

SORBONNE UNIVERSITÉ

MASTER PHYSIQUE FONDAMENTALE ET APPLICATIONS

Rapport de stage de M1 : Reconstruction du boson Z^0 dans le
détecteur ATLAS au LHC

par :

Eva Vuilloz

N°étudiant : 28712737

Encadrant : Frédéric DERUE

Laboratoire de Physique Nucléaire et des Hautes Energies (LPNHE), Paris

19 juin 2024

Table des matières

Introduction	3
1 Contexte scientifique	4
1.1 Le Modèle Standard de la physique des particules	4
1.2 Le Large Hadron Collider et le détecteur ATLAS	5
2 Reconstruction du boson Z^0	6
3 Identification des électrons	8
3.1 Caractérisation et sélection des électrons	8
3.2 Utilisation du Machine Learning	10
Conclusion	13
Références	14

Introduction

Dans le cadre du M1 de Physique Fondamentale et Applications à Sorbonne Université, j'ai réalisé un stage de recherche de deux mois au Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE) de Paris, dans le groupe de recherche lié au détecteur ATLAS (A Torroïdal LHC ApparatuS) [1], sous la supervision de Frederic Derue. Ce stage a été l'occasion pour moi de découvrir le monde de la recherche en physique des particules, et de m'initier à des méthodologies d'analyse propres à ce domaine, en particulier l'étude de l'efficacité d'identification des électrons dans le détecteur. Le LPNHE est une unité mixte de recherche (UMR 7585) de l'Institut National de Physique Nucléaire et de Physique des Particules (IN2P3) du CNRS, de Sorbonne Université et de l'Université Paris-Cité. Il est situé sur le campus Pierre et Marie Curie sur le site de Jussieu. Ce laboratoire est engagé dans plusieurs grands programmes expérimentaux dans les domaines de la physique des particules, des astroparticules et de la cosmologie, notamment le détecteur ATLAS situé au LHC (Large Hadron Collider) [2], au CERN.[3] Le groupe auquel appartient Frederic Derue travaille sur le traitement et l'analyse des données du détecteur ATLAS, qui sera décrit plus en détail dans la suite de ce rapport. La recherche de Mr. Derue est plus particulièrement tournée vers le quark top, la particule la plus massive du Modèle Standard, et la détermination plus précise de sa masse, qui est l'un des paramètres du Modèle.

L'objectif de mon stage était de reconstruire la masse du boson Z^0 , l'une des particules du Modèle Standard responsables de l'interaction faible. Pour cela, j'ai utilisé à la fois des données de collisions enregistrées par le détecteur ATLAS et des données simulées. J'ai aussi effectué un travail sur l'identification des électrons, les produits de désintégration du boson Z^0 . Le but était d'étudier l'efficacité de reconstruction des électrons à l'aide des données simulées, et de comparer une identification des électrons basée sur une série de coupures faite "à la main" et une étude avec du Machine Learning.

La première partie de ce rapport est consacrée au Modèle Standard de la physique des particules et à la présentation du détecteur ATLAS, dont la compréhension est nécessaire pour cerner pleinement le travail que j'ai réalisé ensuite. La deuxième partie est dédiée aux résultats que j'ai obtenus pour la reconstruction du boson Z^0 , et enfin l'étude sur l'efficacité d'identification des électrons sera présentée dans la troisième partie.

1 Contexte scientifique

1.1 Le Modèle Standard de la physique des particules

Le Modèle Standard a été développé dans la deuxième moitié du XXème siècle pour décrire les constituants élémentaires de la matière et leurs interactions.[4] Il est constitué de 17 particules, réparties en 2 groupes : les fermions, de spin demi-entier, qui sont les constituants de la matière, et les bosons, de spin entier, qui sont les vecteurs des interactions. Il est considéré comme robuste du point de vue expérimental mais incomplet. En effet, entre autres choses, il ne décrit pas la gravité et ne permet pas d'expliquer la matière noire et l'énergie noire.

Les fermions sont séparés en deux catégories, les quarks et les leptons, qui sont chacun divisés en 3 familles ou générations. La matière ordinaire est composée d'éléments de la première génération (up et down qui forment les protons et neutrons, et électrons). Les deux autres familles sont composées d'éléments beaucoup plus lourds, donc avec une durée de vie bien plus faible et se désintègrent très rapidement. De plus, à chaque fermion est associée une antiparticule. Les bosons sont les vecteurs des interactions fondamentales. Il en existe trois types : le photon est le boson de l'électromagnétisme, le gluon celui de l'interaction forte, et les W^+ et W^- et le Z^0 sont les bosons de l'interaction faible. Il existe également un autre boson, le boson de Higgs, à l'origine de la masse des particules et qui a été découvert au LHC en 2012. Tous ces éléments sont résumés dans la Figure 1.

