10^2$
ID	78.6 ± 0.1	$10457 \pm 5, 2.10^2$	74.9 ± 0.2	$42974 \pm 3, 3.10^3$
Matching1	74.3 ± 0.2	$22972 \pm 1, 7.10^3$	71.6 ± 0.3	$105877 \pm 1, 5.10^4$
E/p	68.7 ± 0.2	$29194 \pm 2, 4.10^3$	68.4 ± 0.3	$119763 \pm 1, 5.10^4$
TRT	67.4 ± 0.2	$60056 \pm 7, 2.10^3$	67.6 ± 0.3	$155437 \pm 2, 4.10^4$
pas de conversions	67.4 ± 0.2	$95544 \pm 1, 4.10^4$	67.6 ± 0.3	$260912 \pm 4, 9.10^4$

TAB. 6 – *Efficacité et réjection après chaque coupure*

4.3 Comparaison avec les études précédentes

4.3.1 Comparaison avec le TDR

La table 7 est issue de [P. Pralavorio, 1999] et donne les performances obtenues sur des lot d'événements semblables. On notera cependant qu'à basse luminosité l'étude était faite sur des électrons ayant un $E_T = 20$ GeV contre 25 GeV désormais. De plus les réjections étaient estimées sur un lot de 10^6 di-jets contre près de $1,5 \times 10^7$ dans notre étude.

Coupures	basse luminosité		haute luminosité	
	ϵ(e20) (%)	réjection (10^3)	ϵ(e30) (%)	réjection (10^3)
LV1	94.0	0.08 ± 0.001	96.1	0.09 ± 0.001
ECAL	80.9	9.8 ± 1.2	81.1	8.4 ± 1.5
ID	77.4	16.8 ± 2.7	77.2	22.7 ± 6.9
Matching	75.4	40 ± 10	75.3	35.8 ± 13
TRT	68.5	> 150	67.5	> 45

TAB. 7 – *Efficacités et réjections telles que décrites dans [P. Pralavorio, 1999].*

Concernant l'efficacité d'identification des électrons, on peut remarquer que (à basse luminosité) :
- les coupures au niveau LV1 sont plus dures dans notre analyse;
- les coupures au niveau du ECAL sont plus laches dans notre analyse, coupant 96% des électrons restants contre 86% dans le TDR. Les seuils diffèrent en effet un peu. Il faudra plus de travail pour comprendre l'origine des différences.
- l'effet des coupures sur la correspondance trace-cluster et le E/p sont plus dures dans notre analyse, coupant 91% des électrons restants contre 97% dans le TDR. Il semblerait que ce soit la coupure sur $\Delta\eta$ qui soit plus contraignante désormais puisque la coupure était à $|\Delta\eta| < 0.01$ dans [P. Pralavorio, 1999].
- les coupures sur le TRT sont plus laches dans notre analyse, coupant 95% des électrons restants contre 91% dans le TDR.

Ainsi si les efficacités globales sont quasi-identiques, il y a d'importantes différences pour chaque coupure.

que les critères sont plus "durs" à haute luminosité - environnement plus difficile - ce qui entrainera alors un pouvoir de réjection plus élevé.

Concernant la réjection, on peut remarquer que (à basse luminosité) :

- les coupures sur le calorimètre étant plus laches dans notre analyse, la réjection est beaucoup plus faible : $\sim$ 800 contre près de 9800 dans le TDR;
- avant l'application du critère sur le TRT les réjections sont cependant du même ordre : $\sim$ 31000 dans notre analyse contre $\sim$ 40000 dans le TDR.

