

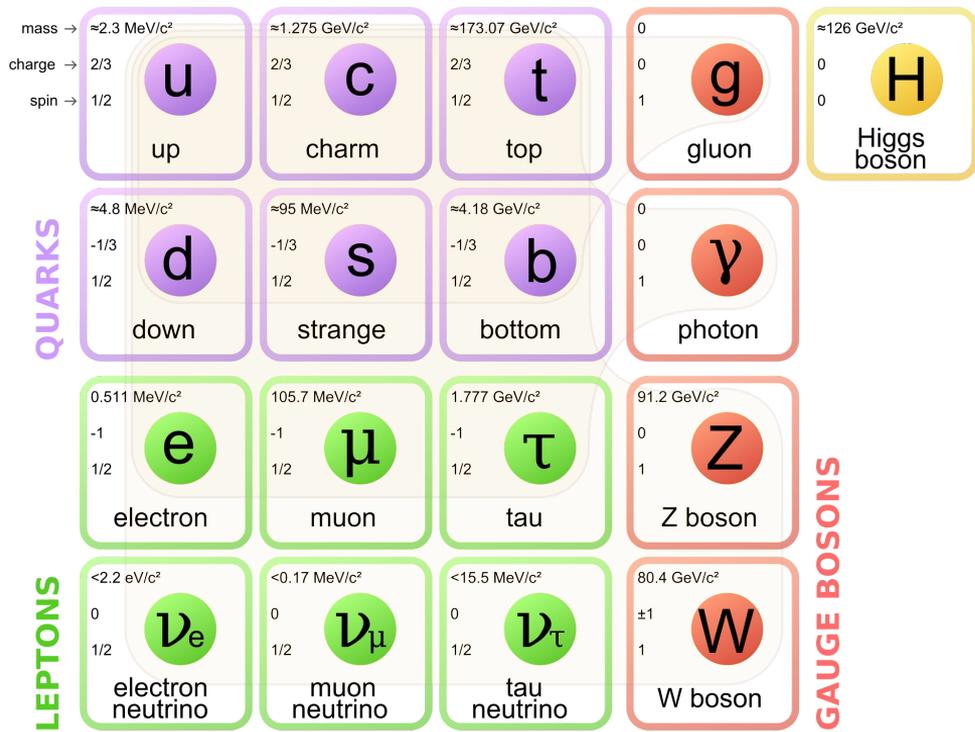
Reconstruction du boson Z^0 dans l'expérience ATLAS et son utilisation comme chandelle étalon



Guillem Joseph Planas – M1 PFA

Encadrant: Frédéric Derue
Laboratoire de Physique Nucléaire
et des Hautes Énergies (LPNHE)
Paris

Le Modèle Standard de la physique des particules



Décrit les constituants de la matière et leurs interactions:

Fermions (spin $1/2$), composent la matière:

- Séparés en quarks et leptons
- 3 générations

Bosons (spin 1), vecteurs des interactions:

- Gluon (Forte)
- Photon (EM)
- W^+ , W^- , Z^0 (Faible)

+ Boson de Higgs (spin 0) qui explique l'origine de la masse des particules

Robuste expérimentalement, il reste incomplet → besoin d'accélérateurs

Le Large Hadron Collider (LHC)

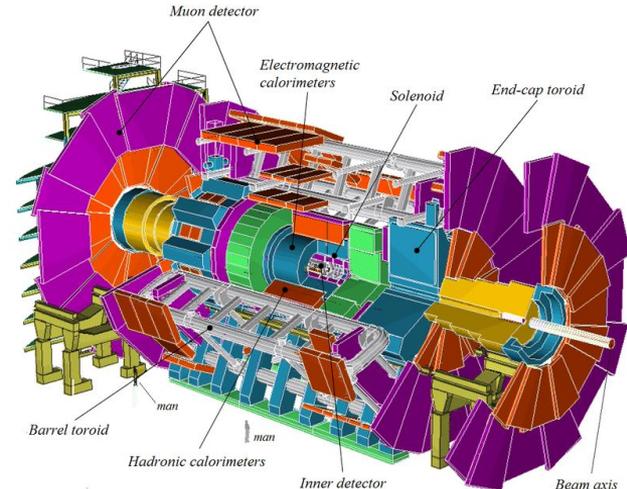
Collisionneur proton-proton, étude du Modèle Standard et au-delà