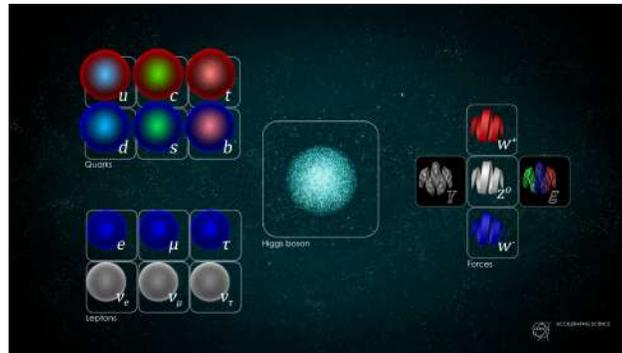


FIGURE 1 – Particules du Modèle standard (Image : Daniel Dominguez/CERN) [4]

Dans ce rapport, nous nous intéressons plus particulièrement au boson Z^0 . Sa masse est importante, d'environ $91,2 \text{ GeV}/c^2$. Sa durée de vie est très courte ($\approx 10^{-25}$)[5], on ne peut donc l'étudier qu'à travers ses produits de désintégration. Il se désintègre en plusieurs canaux : en paires d'électrons (électron + positron), en paires de muons (muon + antimuon) ou en deux quarks qui vont former deux paires de jets de particules. Celui qui va nous intéresser ici est le canal de désintégration en un électron et un positron, avec un rapport d'embranchement de $3.3632 \pm 0.0042 \%$ [5].

En relativité restreinte, la masse et l'énergie sont liées par la relation d'Einstein : $E^2 = (pc)^2 + (mc^2)^2$. Par conservation du quadri-vecteur énergie-impulsion lors de la désintégration, on peut donc écrire :

$$m_Z = \sqrt{\left(\frac{E_{e^-} + E_{e^+}}{c^2}\right)^2 - \left(\frac{p_{e^-} + p_{e^+}}{c}\right)^2} \tag{1.1}$$

Cette masse est la masse invariante du boson Z^0 , et cette formule nous permet de la retrouver à partir de la mesure de l'énergie et de l'impulsion des deux produits de sa désintégration.

1.2 Le Large Hadron Collider et le détecteur ATLAS

Le LHC est un accélérateur de particules circulaire, de 27 km de circonférence, situé près de Genève, en fonctionnement depuis 2009. Il a été construit pour étudier des collisions protons-protons et ions lourds. L'énergie des collisions a varié avec les années, et est actuellement de $\sqrt{s} = 13.6$ TeV. Quatre détecteurs sont placés sur l'anneau, à chaque point de collision des 2 faisceaux, dont ATLAS. L'objectif d'ATLAS est d'étudier l'ensemble des processus physiques en jeu lors des collisions protons-protons.[6] Il entoure le point de collision et est constitué de 4 sous-détecteurs différents comme montré sur la Figure 2. Au plus proche du point de collision, il y a le trajectographe qui permet de reconstruire les traces de particules chargées sans les détruire. Il est immergé dans un champ magnétique de 2 Tesla, qui va courber les trajectoires des particules ce qui permet d'en déduire leur charge et leur impulsion. Autour se trouve le calorimètre électromagnétique, qui mesure l'énergie et la position des électrons et des photons. Ce sont des particules qui vont produire des gerbes électromagnétiques, et vont donc déposer toute leur énergie dans le calorimètre. Il y a ensuite le calorimètre hadronique, qui a le même comportement avec les gerbes hadroniques créées par les hadrons. La couche externe est composée de chambres à muons, situées dans un champ magnétique toroidal - les muons sont des particules qui interagissent très peu avec les précédents détecteurs. On peut ainsi retrouver la nature d'une particule grâce à l'étude de son passage dans les différents sous-détecteurs, comme montré sur la Figure 2.

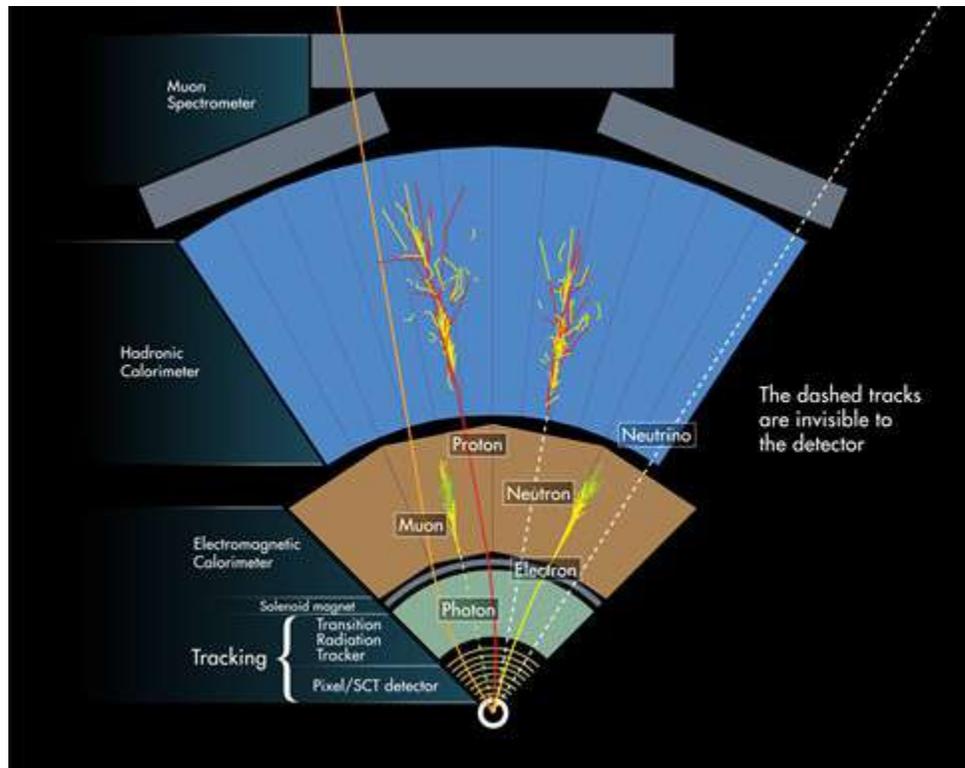


FIGURE 2 – Schéma en coupe des différents détecteurs d'ATLAS [7]

Les données utilisées dans ce rapport sont principalement celles du trajectographe et du calorimètre électromagnétique, puisque l'étude présentée est faite sur les électrons. Dans la suite de ce rapport, les électrons et positrons seront tous appelés électrons sans distinction de charges, leurs signatures dans les détecteurs étant les mêmes.

