

Rapport de stage

Master 1 Physique Fondamentale

---

Étude des électrons provenant du  $J/\Psi$   
dans l'expérience ATLAS

---

*Auteur :*  
SYLVESTRE PIRES

*Superviseurs :*  
FRÉDÉRIC DERUE  
TIMOTHÉE THEVENEUX-PELZER

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Présentation du laboratoire . . . . .	2
1.2	Le groupe ATLAS . . . . .	2
<b>2</b>	<b>Sujet et problématique</b>	<b>3</b>
2.1	Le contexte scientifique . . . . .	3
2.2	Le dispositif expérimental . . . . .	3
2.3	Approche adoptée . . . . .	4
<b>3</b>	<b>Travail personnel réalisé</b>	<b>5</b>
3.1	Les données utilisées . . . . .	5
3.2	Premiers pas . . . . .	5
3.3	Identification des électrons . . . . .	5
3.3.1	Calorimètre hadronique . . . . .	5
3.3.2	Calorimètre électromagnétique . . . . .	6
3.3.3	Détecteur interne . . . . .	6
3.3.4	Combinaison des deux sous-détecteurs . . . . .	6
3.3.5	Efficacité et pureté des coupures . . . . .	7
3.4	Utilisation de TMVA . . . . .	7
3.4.1	Les méthodes d'identification utilisées . . . . .	7
3.4.2	Comparaison des différentes méthodes . . . . .	8
<b>4</b>	<b>Conclusion et discussion</b>	<b>9</b>
4.1	Conclusion . . . . .	9
4.2	Bilan personnel du stage . . . . .	9
<b>5</b>	<b>Annexe</b>	<b>11</b>
5.1	Figure du dispositif expérimental . . . . .	11
5.2	Identification des électrons . . . . .	12
5.3	Utilisation de TMVA . . . . .	13

# 1 Introduction

## 1.1 Présentation du laboratoire

Le **Laboratoire de Physique Nucléaire des Hautes Énergies** [1], situé sur le campus universitaire de Jussieu est une unité mixte de l'**Institut National de Physique Nucléaire** et de **Physique des Particules** [2] et des université Pierre et Marie Curie, et Paris Diderot. Regroupant quatre-vingt seize chercheurs et doctorants, cinquante deux ingénieurs, techniciens et personnels administratifs, ce laboratoire est impliqué dans de nombreux programmes expérimentaux internationaux. Ses activités réunissent différents domaines de recherches :

**La physique des particules** en participant aux analyses de données au sein des expériences CDF et DØ au Tevatron<sup>1</sup> à Fermilab, et BaBar - physique du quark b - au SLAC<sup>2</sup> National Laboratory. Le LPNHE est également présent sur ATLAS<sup>3</sup> et LHCb au CERN, mais aussi sur T2K<sup>4</sup> au Japon et dans le projet de futur accélérateur ILC<sup>5</sup>.

**La cosmologie observationnelle** en s'impliquant dans les programmes d'observations Supernova Legacy Survey - étude de l'énergie noire - du Canadian French Hawai Telescope, dans Nearby Supernova Factory - étude de l'expansion de l'univers - et en s'engageant auprès des projets Supernova Acceleration Probe et Large Synoptic Survey Telescope.

**La nature et l'origine du rayonnement cosmique** avec le réseau de télescopes H.E.S.S.<sup>6</sup> en Namibie et à l'observatoire Auger en Argentine.

Pour l'ensemble de ces collaborations, le laboratoire s'engage dans la recherche et le développement de détecteurs, la participation à la construction, l'étalonnage et la mise en place des instruments, l'analyse des données et leur interprétation physique jusqu'à la publication des résultats.

## 1.2 Le groupe ATLAS

Le groupe ATLAS du LPNHE, composé d'une trentaine de chercheurs dont 7 étudiants en thèse, est fortement impliqué dans le calorimètre électromagnétique du détecteur ATLAS, de sa réalisation au suivi de ses performances et dans les logiciels de reconstruction et d'identification des photons et des électrons. Le groupe s'est aussi engagé dans les activités de grille de calculs, de développement des outils d'analyse de données ou encore de mise au point d'algorithmes de déclenchement d'acquisition.

Le laboratoire prépare et réalise l'analyse de données pour l'étude de la masse et de la section efficace de production du quark top ainsi que la recherche du boson de HIGGS. Il s'est intéressé à la caractérisation complète des divers bruits de fond dans le canal de désintégration en deux photons du boson de HIGGS, *via* des simulations théoriques et l'étude du système de détection des photons, améliorant les perspectives de découverte dans ce canal.