- Situé au CERN, en activité depuis 2008
- 27 km de circonférence, 100 m sous-terre
- Collisions avec une énergie dans le centre de masse de $\sqrt{s} = 13.6$ TeV
- Détecteurs situés à chaque point d'interaction



Le détecteur ATLAS

Expérience située à Genève, étude généraliste du Modèle Standard

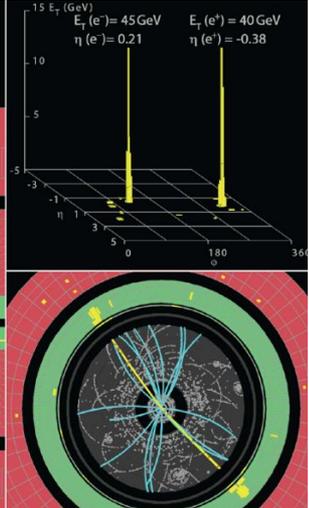
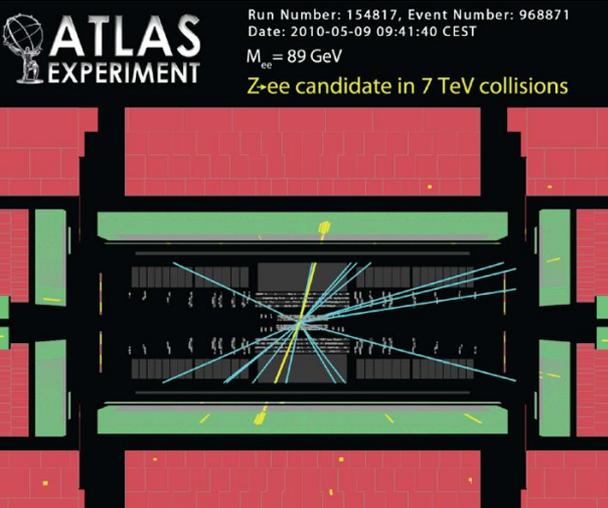
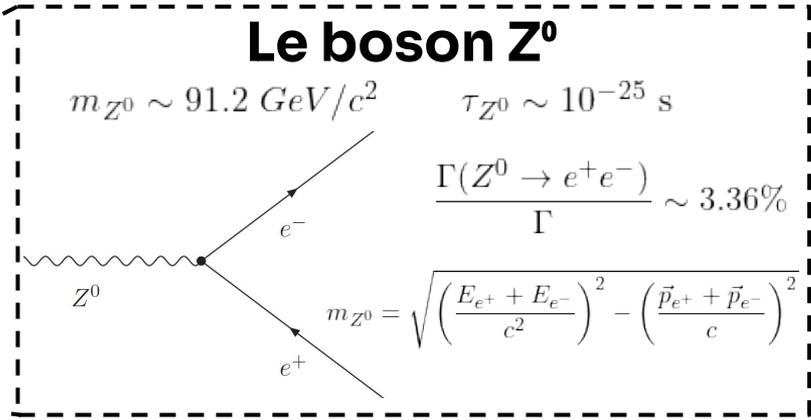


Possède une structure en couches:

- Trajectographe: mesure les trajectoires et les impulsions des particules chargées
- Calorimètres: mesurent l'énergie et la position des particules (électrons, photons, hadrons) grâce à la formation de gerbes
- Chambres à muons: Détectent les muons, capables de traverser toutes les couches précédentes