2 Reconstruction du boson Z^0

Pour la reconstruction de la masse invariante du boson Z^0 , j'ai utilisé deux types de données : des données de collisions enregistrées par le détecteur ATLAS, et des données simulées à la fois les processus de collisions et la description du détecteur, le tout par une méthode Monte-Carlo.[8] Les données de collision sont des données collectées par la collaboration ATLAS en 2018. Elles ont été reconstruites par un logiciel de la collaboration qui permet, à partir des signaux électriques enregistrés par les différents éléments du détecteur, de reconstruire des variables plus facilement exploitables, comme des traces dans le trajectographe ou des amas d'énergie dans les calorimètres. Ce programme donne une liste de candidats électrons, c'est à dire un objet associant une trace à un amas dans le calorimètre ; c'est à partir de ces données reconstruites que j'ai travaillé. Le deuxième type de données sont des données issues d'un générateur Monte-Carlo simulant des collisions protons-protons à 13 TeV produisant uniquement un $Z^0 \rightarrow e^+e^-$. Après cette génération, les listes de particules sont injectées dans une simulation du détecteur ATLAS permettant à la fois de construire un détecteur virtuel (représentant au mieux les sous-détecteurs, leurs matériaux, leur emplacement ...) et de simuler sa réponse au passage de chaque particule. De plus, on a accès à la "vérité", c'est à dire que l'on peut associer l'objet reconstruit à une particule "vraie" (simulée). J'ai ainsi pu savoir si un objet reconstruit correspondait à un électron "vrai" ou non, dans ce cas c'est un bruit de fond.

Pour la reconstruction, j'ai développé un code d'analyse en Python que j'ai tourné dans le cadre d'une plate-forme de calcul de type Jupyterhub.[9]

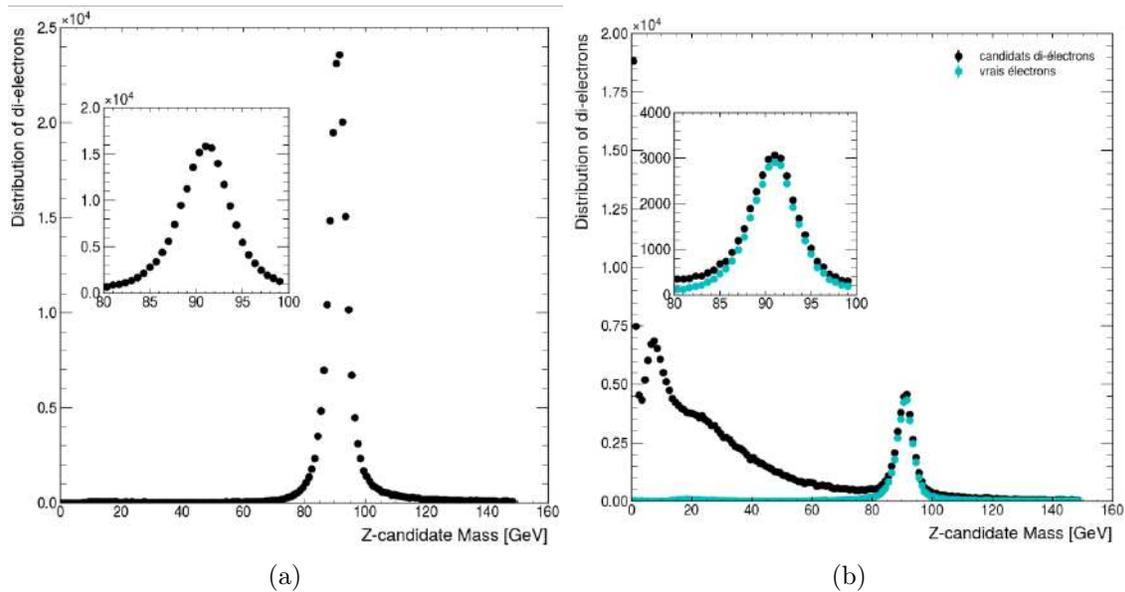


FIGURE 3 – Distributions de la masse invariante des paires d'électrons pour (a) des données de collision et (b) des données simulées de bosons Z^0 dans le canal de désintégration en 2 électrons. Pour les données simulées on montre les distributions pour tous les candidats (en noir) et pour les "vrais" électrons (en bleu). L'encart montre un zoom dans la fenêtre de masse 80-100 GeV.

La figure 3 montre les distributions de la masse invariante des paires d'électrons, de charges opposées, pour des données de collision et des données simulées de bosons Z^0 dans le canal de désintégration en 2 électrons. On voit sur les deux figures un pic autour de 90 GeV, c'est à dire à la masse attendue du boson Z^0 . Autour des pics, on voit un continuum correspondant à des événements de bruit de fond provenant de combinaisons

de paires d'électrons mais qui ne sont pas issus d'une même particule comme le Z^0 , ou lorsqu'au moins l'un des objets sélectionnés n'est pas un électron mais provient d'un hadron mal reconstruit et considéré comme un électron. Les données simulées permettent de savoir si l'objet reconstruit est un "vrai" électron ou d'un autre type. On voit sur la figure 4b la distribution pour les vrais électrons, qui est semblable aux données de collisions. La distribution pour tous les objets montre un bruit de fond important à des valeurs plus petites.