---

1. Collisionneur proton-antiproton  
2. *Stanford Linear Accelerator Center*  
3. *A Toroidal LHC Apparatus*  
4. *Tokai to Kamioka* : physique des neutrinos  
5. *International Linear Collider*  
6. *High Energy Stereoscopic System* : astronomie gamma au TeV

## 2 Sujet et problématique

### 2.1 Le contexte scientifique

Dans les années 1960 et 1970, les nombreuses avancées expérimentales et théoriques dans le domaine de la physique des particules ont conduit au développement du modèle standard [3]. Ce modèle est une théorie quantique des champs décrivant de façon exhaustive les interactions entre les particules connues, et permettant de distinguer deux types de particules : les bosons de jauge qui véhiculent les interactions fondamentales (électromagnétique, faible et forte) et les fermions constituant la matière. Les fermions peuvent être séparés en deux catégories : les quarks et les leptons. Ces deux catégories étant elles mêmes subdivisées en trois familles. Enfin, à chaque particule est associée son antiparticule, de même masse et de même spin, mais de charges opposées.

	1 <sup>ère</sup> famille	2 <sup>ème</sup> famille	3 <sup>ème</sup> famille
Quarks	$\begin{pmatrix} \textit{up } u \\ \textit{down } d \end{pmatrix}$	$\begin{pmatrix} \textit{charm } c \\ \textit{strange } s \end{pmatrix}$	$\begin{pmatrix} \textit{top } t \\ \textit{bottom } b \end{pmatrix}$
Leptons	$\begin{pmatrix} \textit{neutrino } \nu_e \\ \textit{electron } e^- \end{pmatrix}$	$\begin{pmatrix} \textit{neutrino } \nu_\mu \\ \textit{muon } \mu \end{pmatrix}$	$\begin{pmatrix} \textit{neutrino } \nu_\tau \\ \textit{tau } \tau \end{pmatrix}$

TABLE 1 – Les particules de type fermion du modèle standard

La première famille permet de reconstituer toute la matière ordinaire, en effet des associations de quarks  $u$  et  $d$  forment des protons et des neutrons qui avec des électrons formeront des atomes. Ce modèle a également permis d'unifier les interactions électromagnétique et faible en une seule : l'interaction électrofaible. Néanmoins ce modèle n'est pas complet, en effet il n'explique pas pourquoi et comment les particules acquièrent-elles leur masse, mais aussi pourquoi ces masses sont elles si différentes - de 0 à 172GeV -. Ces questions peuvent trouver une explication *via* une brisure spontanée de symétrie, ce que le mécanisme de HIGGS fait en introduisant un nouveau champ scalaire massif, le boson de HIGGS.

La découverte de ce boson est un enjeu important pour valider le modèle standard. Actuellement, ce modèle est compris comme une théorie effective à basse énergie, mais qui doit être liée à une théorie plus fondamentale qui expliquerait par exemple l'origine de la matière noire. Les candidats à la nouvelle physique sont nombreux : supersymétrie, théories à dimensions supplémentaires ou encore théories de gravité quantique. Afin de pouvoir tester ces théories, la communauté scientifique a eu besoin d'accélérateurs de particules pouvant atteindre le TeV ; c'est dans cette optique qu'a été créé le LHC.

### 2.2 Le dispositif expérimental

Le LHC [4], *Large Hadron Collider*, est un collisionneur proton-proton de 27 km de circonférence situé au CERN [5] à Genève. Cet accélérateur est conçu pour atteindre une énergie dans le centre de masse de 14 TeV.

Depuis début 2010, le LHC opère à 7 TeV, accumulant des données jusque fin 2012 pour une première phase d'analyse. La machine sera ensuite arrêtée pour effectuer les travaux nécessaires à la montée à 14 TeV et une augmentation de la luminosité en 2014. Quatre expériences ont été mises en place sur le faisceau du LHC afin d'étudier les particules produites lors des collisions ; LHCb se concentre sur la physique du quark  $b$ , ALICE<sup>7</sup> a pour but d'étudier le plasma de quarks et de gluons, enfin CMS<sup>8</sup> et ATLAS sont des expériences plus généralistes qui concernent le modèle standard et la recherche de nouvelle physique.

Le détecteur ATLAS, de forme cylindrique, mesure 25 m de haut et 44 m de longueur pour un poids total d'environ 7000 tonnes (FIG.2). Il est à ce jour le plus grand des détecteurs construit pour une expérience de physique des particules. Dans le détecteur, les particules se désintègrent

7. *A Large Ion Collider Experiment*

8. *Compact Muon Solenoid*

spontanément en particules secondaires, des leptons, principalement des électrons, des jets de particules, etc... ; c'est ce qui permet de les identifier.

Le plan transverse de ce détecteur est défini par les axes  $x$  et  $y$ , de sorte que l'axe  $z$  soit colinéaire aux faisceaux. La symétrie du détecteur ATLAS permet d'introduire l'angle azimutal  $\phi$  et l'angle polaire  $\theta$ , même si nous préférons employer la pseudo-rapacité  $\eta$ <sup>9</sup>.