Identification et caractérisation d'un électron → données issues du trajectographe + calorimètre électromagnétique

mass → charge → spin →	$\approx 2.3 \text{ MeV}/c^2$ 2/3 1/2 u up	$\approx 1.275 \text{ GeV}/c^2$ 2/3 1/2 c charm	$\approx 173.07 \text{ GeV}/c^2$ 2/3 1/2 t top	0 1 g gluon	$\approx 126 \text{ GeV}/c^2$ 0 0 H Higgs boson
QUARKS	$\approx 4.8 \text{ MeV}/c^2$ -1/3 1/2 d down	$\approx 95 \text{ MeV}/c^2$ -1/3 1/2 s strange	$\approx 4.18 \text{ GeV}/c^2$ -1/3 1/2 b bottom	0 1 γ photon	
LEPTONS	$0.511 \text{ MeV}/c^2$ -1 1/2 e electron	$105.7 \text{ MeV}/c^2$ -1 1/2 μ muon	$1.777 \text{ GeV}/c^2$ -1 1/2 τ tau	$91.2 \text{ GeV}/c^2$ 0 1 Z Z boson	
	$< 2.2 \text{ eV}/c^2$ 0 1/2 ν_e electron neutrino	$< 0.17 \text{ MeV}/c^2$ 0 1/2 ν_μ muon neutrino	$< 15.5 \text{ MeV}/c^2$ 0 1/2 ν_τ tau neutrino	$80.4 \text{ GeV}/c^2$ ± 1 1 W W boson	GAUGE BOSONS



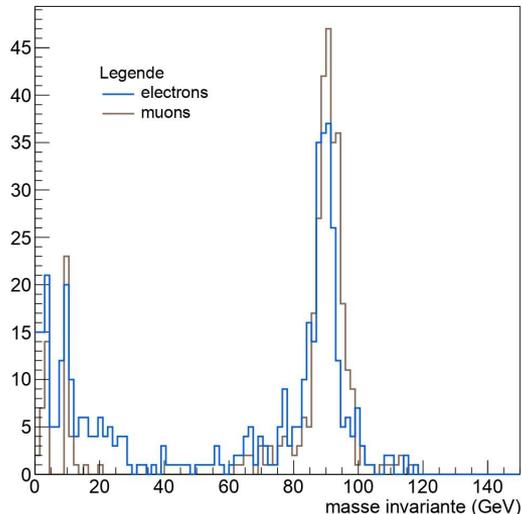
Canal $Z^0 \rightarrow e^+e^-$ dans ATLAS

Le LHC produit des bosons Z^0 en grand nombre

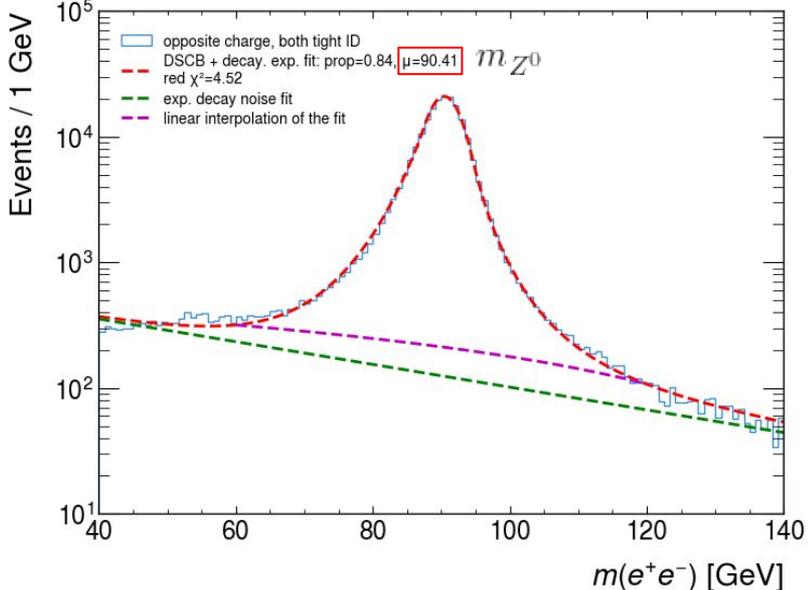
- Peuvent être reconstruits facilement dans sa désintégration en paires d'électrons, muons ou photons

→ peuvent alors être utilisés comme une "chandelle étalon" pour mesurer des efficacités d'identification des électrons ou pour calibrer leur énergie

Reconstruction du boson Z^0



- Masses reconstruites des candidats di-électrons dans des données de collisions dans ATLAS lors du Run 2
- Pics aux masses de certaines résonances (Z^0 J/Ψ Υ)