Après application de la coupure sur le TRT, il reste dans [P. Pralavorio, 1999] seulement un évenement. Lorsqu'il y a si peu d'évenements on donne une limite, un niveau de confiance, par exemple à 90%. En effet une même analyse faite sur d'autres données similaires aurait pu donner 0, 1, 2, ou 3 évenements. La probabilité d'obtenir un tel nombre N d'évenements est régie par la statistique de Poisson. Ainsi, suivant les prescriptions de [PDG, 2002], si on a "vu" un événement, il y a 90% de chances que le "vrai" nombre était compris entre $[0, 105; 3.89]$. En ce qui nous concerne, on a plus d'événements. On peut considérer que l'erreur Poissonienne est de l'ordre de $\sqrt{N}$. Pour obtenir un intervalle à 90% de niveau de confiance on considérera que le "vrai" nombre d'événements est compris entre $N - 1,68\sqrt{N}, N + 1,68\sqrt{N}$. Ainsi, pour 57 événements observés, la limite supérieure est de $\sim$ 70 événements. Celà correspond à une réjection $>$ 60000 à 90% de niveau de confiance. Ce chiffre est donc moindre que le $>$ 150000 indiqué dans le TDR. Nous avons vu cependant qu'un rejet efficace des conversions pourrait augmenter ce nombre.

4.3.2 Comparaison avec une étude DC1

Dans la table 8 on reporte l'efficacité moyenne obtenue par Cédric Serfon du laboratoire de Marseille, qui a fait une analyse similaire. Dès la coupure LV1 nos résultat diffèrent, et surtout pour la coupure LV1 sur les lots à haute luminosité. En effet Cédric trouve une efficacité de 88.5% et nous 94.2%, nous avons remarqué que quand nous nous utilisons les coupures LV1 basse luminosité sur les lots à haute luminosité nous trouvons une efficacité de 85%. Peut être que les différences viennent de celà. Dans son analyse Cédric aurait ainsi utilisé le même LV1 pour les deux lots[4] Toutefois nous retrouvons à peu près les même valeurs d'efficacité après les coupures sur le calorimètre. Pour la basse luminosité celà est normal car nous utilisons les mêmes valeurs de coupures. Pour la haute luminosité Cédric rattrape la coupure importante sur le LV1 par des coupures beaucoup plus lâche sur le calorimètre, ce qui fait qu'au final on retrouve des efficacité similaires. Au final notre efficacité globale est plus faible. Celà est dû essentiellement à la coupure E/p qui est plus sévère dans notre analyse. On utilise $E/p < 2.4$ contre $E/p < 3$ dans [C. Serfon, 2004]. Cependant nos seuils sont eux en accord avec ceux définis par le groupe du système de déclenchement [Diaz-Gomez M. et al., 2004]. Dans [C. Serfon, 2004] est

Coupures	basse luminosité	haute luminosité
LV1	90.2 ± 0.4	88.5 ± 0.3
ECAL	87.2 ± 0.5	86.1 ± 0.3
ID	77.4 ± 0.6	74.6 ± 0.4
Matching1	75.1 ± 0.6	73.2 ± 0.4
E/p	74.3 ± 0.6	72.4 ± 0.5
TRT	73.6 ± 0.6	71.7 ± 0.5

TAB. 8 – *Efficacités après chaque coupure telles que décrites dans [C. Serfon, 2004].*

[4] A la suite de mon stage des contacts plus importants avec Cédric seront entrepris.

reporté une réjection > 66000 à basse luminosité. Notre estimation est en accord puisque nous trouvons une réjection > 60000 à 90% de niveau de confiance.

4.4 Conclusion

Au final nous avons trouvé une réjection > 60000, à 90% de niveau de confiance, est obtenue à basse luminosité. Ce nombre est en bon accord avec une étude similaire récente faite au laboratoire de Marseille (CPPM), il est toutefois près d'un facteur deux plus bas que ceux obtenus lors de la préparation du Technical Design Report de Physique d'Atlas de 1999. Il est aussi encore assez loin de la réjection de 10^5 à basse luminosité nécessaire pour atteindre des performances optimales dans la plupart des analyses de physique.

On voit donc qu'il rester encore du travail á réaliser pour atteindre la réjection de 10^5 à basse luminosité nécessaire pour atteindre des performances optimales dans la plupart des analyses de physique. De nombreuses voies d'amélioration des performances sont envisageables. En particulier un effort plus important doit être fait sur les coupures du calorimètre. Enfin un rejet efficace des conversions est souhaitable.