En partant du point d'interaction des deux faisceaux de protons, on distingue quatre sous-détecteurs [6] :

**Le détecteur interne ou trajectographe** (FIG.3) permet de déterminer la trajectoire des particules chargées et de reconstruire les vertex primaires - point d'interaction des deux faisceaux - et secondaires - point de désintégration des particules -. Il est lui même subdivisé en trois sous-détecteurs qui sont :

- Le détecteur à pixels, fait de 3 couches de pixels.
- Le **Semi-Conductor Tracking**, détecteur à micropistes, composé de 4 couches de micropistes.
- Le **Transition Radiation Tracker** (détecteur à tubes à pailles) permet l'identification des électrons par leur rayonnement cinématique. Dans ce détecteur chaque trace chargée est échantillonnée sur 36 points.

Enfin un aimant solénoïdal de 2T entourant le détecteur interne courbe la trajectoire des particules, permettant la mesure de l'impulsion et de la charge.

**Le calorimètre électromagnétique** (FIG.4) mesure avec précision la position et l'énergie des électrons et des photons. Il est divisé en trois compartiments :

- Le premier permet d'évaluer très précisément l'angle et la position du début de la gerbe.
- Le second, qui contient l'essentiel de la gerbe, sépare les gerbes électroniques de celles issues des particules du bruit de fond.
- Le dernier compartiment mesure les gerbes de très hautes énergies qui ne seraient pas entièrement contenues dans les deux premiers compartiments.

Ce calorimètre est composé d'un assemblage de plaques de plomb pliées en "accordéon" et séparées par des électrodes et de l'argon liquide.

**Le calorimètre hadronique** a été construit pour mesurer l'énergie déposée par des hadrons (protons, pions, kaons, ...) qui développent des gerbes lors de leur passage dans ce calorimètre. Il est composé d'absorbeurs en fer, séparés par des tuiles scintillantes en polystyrène.

**Le spectromètre à muons** (FIG.5) est capable d'identifier et de mesurer l'impulsion des muons, seules particules à atteindre cette partie du détecteur puisqu'elles interagissent très peu avec la matière. Il forme, avec les bobines supraconductrices qui l'entourent, la couche externe du détecteur.

### 2.3 Approche adoptée

Le stage porte sur la reconstruction et l'identification des électrons provenant du  $J/\Psi$  [7] (méson  $c\bar{c}$ , de masse  $3.09616 \text{ GeV}/c^2$ ), ce qui nécessite de comprendre la physique qui régit sa désintégration, d'apprendre comment fonctionne le détecteur et comment on détecte les électrons, pour ensuite créer un programme en C++/ROOT (logiciel permettant le calcul, la manipulation et la visualisation d'histogrammes couramment utilisé dans la communauté de la physique des particules) qui permettra d'automatiser cette identification.

L'approche adoptée sera donc principalement de la programmation pour l'analyse de données simulées. Le but du stage est avant tout d'avoir un aperçu du travail quotidien d'un chercheur, tout en tentant d'être utile en accomplissant un travail d'optimisation et d'amélioration.

---

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

## 3 Travail personnel réalisé

Tout au long du stage, nous nous sommes intéressés à l'étude des électrons provenant des désintégrations  $J/\Psi \rightarrow e^+e^-$ ,  $Z \rightarrow e^+e^-$  et  $MesonB \rightarrow e^+e^-$ .

### 3.1 Les données utilisées

Les données utilisées lors du stage ont été générées par une simulation Monte Carlo fondée sur le programme Pythia [8]. Cette simulation permet la production de particules issues d'une collision proton-proton à 7 TeV (type LHC) ainsi que les différentes désintégrations qui peuvent avoir lieu suite à celle-ci. De fait à la sortie de cette simulation un fichier est généré dans lequel est référencé l'ensemble des particules produites ainsi que leur quadri-impulsion. Après cette génération de données, le logiciel GEANT4 décrit la géométrie des détecteurs et simule les réponses que laisserait chaque particule dans ceux-ci. Enfin un logiciel de la collaboration ATLAS reconstruit, pour chaque détecteur, des particules à partir des informations issues de la simulation de l'appareil. Cela permet d'avoir accès à des traces - pour les particules chargées - et aux cellules dans les deux calorimètres contenant les dépôts d'énergie.

Nous avons donc à notre disposition plusieurs fichiers (un par simulation) contenant une liste de toutes les particules simulées, munies de leurs caractéristiques, intrinsèques et enregistrées par les détecteurs.

### 3.2 Premiers pas

Après avoir visualisé quelques données d'un fichier en utilisant ROOT [9], un programme en C++ a été écrit afin de pouvoir automatiser et optimiser la lecture de la liste des particules et de leurs caractéristiques. En effet, parmi toutes les variables propres à une particule, seule une trentaine d'entre elles nous intéressera pendant la durée de ce stage. Le premier objectif de ce programme est donc de lire les fichiers simulés et de tracer les histogrammes des variables sélectionnées. Nous pouvons donc déjà vérifier par exemple que la variable  $\phi$  est constante ; ce qui confirme que la distribution des particules est bien isotrope.