Ajustement des courbes

- Estimation précise de la masse du Z^0 , nombre d'événements de signal et de bruit de fond
- Utilisation d'une fonction incluant le signal (Double Sided Crystal Ball) et le bruit (exponentielle décroissante)
- Interpolation linéaire du bruit de fond sous le pic du Z^0
- Estimation de la masse: proche des valeurs connues

Méthode de tag & probe

- Sélection de 2 électrons:
 - 1 tight (1e tag), bien identifié
 - 1 probe, pas encore identifié
- Le pic du Z^0 est très bien reconstruit avec peu de bruit de fond
- La zone autour du Z^0 (± 10 GeV) donne un lot d'électrons probe avec lequel on peut mesurer l'efficacité d'identification des électrons
- L'efficacité d'identification des électrons est peu sensible à l'estimation du (petit) bruit de fond

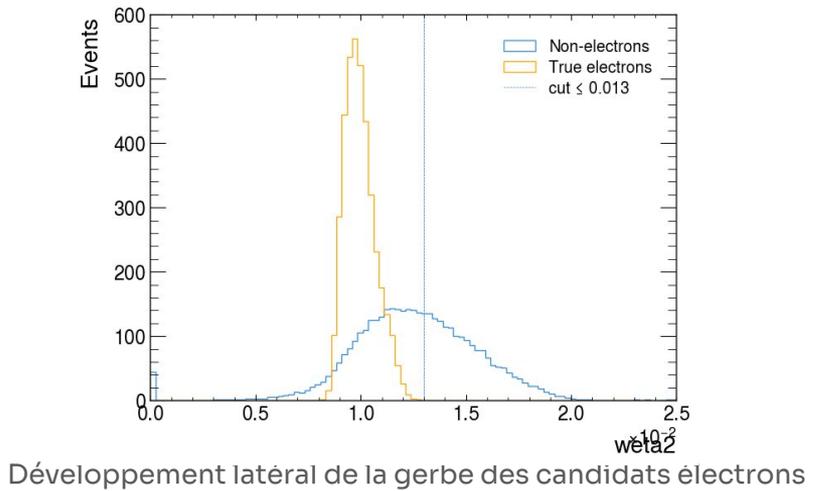
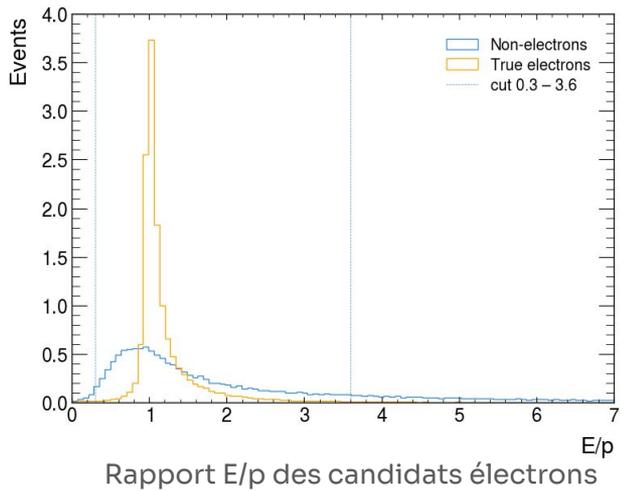
Données simulées par Monte-Carlo

Échantillon Monte Carlo: collisions proton-proton à $\sqrt{s} = 13 \text{ TeV}$, donnant uniquement $Z^0 \rightarrow e^+e^-$

- Événements passés dans une simulation complète (géométrie, matériaux, sous-détecteurs) d'ATLAS
→ Simulation de la réponse du détecteur au passage des particules
- Accès à l'information "vérité": nature réelle de chaque particule simulée
- Accès aux variables caractérisant les particules et leurs dépôts d'énergie

Identification des électrons

Relier une trace dans la trajectographe à un amas dans le calorimètre. Les caractéristiques de cet amas servent à discriminer les vrais électrons des faux



Évaluation des sélections

Des coupures rectangulaires ont été effectuées sur 6 variables

- Une coupure de reconstruction sur l'impulsion transverse
- 5 coupures d'identification sur des variables discriminantes