5 Résumé du travail effectué au cours du stage

Comme on l'a vu dans la partie 1, afin de permettre l'étude du quark top, il faut pouvoir séparer les électrons des jets, c'est sur ce problême que j'ai travaillé durant mon stage au sein du groupe Atlas du LPNHE.

L'identification des électrons et la réjection des jets est d'une importance cruciale dans Atlas. Une réjection de l'ordre de 10^5 est nécessaire.

Dans ce stage, un algorithme d'identification d'électrons isolés, dans une gamme d'énergie moyenne ($E_T = 25$ GeV) a été utilisé. Celui-ci utilise la caractérisation du développement latéral et longitudinal d'une gerbe dans le calorimètre électromagnétique grâce à la géométrie de ce détecteur, en particulier sa granularité. L'algorithme utilise de plus les informations du trajectographe, en associant une trace chargée, reconstruite dans le détecteur interne, avec un cluster du calorimètre électromagnétique. Un ensemble de variables discriminantes est défini. Des seuils sont définis pour identifier les électrons avec une efficacité de l'ordre de 70%. En appliquant la même procédure à des lots dévénements contenant des di-jets, on obtient le pouvoir de rejet de l'algorithme a ces jets. Une réjection > 60000, à 90% de niveau de confiance, est obtenue à basse luminosité. Ce nombre est en bon accord avec une étude similaire récente faite au laboratoire de Marseille (CPPM). Ce nombre est toutefois près d'un facteur deux plus bas que ceux obtenus lors de la préparation du Technical Design Report de Physique d'Atlas de 1999. De nombreuses voies d'améliorations des performances sont envisageables. En particulier un effort plus important doit être fait sur les coupures du calorimètre. Enfin un rejet efficace des conversions est souhaitable.

A Coupures sur les variables

Les tables 9 et 10 décrivent les seuils appliqués à basse et haute luminosité.

	$Hadr$	$R_\eta(37)$	ω_{η^2}	ΔE_{max2}	ΔE	$\omega_{3\text{strips}}$	ω_{tot1}	F_{side}
$0 < \|\eta\| < 0.75$	< 0.02	> 0.90	< 0.012	> 0.25	> 0.25	> 0.76	> 2.7	> 0.25
$0.75 < \|\eta\| < 1.4$	< 0.01	> 0.90	< 0.012	> 0.50	> 0.15	> 0.80	> 3.3	> 0.5
$1.52 < \|\eta\| < 1.8$	< 0.02	> 0.90	< 0.012	> 1.00	> 0.40	> 0.80	> 3.5	> 1.0
$1.8 < \|\eta\| < 2.0$	< 0.02	> 0.91	< 0.012	> 0.50	> 0.25	> 0.76	> 2.2	> 0.5
$2.0 < \|\eta\| < 2.47$	< 0.02	> 0.91	< 0.013	> 0.40	> 0.20	> 0.74	> 1.6	> 0.4

TAB. 9 – *Coupures appliquées sur les variables du calorimètre, suivant η à basse luminosité pour des électrons de $E_T = 25$ GeV.*

	$Hadr$	$R_\eta(37)$	ω_{η^2}	ΔE_{max2}	ΔE	$\omega_{3\text{strips}}$	ω_{tot1}	F_{side}
$0 < \|\eta\| < 0.75$	< 0.03	> 0.90	< 0.012	> 0.25	> 0.45	> 0.76	> 3.5	> 0.35
$0.75 < \|\eta\| < 1.4$	< 0.02	> 0.90	< 0.012	> 0.50	> 0.25	> 0.80	> 3.5	> 0.50
$1.52 < \|\eta\| < 1.8$	< 0.04	> 0.90	< 0.013	> 0.90	> 0.50	> 0.80	> 2.8	> 0.68
$1.8 < \|\eta\| < 2.0$	< 0.04	> 0.90	< 0.013	> 0.90	> 0.80	> 0.76	> 4.0	> 0.30
$2.0 < \|\eta\| < 2.47$	< 0.02	> 0.89	< 0.013	> 1.00	> 0.75	> 0.74	> 2.0	> 0.225

TAB. 10 – *Coupures appliquées sur les variables du calorimètre, suivant η à haute luminosité pour des électrons de $E_T = 30$ GeV.*

La table 11 présente les coupures appliquées sur les variables du détecteur interne.