Comme nous utilisons des données simulées nous avons accès à la "vérité", soit la liste des particules générées. Il est donc possible d'associer une particule "vraie" à un objet reconstruit, cela permet par exemple de savoir directement si la particule sélectionnée est bien un électron provenant de la bonne particule mère ; c'est à ces électrons que nous ferons référence avec le terme "vrais électrons". Grâce à cette information, nous avons pu tracer des histogrammes des variables propres aux électrons, puis à ceux provenant de  $Z$  ou de  $J/\Psi$ .

### 3.3 Identification des électrons

Une partie conséquente du stage a été dédiée à la sélection de variables discriminantes pour pouvoir séparer les électrons provenant de désintégration des particules mères -  $J/\Psi$ ,  $Z$ ,  $MesonB$  - des autres particules pour lesquelles nous ne nous intéressons pas - hadrons, photons, électrons ne provenant pas de ces désintégrations - et qui composent le bruit de fond. Pour ce faire, la méthode la plus directe et la plus intuitive est de procéder à une série de coupures sur un certain nombre de variables que nous choisirons en fonction de leur efficacité à séparer les deux signaux.

Les premières des coupures à effectuer, et ce indépendamment des autres, sont une coupure sur l'énergie transverse,  $E_T$  et une sur la pseudo-rapacité. La première doit être supérieure à 20GeV dans le cas de  $Z \rightarrow e^+e^-$ , et supérieure à 3GeV dans les deux autres cas ( $J/\Psi$ ,  $MesonB$ ) ; dès lors énormément de particules lentes et indésirables sont supprimées. Quant à la deuxième, sa valeur absolue doit être inférieure à 2.5.

Ensuite on peut distinguer les variables discriminantes selon quatre catégories.

#### 3.3.1 Calorimètre hadronique

Le LHC a été construit de telle sorte que les électrons ne déposent qu'une faible fraction de leur énergie dans le calorimètre hadronique. C'est pourquoi on peut définir une variable décrivant si l'objet fuit dans le calorimètre hadronique comme étant le rapport (dans une fenêtre de granularité

de  $0,2 \times 0,2$ ) entre l'énergie transverse déposée dans le premier compartiment de ce calorimètre et celle déposée dans le calorimètre électromagnétique. Ce sera l'unique variable liée au calorimètre hadronique que nous utiliserons.

### 3.3.2 Calorimètre électromagnétique

Les électrons provenant d'une des désintégrations qui nous intéresse n'interagissent pas de la même façon avec la matière que les particules composant le bruit de fond, c'est pourquoi ces électrons ne déposent que peu d'énergie par cellules dans le calorimètre électromagnétique. Ils sont donc caractérisés par des gerbes bien plus fines que celle des particules issues du bruit de fond. Il apparaît essentiel d'exploiter ce fait physique pour séparer ces électrons des autres particules.

Les variables  $R_\eta$  et  $w\eta^2$  ont été utilisées dans ce but précis ; la première est définie comme étant le rapport entre l'énergie déposée dans une fenêtre de  $3 \times 7$  et  $7 \times 7$  cellules du second compartiment du calorimètre. C'est une des variables les plus discriminantes que nous avons à notre disposition, et elle s'avèrera très puissante par la suite. Quant à la seconde, il s'agit de la largeur latérale de la gerbe, calculée sur une fenêtre de  $3 \times 5$  cellules.

Nous nous servirons également de la variable  $E_{S1}/E_{tot}$  correspondant à la fraction d'énergie déposée dans le premier compartiment longitudinal du calorimètre, centrée autour de la cellule contenant la gerbe, divisée par l'énergie totale qui est définie comme la somme des énergies dans les trois compartiments.

La variable  $E_{S3}/E_{tot}$  - fraction d'énergie déposée dans  $3 \times 3$  cellules du dernier compartiment - sera elle aussi utilisée.

Enfin une dernière variable propre au calorimètre électromagnétique servira à faire des coupures discriminantes entre le signal (les électrons) et le bruit de fond, il s'agit de  $w_{stot}$  qui est la largeur totale de la gerbe.

### 3.3.3 Détecteur interne

Parmi les variables propres au détecteur interne, nous avons utilisé nPix et nSi, qui sont respectivement le nombre de coups le long de la trajectoire de la particule dans le détecteur à pixels et la somme des coups dans le détecteur à pixels, et dans le détecteur à micropistes. Grâce à ces deux variables nous pouvons garder les traces de bonnes qualités.

Une variable dépendant du TRT (détecteur à tubes à pailles) a également été utilisée, nous l'appellerons RatioTRT, et correspond au rapport entre le nombre de coups de haut seuils et le nombre total de coups dans le TRT.