De même, la reconstruction de la masse invariante du Z^0 peut se faire avec d'autres canaux, comme la désintégration en deux muons, comme montré sur la Figure 4.

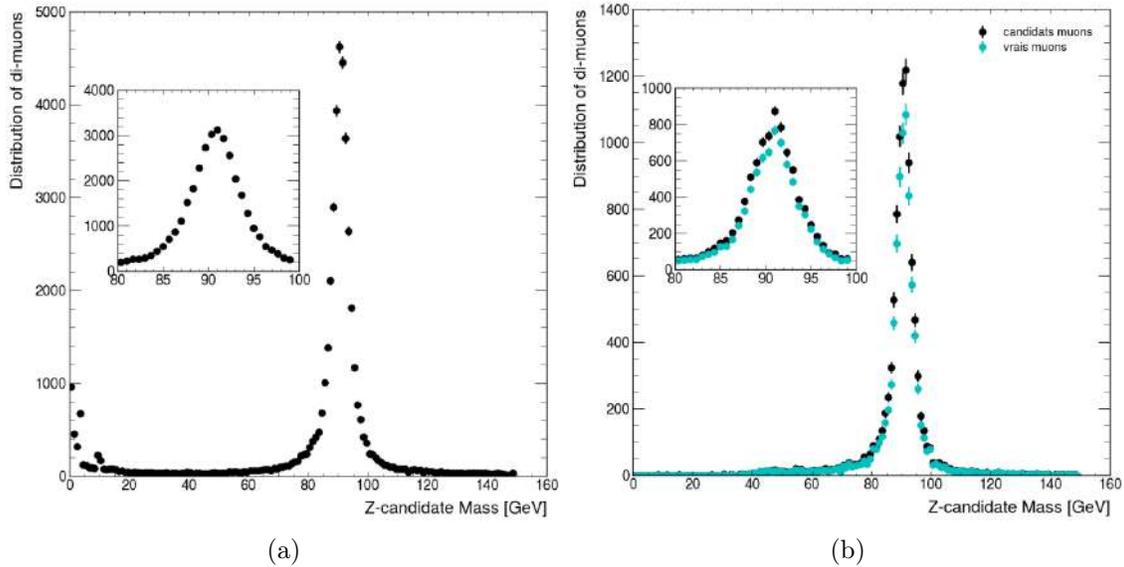


FIGURE 4 – Distributions de la masse invariante des paires de muons pour (a) des données de collision et (b) des données simulées de bosons Z dans le canal de désintégration en 2 muons. Pour les données simulées on montre les distributions pour tous les candidats (en noir) et pour les "vrais" muons (en bleu). L'encart montre un zoom dans la fenêtre de masse 80-100 GeV.

En plus du "pic de Z^0 " à 90 GeV, on distingue deux autres petits pics situés entre 0 et 10 GeV. Ces pics, comme on le voit sur la Figure 5 correspondent à deux autres particules qui peuvent également se désintégrer en deux muons, le J/ψ (état lié de quarks $c\bar{c}$) qui a une masse d'environ 3.09 GeV et le Υ (état lié de quarks $b\bar{b}$) qui a une masse d'environ 9.46 GeV. On ne voit pas ces particules dans la figure 3 car il est plus difficile de reconstruire des électrons que des muons.

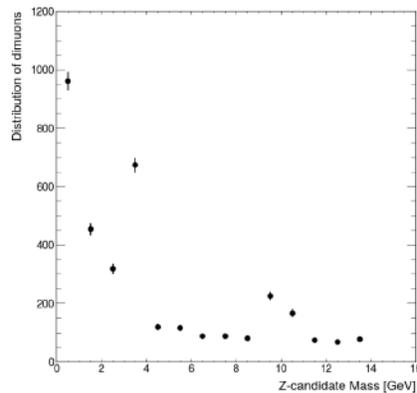


FIGURE 5 – Distribution de la masse invariante des paires de muons pour des valeurs inférieures à 15 GeV

3 Identification des électrons

Comme nous venons de le voir, tous les objets reconstruits comme des électrons n'en sont pas - ils associent simplement des traces dans les trajectographes à des amas d'énergie dans le calorimètre électromagnétique, et/ou ils ne proviennent pas de la désintégration d'un boson Z. Une étape supplémentaire à la reconstruction est nécessaire et est celle de l'identification des électrons. Comme j'ai pu l'étudier dans les premiers jours du stage avec l'inspection visuelle de plusieurs centaines d'événements, un électron se présente non seulement comme une association trace-amas, mais en plus cet amas est en général fin et étroit et n'atteint pas le calorimètre hadronique. Les données à disposition contiennent des informations qui permettent de caractériser le développement énergétique de l'amas (on parle de gerbe électromagnétique), en particulier le développement latéral (sa largeur) et longitudinal (sa profondeur dans le calorimètre). Pour cette étude nous utilisons les données de simulation, qui permettent pour chaque objet (candidat électron) de connaître son origine "vraie", donc s'il s'agit d'un vrai électron ou d'un hadron (bruit de fond). Nous allons donc utiliser les informations caractérisant les gerbes pour séparer les électrons des hadrons. Parmi ces informations, plusieurs variables vont se révéler particulièrement déterminantes pour la sélection, et seront donc préférées aux autres pour l'application des critères.