	$\|\Delta\eta\|$	$\|\Delta\phi\|$	E/p (e25)		E/p (e30)		N_{TR}/N_{str}
$0 < \|\eta\| < 0.75$	< 0.005	< 0.02	$0.8 <$	< 1.3	$0.7 <$	< 1.7	> 0.06
$0.75 < \|\eta\| < 1.4$	< 0.003	< 0.02	$0.8 <$	< 1.3	$0.7 <$	< 1.7	> 0.08
$1.52 < \|\eta\| < 1.8$	< 0.004	< 0.02	$0.7 <$	< 2.5	$0.7 <$	< 2.7	> 0.11
$1.8 < \|\eta\| < 2.0$	< 0.006	< 0.02	$0.7 <$	< 2.5	$0.7 <$	< 2.7	> 0.10
$2.0 < \|\eta\| < 2.47$	< 0.003	< 0.02	$0.7 <$	< 2.5	$0.7 <$	< 2.7	> 0.05

TAB. 11 – *Coupures appliquées sur les variables du détecteur interne, suivant η.*

B Efficacité et rejection suivant η

Efficacité en % à basse luminosité							
η	Clusters initiaux	LV1	ECAL	ID	Match1	E/p	TRT
tout	4323	91.9	87.9	78.6	74.3	68.7	67.4
$0 < \|\eta\| < 0.75$	1395	87.9	84.4	73.5	71.2	66.9	66.2
$0.75 < \|\eta\| < 1.4$	1168	93.8	90.2	83.4	80.3	71.1	69.0
$1.52 < \|\eta\| < 1.8$	556	91.4	88.5	76.8	74.1	72.7	70.9
$1.8 < \|\eta\| < 2.0$	355	93.5	87.9	76.1	74.6	70.4	70.1
$2.0 < \|\eta\| < 2.47$	849	95.8	89.3	82.3	71.1	65.1	63.8
Efficacité en % à haute luminosité							
η	Clusters initiaux	LV1	ECAL	ID	Match1	E/p	TRT
tout	2134	94.2	86.9	74.9	71.6	68.4	67.6
$0 < \|\eta\| < 0.75$	700	92.6	87.4	72	70.3	68.7	68.7
$0.75 < \|\eta\| < 1.4$	581	96.4	85.0	74.9	71.6	68.0	66.6
$1.52 < \|\eta\| < 1.8$	249	93.2	84.7	72.3	70.7	68.7	67.1
$1.8 < \|\eta\| < 2.0$	184	95.1	89.1	73.4	72.3	67.4	66.3
$2.0 < \|\eta\| < 2.47$	420	94.3	89.0	81.9	74.0	68.8	67.8

TAB. 12 – *Nombre de clusters initiaux et éfficacité d'identification des électrons sur le lot d'électrons de 25 GeV à basse luminosité. Les résultats sont donnés par intervalle de pseudorapidité, après chaque étapes d'analyse.*