### 3.3.4 Combinaison des deux sous-détecteurs

Enfin, la dernière catégorie de coupure lie des données du sous-détecteur et du calorimètre électromagnétique. Nous nous sommes servis de la variable  $E/p$  qui est définie comme le rapport entre l'énergie déposée dans le calorimètre électromagnétique et l'impulsion reconstruite dans le détecteur interne. Ce rapport est proche de 1 pour les électrons, leur masse des électrons étant négligeable.

Nous utiliserons aussi la variable  $\Delta\phi^2$  qui est la différence entre le  $\phi$  de la trace dans le détecteur interne et celle dans le second compartiment du calorimètre.

A titre d'illustration, les coupures appliquées pour sélectionner les électrons provenant du  $J/\Psi$  sont regroupées dans le tableau 2.

$R_\eta$	$w\eta^2$	$E_{S1}/E_{tot}$	$E_{S3}/E_{tot}$	$w_{stot}$	<i>hadleak</i>	<i>nPix</i>	<i>nSi</i>	<i>RatioTRT</i>	$E/p$	$\Delta\phi^2$
$> 0,65$	$< 0,015$	$> 0,1$	$< 0,25$	$< 5$	$< 0,1$	$> 2$	$> 7$	$> 0,07$	$< 2,5$	$< 0,05$

TABLE 2 – Valeurs des coupures appliquées sur les variables discriminantes dans l'algorithme d'identification des électrons provenant du  $J/\Psi$

Enfin, la figure 6 représentant les distributions de  $R_\eta$ , E/P, RatioTRT et des fuites hadroniques a été mise en annexe.

### 3.3.5 Efficacité et pureté des coupures

Une partie du stage a été dévolue à l'optimisation des coupures. Nous avons utilisé l'efficacité et la pureté d'une sélection pour estimer si une variable était suffisamment discriminante et ensuite quelle était la meilleure valeur de coupure à effectuer.

L'efficacité d'identification des électrons, que nous noterons  $\varepsilon_e$ , est définie comme  $\varepsilon_e = \frac{N_e^{cut}}{N_e^i}$  ; où  $N_e^{cut}$  est le nombre de "vrais électrons" identifiés après la coupure, et  $N_e^i$  le nombre initial de "vrais électrons". On peut également calculer l'incertitude, qui suit une statistique de POISSON :  $\sigma(\varepsilon_e) = \sqrt{\frac{\varepsilon_e(1-\varepsilon_e)}{N_e^i}}$ .

La pureté d'identification des électrons est déterminée comme  $P_e = \frac{N_e^{cut}}{N_{tot}^{cut}}$  ; où  $N_{tot}^{cut}$  est le nombre total de particules après la coupure. On associe à la pureté son incertitude :  $\sigma(P_e) = \sqrt{\frac{P_e(1-P_e)}{N_{tot}^{cut}}}$ . Le facteur de rejet des particules du bruit de fond sera également utilisé. Il est défini comme  $R_{bkg} = \frac{N_{bkg}^i}{N_{bkg}^{cut}}$  ; où  $N_{bkg}^i$  est le nombre initial de particules du bruit de fond, et  $N_{bkg}^{cut}$  le nombre de particules du bruit de fond après la coupure. Son incertitude statistique donnée par  $\sigma(R_{bkg}) = \frac{R_{bkg}}{\sqrt{N_{bkg}^{cut}}}$  lui sera associée.

Les résultats obtenus dans le cas des électrons provenant du  $J/\Psi$  et avec les coupures définies précédemment sont présentés dans le tableau 3. A titre indicatif, nous avons comparé nos coupures aux officielles, instaurées par le groupe ATLAS.

	$\varepsilon_e$ (en %)	$P_e$ (en %)	$R_{bkg}$
Coupure tableau 2	$57,1 \pm 0,2$	$98,6 \pm 0,1$	$31,4 \pm 0,5$
ATLAS Loose	$82,6 \pm 0,1$	$84,1 \pm 0,1$	$4,0 \pm 0,1$
ATLAS Medium	$76,0 \pm 0,1$	$96,0 \pm 0,1$	$9,0 \pm 0,1$
ATLAS Tight	$60,0 \pm 0,2$	$99,1 \pm 0,1$	$93,5 \pm 2,3$

TABLE 3 – *Efficacité et pureté d'identification des électrons avec les coupures définies dans le tableau 2 et les coupures officielles.*

Notre sélection "manuelle" est proche de la sélection dite "Tight" permettant une grande pureté, avec une efficacité  $\sim 60\%$ . Cependant le facteur de rejet du bruit de fond est bien plus faible.

Les données simulées utilisées dans la première partie du stage étaient pourvues de peu de bruit de fond ( $\sim 24\%$  de la totalité des particules analysées), néanmoins lors d'une désintégration  $J/\Psi \rightarrow e^+e^-$  dans le détecteur ATLAS il y a d'autres particules se désintègrent ; c'est pourquoi nous avons décidés de "rajouter" des hadrons dans les particules composant le bruit de fond en utilisant des données d'autres désintégrations. Même si les coupures ATLAS et notre coupure "manuelle" avaient permis précédemment d'obtenir de bons résultats, nous avons préféré utiliser des méthodes d'analyse statistique afin de séparer les électrons provenant du  $J/\Psi$  des particules composant le bruit de fond dans le cas ou celui-ci représente  $\sim 54\%$  des données.