3.1 Caractérisation et sélection des électrons

Le but ici est donc de choisir un certain nombre de variables discriminantes, sur lesquelles appliquer des critères de sélection correspondant à des coupures rectangulaires sur les distributions pour ne garder que des "vrais" électrons.

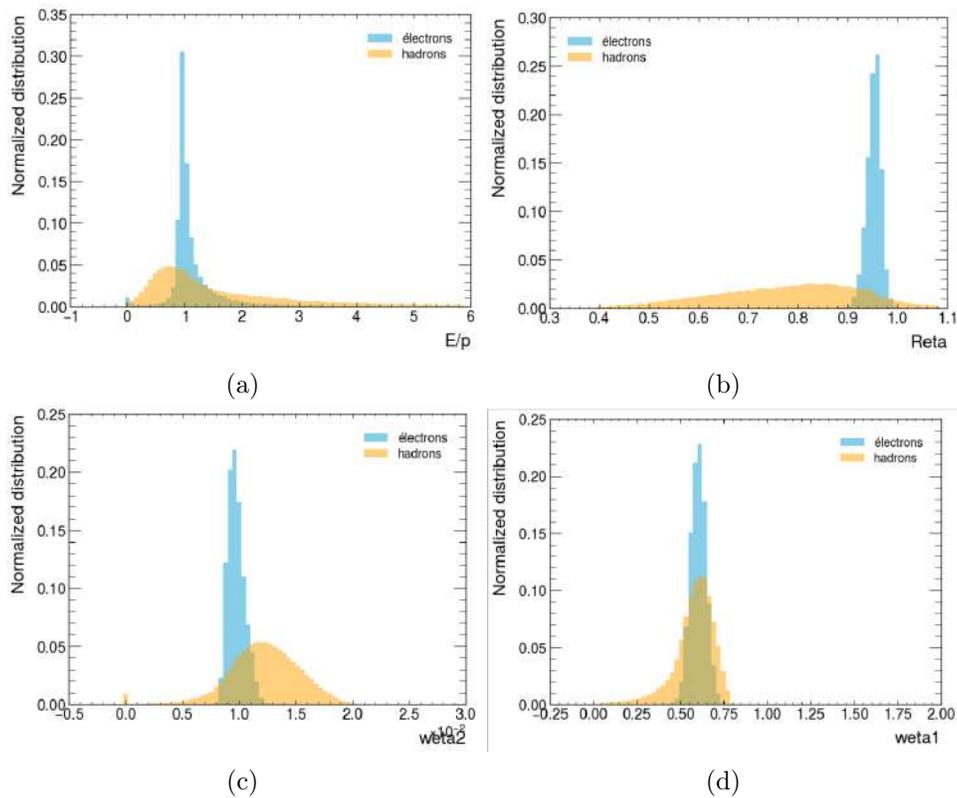


FIGURE 6 – Distributions de (a) E/p , (b) $Reta$, (c) $weta2$ et (d) $weta1$ pour les électrons et pour les hadrons

Sur la figure 6 on voit les distributions de 4 variables différentes, pour les vrais électrons et pour les hadrons.

Le rapport E/p est le rapport de l'énergie mesurée dans le calorimètre sur l'impulsion mesurée dans le trajectographe. Pour cette distribution on voit que les électrons piquent autour de 1, ce qui est attendu du fait de la très faible masse de l'électron. Des queues existent dans la distribution, à basse énergie par des effets de résolution et à plus haute énergie par des pertes sur la mesure de E par le rayonnement de freinage. La variable R_{eta} décrit le développement latéral de la gerbe dans le calorimètre. Les électrons ont une gerbe électromagnétique étroite et donc une valeur proche de 1, alors que les hadrons ont une gerbe beaucoup plus large. En effet, ceux-ci ne déposent qu'une partie de leur énergie dans le calorimètre électromagnétique et ont donc une distribution plus large. Les variables $\omega_{\eta 1}$ et $\omega_{\eta 2}$ sont les largeurs latérales des gerbes, respectivement dans les compartiments avant et milieu du calorimètre électromagnétique. En comparant les distributions des Figures 6d et 6c, on voit que $\omega_{\eta 2}$ est une meilleur variable discriminante que $\omega_{\eta 1}$, puisque pour celle-ci les pics des électrons et des hadrons sont confondus. Une coupure sur cette variable ne permettra donc pas une réduction du bruit de fond significative.

En tout j'ai utilisé 5 variables sur lesquelles j'ai déterminé des critères de sélection et d'identification des électrons. En plus de E/p , R_{eta} et $\omega_{\eta 2}$, j'ai également utilisé les fuites hadroniques caractérisées par le rapport R_{had} et la largeur totale de la gerbe ω_{tot1} dans le premier compartiment du calorimètre électromagnétique. Les gerbes provenant d'électrons ne vont quasiment pas déposer d'énergie dans le calorimètre hadronique, car la quasi-totalité de leur énergie aura été déposée avant dans le calorimètre électromagnétique. La distribution des électrons pour R_{had} sera donc très piquée autour de 0, tandis qu'elle sera beaucoup plus large pour les hadrons qui eux vont déposer leur énergie principalement dans le calorimètre hadronique.

Les coupures appliquées sont présentées dans le Tableau 1.