Coupure
$p_T \geq 5 \text{ GeV}$
$E/p \in [0.3, 3.6]$
$R_{had} \leq 0.04$
$R_\eta \geq 0.9$
$\omega_{\eta 2} \leq 0.013$
$\omega_{tot1} \leq 3.2$

La qualité de la sélection est déterminée par les grandeurs suivantes:

- Efficacité d'identification des électrons

$$\epsilon_e = \frac{N_e^{\text{selec}}}{N_e^{\text{tot}}}$$

- Facteur de rejet du bruit de fond

$$r_b = \frac{1}{\epsilon_b} = \frac{N_b^{\text{tot}}}{N_b^{\text{selec}}}$$

Résultat
$\epsilon_e = 96.8 \pm 0.1\%$
$r_b = 23 \pm 0.5$
$1 - \epsilon_b = 95.6 \pm 0.1\%$

Les coupures choisies permettent de garder la majorité des vrais électrons en éliminant presque tout le bruit de fond

Identification des électrons avec du Machine Learning

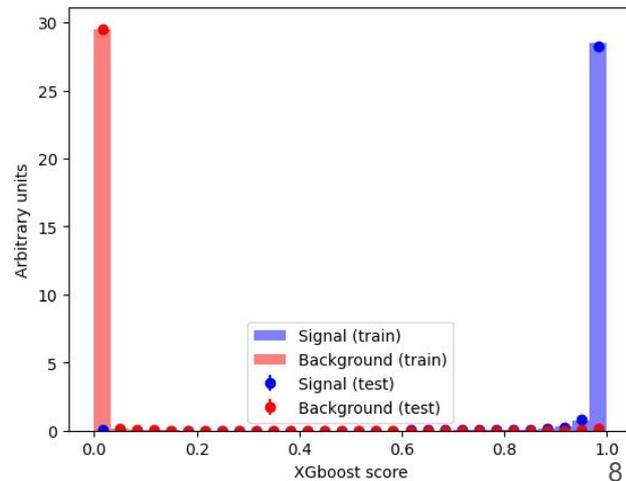
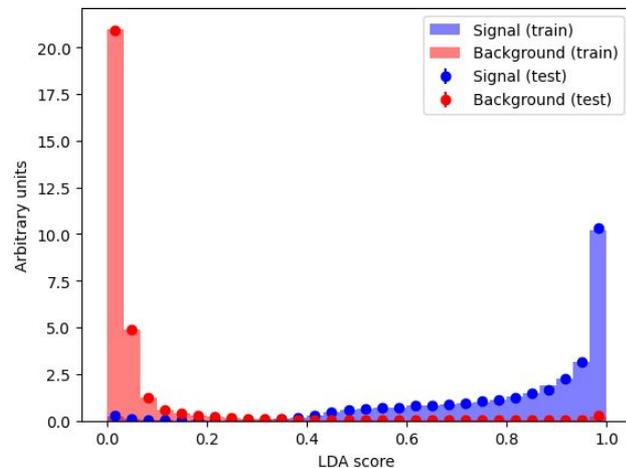
Tâche de classification, qui est une forme d'apprentissage supervisé → les algorithmes s'entraînent et sont testés sur 2 lots différents

- **Entraînement**: L'algorithme construit une variable dont la valeur est très différente entre les candidats (leur statut de vérité lui est connu), en se basant sur les variables discriminantes

- **Test**: L'algorithme essaye de classer des candidats de classe inconnue

Utilisation de 3 méthodes différentes de la bibliothèque Scikit-learn:

- **Likelihood** (Gaussian Naive Bayes):
 - Hypothèse de distribution gaussienne et indépendance
- **Discriminant de Fisher** (LDA):
 - Projette les données sur un axe maximisant la séparation entre candidats
- **Arbres de décision boostés** (BDT):
 - Division de l'espace de données selon des coupures successives pour maximiser la séparation des candidats. Ceci est répété cherchant à minimiser une certaine fonction qui normalement traduit l'erreur



Évaluation des méthodes

Utilisation de la courbe ROC (Receiver Operating Characteristic) → traduit la capacité à éliminer le bruit de fond en fonction de l'efficacité d'identification du signal