## 3.4 Utilisation de TMVA

Le logiciel TMVA<sup>10</sup> [10] permet de prendre en compte la forme des distributions des variables discriminantes et de déterminer si elles sont corrélées. En utilisant ces données, le logiciel améliore l'identification des électrons grâce à plusieurs méthodes d'analyse statistique.

### 3.4.1 Les méthodes d'identification utilisées

Parmi les nombreuses méthodes d'identification présentent dans TMVA, nous en avons utilisé quatre.

<sup>10</sup> Toolkit for Multivariate Data Analysis



**Optimisation des coupures** La première méthode que nous avons utilisé est comparable au programme que nous avons écrit, elle permet de faire des coupures sur les variables en optimisant le rapport signal sur bruit. Cependant elle est bien plus puissante que notre programme car elle est automatique, et ne dépend donc pas de l'appréciation de l'utilisateur.

**Likelihood** Cette méthode d'analyse statistique est fondée sur un rapport de vraisemblance (*likelihood*) défini à partir de la distribution de densité de probabilité des variables discriminantes du signal et du bruit de fond. Cette méthode est optimale lorsque les variables ne sont pas corrélées.

**Boosted Decision Trees** La méthode BDT possède une architecture en forme d'arbre (FIG.7A). Une coupure est appliquée au niveau de chaque nœud permettant de séparer le signal du bruit de fond.

**Multi-Layer Perceptron** Le réseau de neurones (MLP) est une méthode d'analyse statistique non-linéaire qui présente une structure en couches (FIG.7B).

### 3.4.2 Comparaison des différentes méthodes

Les trois méthodes d'analyses statistiques utilisent des fonctions discriminantes que nous ne détaillerons pas dans ce rapport. Elles permettent de séparer au mieux les événements composant le signal de ceux composant le bruit de fond. Le figure 8 en annexe montre la distribution du likelihood.

Pour comparer les différentes méthodes d'analyses, nous avons utilisé des algorithmes déjà présents dans TMVA, qui tracent l'efficacité d'identification des électrons en fonction du facteur de rejet des particules composant le bruit de fond, défini par TMVA comme  $1 - \frac{1}{R_{bkg}}$  où  $R_{bkg}$  correspond au facteur de rejet utilisé précédemment. On observe très clairement que le likelihood est la méthode d'analyse statistique la plus performante dans notre cas ; cela vient sûrement du fait que les variables que nous utilisons sont très peu corrélées, rendant le likelihood optimal. Les performances de nos coupures sur ces données auraient plus faibles que celles des méthodes d'analyses statistiques puisque nous aurions traité les variables discriminantes individuellement, sans prendre en compte la forme de leur distribution ni leur possible corrélation.

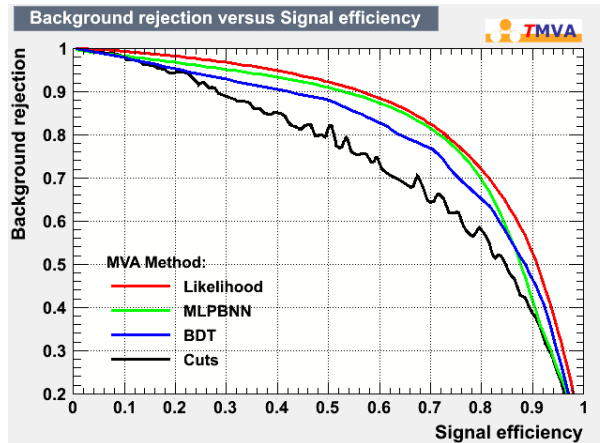


FIGURE 1 – Efficacité d'identification des électrons en fonction du facteur de rejet des particules composant le bruit de fond.

## 4 Conclusion et discussion

### 4.1 Conclusion

Le développement du programme en C++/ROOT a permis une identification rapide des électrons provenant du  $J/\Psi$  en effectuant des coupures sur des données provenant des différents composants du détecteur ATLAS. L'avantage de cette méthode réside dans le fait qu'elle soit facile à comprendre et qu'elle permet de se familiariser avec les différentes sous-parties du détecteur. Cependant, même si le programme obtient une sélection de particules pures avec une bonne efficacité, le facteur de rejet des particules du bruit de fond est assez faible, ce qui rend l'algorithme peu efficace si l'on veut par exemple reconstruire les particules mères des électrons.