Variable	Sélection
E/P	0 - 7
R_{η}	0.85 - 1.05
R_{had}	-0.04 - 0.04
$w_{\eta 2}$	0.008 - 0.012
w_{tots1}	0.8 - 3.5

TABLE 1 – Critères de sélection appliqués sur les données simulées

Nous pouvons estimer l'impact de ces critères sur les données à l'aide de l'efficacité d'identification des électrons et du facteur de rejet des hadrons :

$$\epsilon_e = \frac{N_{el}^{selec}}{N_{el}^{vrai}} \quad \text{et} \quad r_{rejet} = \frac{1}{\epsilon_{hadrons}} = \frac{N_{hadrons}}{N_{hadrons}^{selec}}$$

Avec les coupures effectuées dans le Tableau 1, nous obtenons une efficacité de $\epsilon_e = 0.96$ et un facteur de rejet d'environ 30. Nous pouvons donc conclure que la sélection effectuée a une efficacité très haute (proche de 1) tout en permettant d'éliminer une grande partie du bruit de fond ($\approx 97\%$). On notera que la variable E/p est en pratique peu utilisée. Couper sur cette variable abaisse assez rapidement l'efficacité. Pour mieux faire, nous allons maintenant étudier des sélections faites avec du Machine Learning.

3.2 Utilisation du Machine Learning

L'étude faite ici est un exemple de problème de classification. Pour cette partie, j'ai utilisé le package scikit-learn [10] en Python, qui possède différents modules d'apprentissage. Toutes les méthodes d'apprentissage fonctionnent sur le même principe. Le lot de données est séparé en deux parties, un échantillon d'entraînement et un échantillon de test. En fournissant au module le premier échantillon et la composition en vrais électrons et en bruit de fond, le module va s'entraîner sur toutes les variables données pour déterminer des critères définissant des électrons. Puis l'échantillon de test est utilisé avec le module, cette fois de composition inconnue, et c'est le module qui va séparer à l'aide de ses propres critères les électrons des autres particules. Différents types de modules, d'algorithmes, ont ainsi été testés afin de comparer leurs performances.

La méthode la plus simple est d'utiliser l'ensemble de l'information de chaque distribution en particulier sa forme qui peut être différente pour le signal (électron) et le bruit de fond. Chaque distribution est normalisée à l'unité et peut ainsi être traitée comme une densité de probabilité, d'être du signal ou du bruit de fond. En utilisant plusieurs variables discriminantes, on peut multiplier ainsi les densités de probabilités de chaque distribution pour n'obtenir en fin de course qu'une seule variable discriminante indiquant la probabilité d'être du signal ou du bruit de fond. C'est le principe des méthodes de vraisemblance ou likelihood qui sont implémentées dans scikit sous le nom de GNB.

Cette méthode fonctionne bien si les distributions de signal et de bruit de fond sont bien séparées selon l'axe x des distributions obtenues en 1D (comme sur les figures 6a, 6b et 6c). Si ce n'est pas le cas, il faut trouver un axe de séparation dans l'espace des données d'entrée de manière à ce que la projection de cet axe dans l'espace des données de sortie sépare au maximum le signal du bruit de fond. C'est la méthode du Fisher's Linear Discriminant (LDA).[11] Ces méthodes fonctionnent bien si les variables ne sont pas corrélées ou si la corrélation est linéaire. Pour des cas plus complexes de corrélation alors il faut utiliser des méthodes plus complexes comme les réseaux de neurones. Par manque de temps, ces méthodes n'ont pas pu être testées.

Enfin des méthodes simples sur le principe, mais complexes dans la pratique, sont des arbres décision.[12] Ceux-ci correspondent à des coupures effectuées sur chaque variable, comme nous l'avons fait à la main, mais un arbre de décision va diviser les données petit à petit. A chaque division, une question sur les variables sera posée, et une nouvelle division sera faite selon la réponse à la question, le but final étant de séparer entièrement le signal du bruit de fond. Il existe différents modèles d'arbres de décision, et plusieurs ont été testés ici. Les modules utilisés sont implémentés dans scikit sous les noms XGB, LightGBM, GBDT et AdaBoost.

Nous pouvons comparer leurs capacités selon les distributions obtenues en sortie des différents modules, comme sur la Figure 7.