Réjection dans le calorimètre à basse luminosité					
η	Clusters initiaux	LV1		ECAL	
tout	4203960	962480	4.4	5230	803.8
$0 < \|\eta\| < 0.75$	1455327	333191.5	3.9	1678	867.3
$0.75 < \|\eta\| < 1.4$	1148404	262614.8	4.4	1608	714.2
$1.52 < \|\eta\| < 1.8$	479677	99348.6	4.8	581	825.6
$1.8 < \|\eta\| < 2.0$	347707	76929.6	4.5	327	1063.3
$2.0 < \|\eta\| < 2.47$	772846	15322	5.0	1036	745.9
Réjection dans le calorimètre à haute luminosité					
η	Clusters initiaux	LV1		ECAL	
tout	7305540	893098	8.2	1947	3752.2
$0 < \|\eta\| < 0.75$	2695117	359349	7.5	561	4804.1
$0.75 < \|\eta\| < 1.4$	2010620	257772	7.8	674	2983.1
$1.52 < \|\eta\| < 1.8$	788664	83017	9.5	174	4532.5
$1.8 < \|\eta\| < 2.0$	557660	61962	9.0	107	5211.8
$2.0 < \|\eta\| < 2.47$	1253480	131945	9.5	431	2908.3

TAB. 13 – *FINAL : Nombre de clusters initiaux, nombres de clusters et réjection après les coupures LV1 et calorimètre. Les résultats sont donnés par intervalle de η, à basse et haute luminosité.*

Réjection dans le détecteur interne à basse luminosité									
η	Clusters initiaux	ID		Match1		E/p		TRT	
tout	4203960	402	10457	183	22972	144	29194	70	60056
$0 < \|\eta\| < 0.75$	1455327	124	11736	63	23100	51	28536	32	45479
$0.75 < \|\eta\| < 1.4$	1148404	156	7361	60	19140	42	27343	10	114840
$1.52 < \|\eta\| < 1.8$	479677	33	14535	21	22841	21	22842	8	59960
$1.8 < \|\eta\| < 2.0$	347707	16	21731	15	23180	12	28976	4	86927
$2.0 < \|\eta\| < 2.47$	772846	73	10586	24	32201	18	42936	16	48303
Réjection dans le détecteur interne à haute luminosité									
η	Clusters initiaux	ID		Match1		E/p		TRT	
tout	7305540	170	42974	69	105877	61	119763	47	155437
$0 < \|\eta\| < 0.75$	2695117	44	61253	20	134756	19	141848	14	192508
$0.75 < \|\eta\| < 1.4$	2010620	76	26455	27	74467	23	87418	17	118272
$1.52 < \|\eta\| < 1.8$	788664	7	112666	3	262888	3	262888	2	394332
$1.8 < \|\eta\| < 2.0$	557660	11	50696	6	92943	4	139415	2	278830
$2.0 < \|\eta\| < 2.47$	1253480	32	39171	13	96421	12	104457	12	104457

TAB. 14 – *FINAL : Nombre de clusters initiaux, nombres de clusters et réjection après les coupures sur le détecteur interne. Les résultats sont donnés par intervalle de η, à basse et haute luminosité.*

Références

[ATLAS TDR Phys Vol 2, 1999] ATLAS collaboration *TDR ATLAS detector and physics performance, vol. 2*, 1999, CERN/LHCC 99-15, ATLAS-TDR-15

[ATLAS TDR Phys Vol 1, 1999] ATLAS collaboration *TDR ATLAS detector and physics performance, vol. 1*, 1999, CERN/LHCC 99-14, ATLAS-TDR-14

[T. Sjöstrand et al., 2002] T. Sjöstrand et al. *The complete PYTHIA 6.206 manual*, 2002, LU TP 01-21 [hep-ph/0108264]

[R. Brun et al., 1986] R. Brun et al. *GEANT 3*, 1986, CERN/DD/EE/84-1

[P. Pralavorio, 1999] P. Pralavorio. *Electron/jet separation with the Atlas detector*, 1999, ATL-PHYS-99-015

[C. Serfon, 2004] C. Serfon *Séparation électron-jets*, 2004, présentation aux Journées de Physique Atlas France

[Diaz-Gomez M. et al., 2004] Diaz-Gomez M. et al.. *The e-γ analysis framework*, 2004, ATL-COM-DAQ-2004-010

[PDG, 2002] K. Hagiwara et al., *Particle Data Group*, 2002, Phys Rev D 66, 010001