C'est pourquoi nous avons utilisé un logiciel d'analyse statistique qui, en prenant en compte la forme des distributions ainsi que les corrélations entre les différentes variables, a grandement amélioré le facteur de rejet des particules non souhaitées. Une analyse plus précise dans les intervalles de pseudo-rapacité  $|\eta| < 0,8$  et  $|\eta| > 1,52$  devrait permettre d'améliorer un peu plus l'identification des électrons, et sera réalisée dans les prochains jours. Néanmoins certaines zones d'ombres persistent, notamment sur la compréhension des algorithmes que TMVA utilise et mériteraient qu'on s'y attarde, mais la durée du stage ne l'a pas permis.

### 4.2 Bilan personnel du stage

Ce stage de fin de master 1 a été une très bonne expérience dans le milieu de la recherche, milieu dans lequel j'espère évoluer plus tard. Pendant ces six semaines, je me suis investi dans un projet humain à grande échelle, aux enjeux scientifiques importants, et à la pointe de la recherche actuelle en physique des particules. Les connaissances que j'ai pu acquérir au cours de ce stage me permettront sans aucun doute d'aborder plus sereinement mon année de master 2 Noyaux Particules Astroparticules et Cosmologie. En effet, en plus d'avoir consolidé les bases de physique de particules qui nous ont été enseignées cette année avec le module Physique Nucléaire et des Particules, j'ai pu apprendre un nouveau langage de programmation, le C++, apprendre le Latex, découvrir ROOT mais aussi comprendre la structure du détecteur ATLAS et la façon de procéder pour pouvoir analyser des données.

Plus important encore, j'ai découvert ce qu'était la vie d'un laboratoire et d'un groupe de recherche ; travail d'équipe, collaborations internationales, réunions hebdomadaires, séminaires ...

En conclusion, ce stage m'a permis de conforter mon orientation vers la physique des particules en abordant ce domaine d'un point de vue expérimental et non plus seulement théorique comme ce fut le cas tout au long du master 1.

## Références

- [1] Laboratoire de Physique Nucléaire et de Hautes Énergies : *Rapport d'activité 2008-2009*, 2<sup>e</sup> trimestre 2011
- [2] Site Internet de l'Institut National de Physique Nucléaire et de Physique des Particules : <http://www.in2p3.fr>
- [3] L. VALENTIN Noyaux et particules : modèles et symétries. (1997)
- [4] L. EVANS et P. BRYANT : *LHC Machine*, Journal of Instrumentation, Volume 3 (2008)
- [5] Organisation européenne pour la recherche nucléaire. Site Internet : [www.cern.ch](http://www.cern.ch)
- [6] F. DERUE, Reconstruction et identification des électrons dans l'expérience Atlas. Participation à la mise en place d'un Tier 2 de la grille de calcul., Habilitation à diriger des recherches, Université Paris-Diderot, <http://tel.archives-ouvertes.fr/tel-00265729/fr/>
- [7] K NAKAMURA *et al* (Particle Data Group) 2010 *J. Phys. G : Nucl. Part. Phys.* **37** 075021
- [8] Site Internet de Pythia : <http://home.thep.lu.se/~torbjorn/Pythia.html>
- [9] Site Internet de ROOT : <http://root.cern.ch/>
- [10] A. HOECKER *et al* TMVA - toolkit for multivariate data analysis - <http://arxiv.org/abs/physics/0703039>

## 5 Annexe

### 5.1 Figure du dispositif expérimental

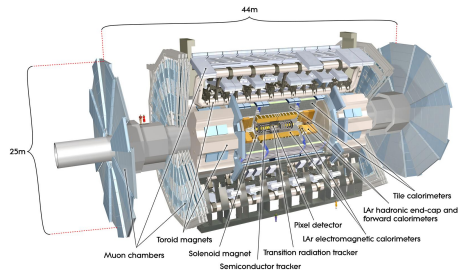


FIGURE 2 – Architecture du détecteur ATLAS.

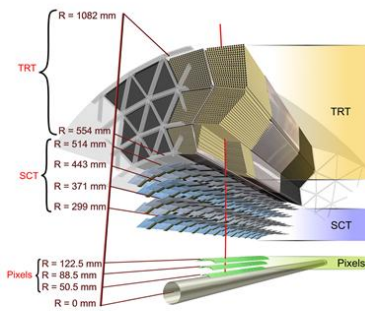


FIGURE 3 – Architecture du détecteur interne.

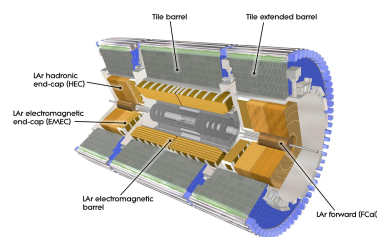


FIGURE 4 – Architecture du calorimètre électromagnétique.