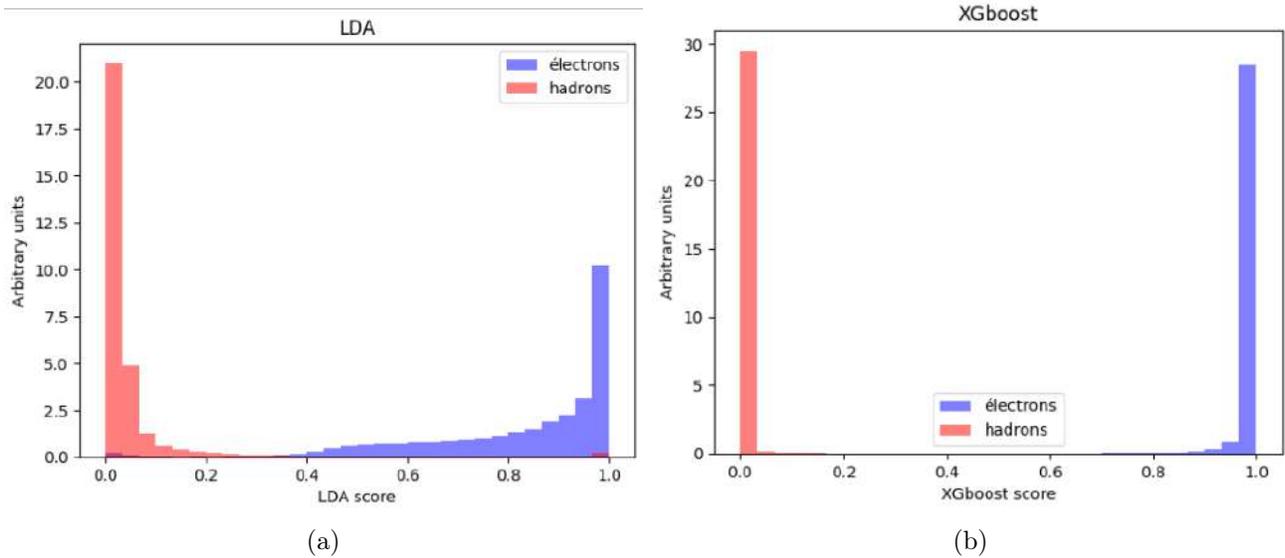


FIGURE 7 – Comparaison des distributions de sortie de deux méthodes de classification, en a) en sortie du LDA et b) en sortie d'un arbre de décision, ici le module XGB

Sur la figure 7, les distributions de sortie avec les données de l'échantillon d'entraînement sont représentés. En comparant les figures 7a et 7b, on voit que les distributions des électrons et des hadrons sont parfaitement séparées sur la figure 7b, alors que à la sortie du LDA les distributions sont plus diffuses et la séparation est moins évidente. La méthode des arbres de décision semble ainsi plus performante que la méthode du Fisher's Linear Discriminant.

Nous pouvons également comparer ces méthodes en traçant une courbe ROC, qui nous permettra également de comparer les méthodes de Machine Learning aux critères de sélection appliqués dans la partie précédente. Une courbe ROC (Receiver Operating Characteristics) représente la capacité à éliminer le bruit de fond en fonction de l'efficacité du signal. On peut ainsi lire sur la courbe la performance de l'outil utilisé, pour une efficacité ou un facteur de rejet donné. L'AUC (Area Under the Curve) est l'aire sous la courbe ROC et donne la performance globale du classificateur. En effet, lorsque la courbe ROC est haute, cela signifie que l'efficacité et le facteur de rejet sont très hauts. Une augmentation de l'AUC indique donc une amélioration des capacités discriminatoires, avec un maximum de 1 (100%).

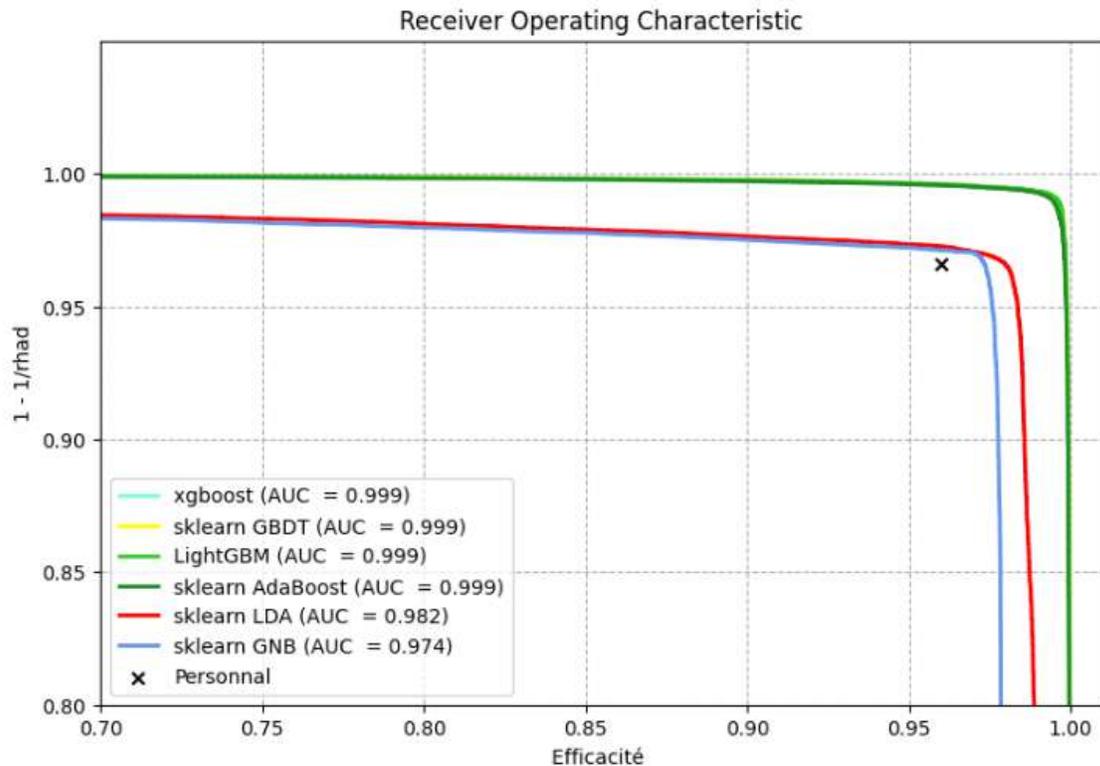


FIGURE 8 – Courbe ROC pour les méthodes de vraisemblance, de LDA et d’arbres de décision. La sélection effectuée dans la partie 3.1 est également tracée.

