

Université de Montréal

Étude de la faisabilité d'une recherche de nouvelle
physique dans l'expérience ATLAS à l'aide d'un
auto-encodeur variationnel

par

Jacinthe Pilette

Département de physique
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en Physique

30 avril 2021

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Étude de la faisabilité d'une recherche de nouvelle physique dans l'expérience ATLAS à l'aide d'un auto-encodeur variationnel

présenté par

Jacinthe Pilette

a été évalué par un jury composé des personnes suivantes :

Claude Leroy

(président-rapporteur)

Jean-François Arguin

(directeur de recherche)

Jean-François Carrier

(membre du jury)

Résumé

Depuis la découverte du boson de Higgs en 2012, les physiciens et physiciennes des particules tentent de trouver des signes de nouvelle physique. Bien que le modèle standard décrivant les forces et particules aient été confirmé par les expériences telles que ATLAS au Grand collisionneur de hadrons (LHC), ce modèle décrit seulement 5% de la matière de notre Univers. Face à l'absence d'excès dans les recherches concentrées sur des modèles de nouvelle physique, l'intelligence artificielle pourrait être la solution.

Ce mémoire s'inscrit dans une perspective novatrice de recherche générale de nouvelle physique indépendante des modèles théoriques dans les données du détecteur ATLAS, par l'utilisation de l'apprentissage machine. Ici, nous nous intéressons à l'application de réseaux de neurones dans les données de simulation de collision proton-proton à $\sqrt{s} = 13$ TeV du détecteur ATLAS. Nous étudierons la performance d'auto-encodeurs variationnels dans les jets boostés, ces jets qui pourraient cacher des signes de nouvelle physique. Pour analyser la performance de notre réseau, nous entraînons celui-ci sur les quadri-vecteurs de jets issus de gluons et de quarks légers. Nous tentons de retrouver un signal test, les jets issus de quarks top boostés, dans les données de simulation en effectuant des sélections sur les scores d'anomalie retournés par le réseau.

Nos modèles d'auto-encodeurs variationnels atteignent une bonne discrimination entre le bruit de fond et le signal du quark top. Nous devons toutefois améliorer le rejet du bruit de fond pour purifier notre signal en fonction de nos sélections.

Mots-clés : physique des particules, LHC, ATLAS, apprentissage machine, réseaux de neurones, jets boostés

Abstract

Since the discovery of the Higgs boson in 2012, particle physicists are searching for signs of new physics. Although the standard model describing forces and particles has been confirmed by experiments like ATLAS at the Large Hadron Collider (LHC), it only describes 5% of the matter of the universe. Facing the absence of excess in searches for new physics model, artificial intelligence could be the solution.

This master thesis is part of a novel general model-independent search for new physics in the ATLAS detector data using machine learning. Here, we are interested in the application of neural networks in $\sqrt{s} = 13$ TeV proton-proton collision ATLAS simulation data. We study the performance of variational auto-encoders in boosted jets, who might be hiding signs of new physics. To analyze our network performance, we train the network on jets four-vectors coming from gluons and light quarks. We try to find a test signal, boosted top quark jets, in our simulation data by applying selections on the anomaly scores returned by our network.

Our variational auto-encoder reach a good discrimination between the background and the top quark signal. However, we need to improve background rejection to purify our signal as a function of our selections.

Keywords : particle physics, LHC, ATLAS, machine learning, neural networks, boosted jets

Table des matières

Résumé	5
Abstract	7
Liste des tableaux	13
Table des figures	15
Liste des sigles et des abréviations	21
Remerciements	25
Commentaires	27
Introduction	29
Chapitre 1. Le Modèle standard	31
1.1 Le Modèle standard	32
1.1.1 Invariance de jauge locale	32
1.1.2 Les particules	34
1.1.3 La force électromagnétique	36
1.1.4 La force nucléaire faible	38
1.1.5 La force nucléaire forte	39
1.1.6 La théorie électrofaible et le mécanisme de Higgs	43
1.2 La physique au-delà du modèle standard	45
1.2.1 Les limites du modèle standard	45
1.2.2 Les théories au-delà du modèle standard	47
1.2.3 Conclusion	50
Chapitre 2. Le Grand collisionneur de hadrons et l'expérience ATLAS....	53
2.1 Le Grand collisionneur de hadrons	53
2.1.1 L'accélération	54
2.1.2 La luminosité	56

2.1.3	Les expériences	57
2.2	Le détecteur ATLAS	57
2.2.1	Le système de coordonnées	58
2.2.2	Le détecteur interne	59
2.2.3	Les calorimètres	62
2.2.4	Le spectromètre à muons	64
2.2.5	Les détecteurs Timepix	65
2.2.6	Le système d'acquisition de données	66
2.3	La reconstruction d'événements	67
2.3.1	Les électrons et les photons	67
2.3.2	Les jets hadroniques	68
2.3.3	Les muons	71
2.3.4	L'énergie manquante	72
2.4	Enjeux démographiques et environnementaux au CERN	72
2.4.1	Conclusion	76
Chapitre 3. Motivations physiques et méthodologie		77
3.1	Choix de signal et de bruit de fond	77
3.1.1	Les jets boostés et les résonances	78
3.1.2	Signal et bruit de fond	79
3.2	Distance de Wasserstein	82
3.3	Les réseaux de neurones	86
3.3.1	Les réseaux de neurones	87
3.3.2	Entraînement et descente du gradient	89
3.3.3	Généralisation	91
3.3.4	Les auto-encodeurs	92
3.4	Identification des jets de quark top boostés avec modèle supervisé	94
3.5	Diversité et inclusion en intelligence artificielle	97
3.5.1	Conclusion	99
Chapitre 4. Résultats et discussion		101
4.1	Description des échantillons de simulation	101