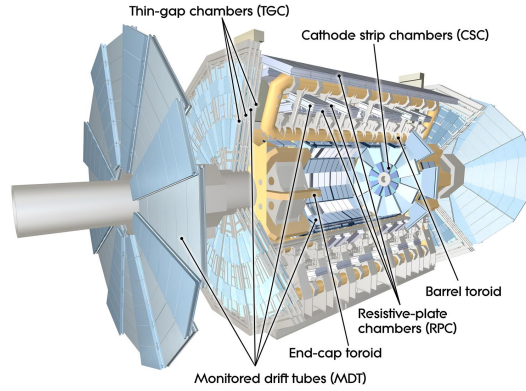


FIGURE 5 – Architecture du détecteur à muons.

## 5.2 Identification des électrons

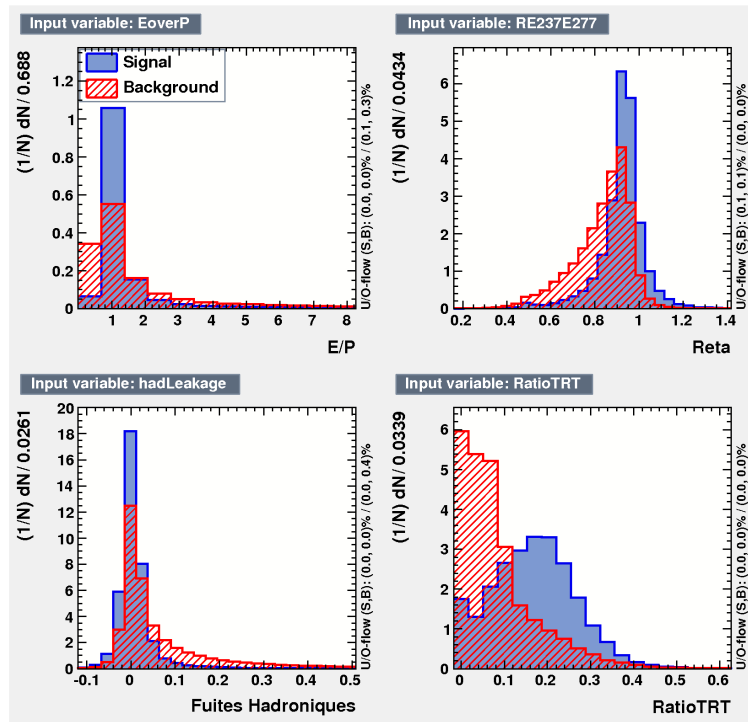


FIGURE 6 – Distributions de  $E/P$ , de  $R_\eta$ , des fuites hadroniques et du RatioTRT; en bleu les électrons, en rouge le bruit de fond.

### 5.3 Utilisation de TMVA

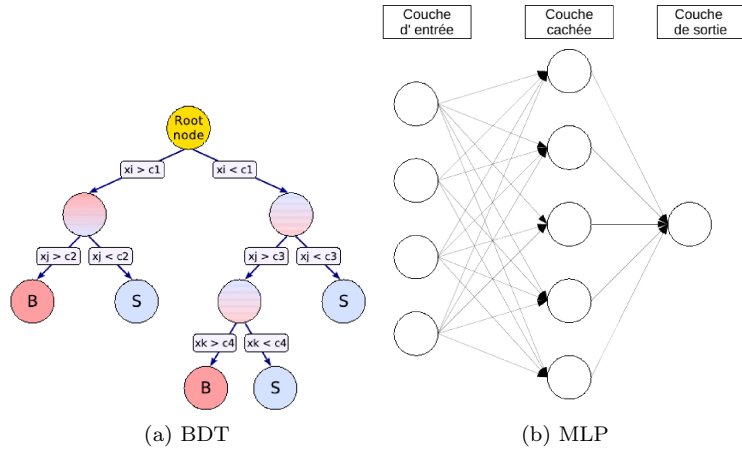


FIGURE 7 – Schéma d'un Boosted Decision Tree et architecture d'un réseau de neurones à une couche cachée.

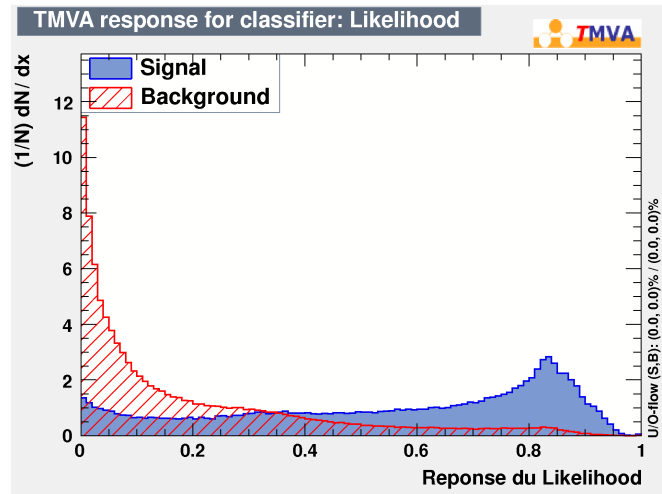


FIGURE 8 – Distributions des fonctions discriminantes du likelihood ; en bleu les électrons, en rouge le bruit de fond.