Sur la figure 8, les performances des différents modèles sont représentés avec pour chacun la relation entre l’efficacité et le facteur de rejet. Nous pouvons constater que la méthode de vraisemblance (GNB) est la moins performante, tandis que tous les modules implémentant une méthode d’arbre de décision se confondent et sont les plus performants. Nous pouvons également les comparer avec leur AUC, avec une AUC de 0.999 pour les arbre de décision qui est supérieure à l’AUC de la méthode de vraisemblance (0.974). Parmi les modules d’arbre de décision, on observe une vraie différence au niveau du temps d’entraînement entre AdaBoost, qui dure 245 secondes, et les trois autres (XGBoost, GBDT et LightGBM) qui durent à peine 1 seconde (respectivement 0.65, 0.78 et 1 seconde). Il est attendu que le module le plus performant soit le XGBoost, c’est le plus récent et le plus utilisé, mais si ce n’est sur le léger décalage en temps d’entraînement sa performance ne se démarque pas vraiment ici.

Nous pouvons également voir que malgré leurs différences de performance toutes les méthodes de Machine Learning restent meilleures que la sélection faite "à la main" à l’aide de coupures dans la section 3.1. Cette constatation n’est pas surprenante, nous nous attendions à ce que la sélection faite à l’aide du Machine Learning soit plus efficace que la sélection à l’aide de critères basés sur des coupures rectangulaires sur les distributions.

Conclusion

Cette immersion de 8 semaines au LPNHE m'a permis d'avoir une première impression du travail de chercheur, ce qui était très important pour moi car mon objectif est de continuer un parcours universitaire après le Master, en particulier dans le domaine de la physique des particules et/ou de la cosmologie. Le travail présenté est celui d'une analyse classique du domaine, permettant dans un temps court d'aborder la production des particules sur accélérateurs et la reconstruction de particules aussi fondamentales qu'un boson Z^0 , qui porte l'interaction faible. Le but était ainsi de se familiariser avec une méthodologie de recherche propre à la physique des particules.

La reconstruction de particules, en particulier à l'aide de leur masse invariante avait été abordée de manière théorique en cours de physique des particules, car c'est justement la méthode utilisée pour la découverte de nouvelles particules. En la mettant en application pendant ce stage, j'ai pu me rendre compte de l'importance de cette méthode, mais également de la complexité derrière le mécanisme de sélection des produits de désintégration étudiés, malgré l'aspect assez simple au premier abord de cette méthode telle qu'abordée dans les tous premiers jours du stage lors de l'inspection visuelle d'événements.

L'identification des électrons a été la première expérience du Machine Learning que j'ai eu. Le Machine Learning est de plus en plus utilisé en recherche. C'est pourquoi améliorer ce genre de compétences me semble crucial pour continuer mon parcours en recherche en physique, et encore plus particulièrement en apprendre plus sur les applications du Machine Learning en physique des particules.

De manière plus générale, j'ai pu découvrir l'environnement de travail au sein d'un laboratoire de recherche, en participant aux différentes réunions du groupe ATLAS et de l'entièreté du laboratoire. J'ai pu également constater l'échange permanent des chercheurs autour des diverses avancées dans le domaine de la physique des particules et des hautes énergies, et de la cosmologie, en assistant aux nombreux séminaires et présentations des doctorants organisés par le laboratoire.

Je souhaite ici remercier mon maître de stage, Mr Frédéric Derue, pour son encadrement et sa disponibilité tout au long de ce stage, et également l'équipe du LPNHE de m'avoir accueillie avec gentillesse au sein du laboratoire.

Références

- [1] The ATLAS experiment, Monica Lynn Dunford and Peter Jenni (2014), Scholarpedia, 9(10) :32147, http://www.scholarpedia.org/article/The_ATLAS_experiment
- [2] Large Hadron Collider, site web de LHC France, <https://www.lhc-france.fr/>
- [3] CERN, Organisation Européenne pour la Recherche Nucléaire, <https://home.cern/fr>
- [4] CERN. «Le Modèle standard», 29 avril 2024. <https://home.cern/fr/science/physics/standard-model>.
- [5] Particle Data Group. «Particle Data Group». Consulté le 6 juin 2024. <https://pdg.lbl.gov/2024/listings/contents-listings.html>.
- [6] Theveneaux-Pelzer, Timothée. «Etudes sur la reconstruction des électrons et mesure de la section efficace de production de paires de quarks top dans les canaux dileptoniques dans l'expérience ATLAS auprès de LHC». Thèse de doctorat, Université Pierre et Marie Curie, 2012.
- [7] Pequeno, Joao, et Paul Schaffner. «How ATLAS detects particles : diagram of particle paths in the detector», 2013. <https://cds.cern.ch/record/1505342>.
- [8] ATLAS open data, <https://atlas.cern/Resources/Opendata>
- [9] JupyterHub, <https://jupyter.org/hub>
- [10] scikit-learn, Machine Learning in Python, <https://scikit-learn.org>
- [11] Ambika. «Linear Discriminant Analysis (LDA) in Machine Learning : Example, Concept and Applications». (blog), 13 septembre 2023. <https://medium.com/aimonks/linear-discriminant-analysis-lda-in-machine-learning-example-concept-and-applications-37f27e7c7e98>.
- [12] Gaurav. «An Introduction to Gradient Boosting Decision Trees». Machine Learning Plus (blog), 12 juin 2021. <https://www.machinelearningplus.com/machine-learning/an-introduction-to-gradient-boosting-decision-trees/>.