· Performance évaluée par l'AUC (Area Under the Curve): le maximum est 1

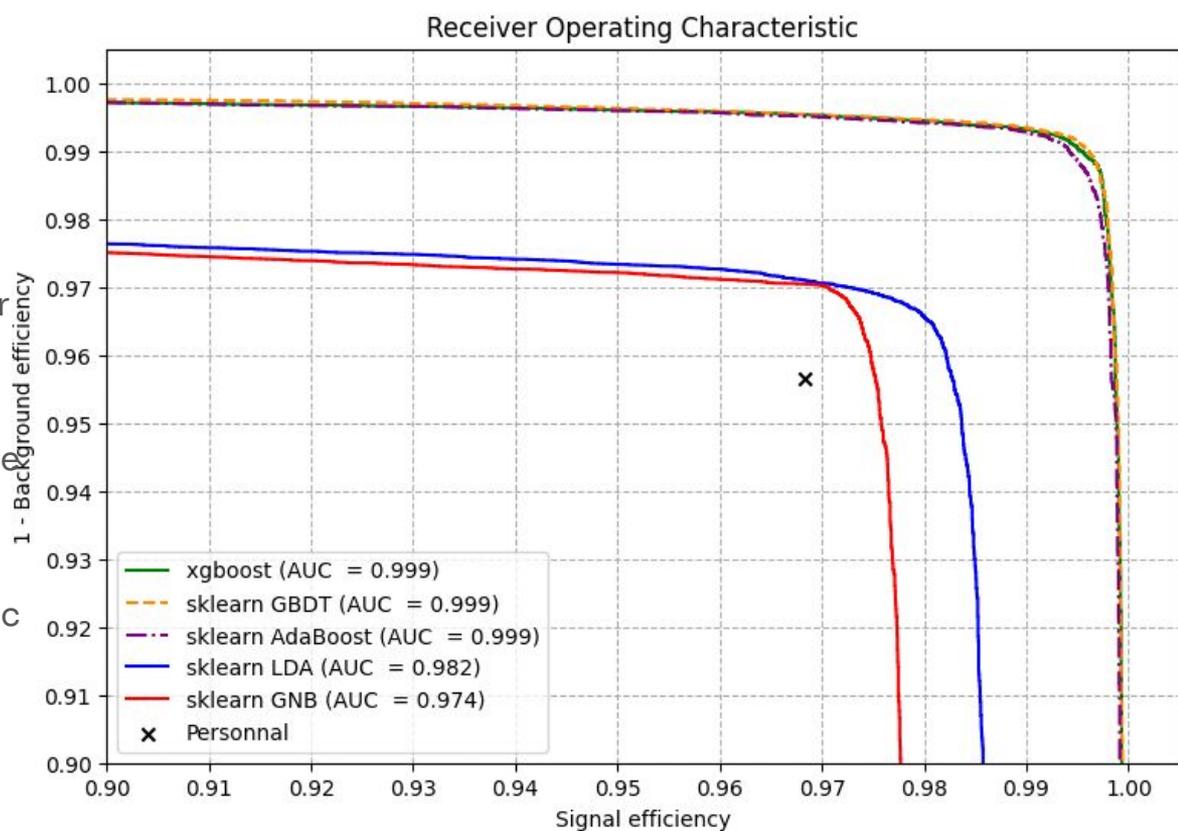
- Le résultat des coupures personnelles est affiché → très en dessous des algorithmes de machine learning et difficile à optimiser

- **Likelihood** et **LDA** offrent des performances similaires, correctes mais avec des limitations

- Les **BDT** sont les plus performants, se différencient par leur temps d'entraînement

· Les méthodes de machine learning permettent de définir les critères tight utilisés pour traiter les données issues d'ATLAS (possibilité de choisir différents points de fonctionnement)

· Les réseaux de neurones prennent de l'ampleur mais n'ont pas pu être étudiés lors de ce stage



Conclusion

- Analyse d'événements d'ATLAS: identification de la signature de la désintégration du Z^0 en di-électrons
- Reconstruction d'événements sur des données et sur de la simulation
- Étude d'efficacité avec la méthode tag & probe
- Compréhension des méthodes de sélection et limites des coupures simples, intégration du machine learning
- Découverte du travail en laboratoire et confirmation de mon intérêt pour la recherche en physique des particules