4.2	Architecture du réseau et méthode d'analyse	104
4.2.1	Description de l'architecture	104
4.2.2	Méthode d'analyse	106
4.3	Résultats pour l'auto-encodeur variationnel	109
4.4	Auto-encodeur variationnel avec exposition aux valeurs aberrantes	114
4.4.1	Méthode d'exposition aux valeurs aberrantes	114
4.4.2	Résultats pour l'auto-encodeur variationnel avec exposition aux valeurs aberrantes	116
4.4.3	Améliorations futures	122
4.4.4	Conclusion	123
Conclusion		125
Références bibliographiques		127

Liste des tableaux

1	Description des échantillons de dijets QCD. Les événements sont générés par tranches d'impulsion correspondant aux divers échantillons JZW.	103
2	Description des échantillons de jets de quark top boostés. Les divers échantillons incluent les événements générés par tranches de masse invariante des deux quarks top initiaux (m_{tt}).	103
3	Hyperparamètres de l'auto-encodeur variationnel.	106
4	Signification statistique pour le beta-VAE en fonction de l'efficacité du bruit de fond (MSE). L'efficacité du signal/bruit de fond (ϵ_S/ϵ_B) correspond au taux de signal/bruit de fond restant après la sélection. Le nombre de signal (S) ou de bruit de fond (B) correspond au nombre de jets restant dans la région de masse d'intérêt de (140,200) GeV/ c^2 . Les chiffres correspondent au nombre de jets attendus dans 36 fb ⁻¹ dans les vraies données.	113
5	Signification statistique pour le beta-VAE en fonction de l'efficacité du bruit de fond (JSD). Les chiffres correspondent au nombre de jets attendus dans 36 fb ⁻¹ dans les vraies données.	113
6	Signification statistique pour le beta-VAE en fonction de l'efficacité du bruit de fond (KLD latent). Les chiffres correspondent au nombre de jets attendus dans 36 fb ⁻¹ dans les vraies données.	114
7	Signification statistique pour le KL-OE-VAE en fonction de l'efficacité du bruit de fond. Les chiffres correspondent au nombre de jets attendus dans 36 fb ⁻¹ dans les vraies données.	120
8	Signification statistique pour le MSE-OE-VAE en fonction de l'efficacité du bruit de fond. Les chiffres correspondent au nombre de jets attendus dans 36 fb ⁻¹ dans les vraies données.	121
9	Comparaison des significations statistiques entre le beta-VAE, le KL-OE-VAE et le TopoDNN en fonction de l'efficacité du bruit de fond.	

Les chiffres correspondent au nombre de jets attendus dans 36 fb^{-1} dans les vraies données..... 121

Table des figures

1	Le tableau du modèle standard. Cette visualisation regroupe les particules par leurs caractéristiques : la masse, le spin et la charge électrique. Les masses sont mesurées en méga électronvolts par c^2 ($1 \text{ MeV} = 1.602 \times 10^{-13} \text{ J}$ et $c = 3 \times 10^8 \text{ m/s}$). La charge est en unité de charge élémentaire ($e = 1.602177 \times 10^{-19} \text{ C}$) [157].	34
2	Le vertex principal de l'électrodynamique quantique (QED). Ce vertex permet les interactions entre fermion(chargé)-fermion(chargé)-photon. Dans le cas présent, deux électrons sont impliqués, mais ceux-ci peuvent également être des muons, des taus ou des quarks [88].	37
3	Diagrammes de Feynman d'ordre inférieur d'interactions de la QED. L'annihilation de paires (gauche), la production de paires (centre) et la Diffusion Compton (droite) [88].	37
4	Les vertex principaux de la force nucléaire faible. À gauche se trouve le vertex de base pour les courants chargés et à droite celui-ci pour les courants neutres [88]. Le courant chargé peut modifier la saveur des quarks et changer un lepton en son neutrino associé.	38
5	Exemples de diagrammes de Feynman pour des interactions sous la force faible. À gauche, un courant chargé impliquant des hadrons et des leptons lors d'une désintégration d'un pion (π). À droite, un canal impliquant que des leptons lors de la désintégration du tau (τ) par le W [148].	39
6	Les vertex principaux de la chromodynamique quantique (QCD). À gauche, le vertex qui permet les interactions entre quark-quark-gluon. Au centre, les interactions à trois et à droite, à quatre gluons [148].	40
7	Mesure de la quantité α_s en fonction de l'échelle d'énergie Q (GeV). Cette figure illustre la liberté asymptotique, caractéristique de la force forte. À mesure que l'énergie augmente, la constante de couplage diminue et la force forte devient de plus en plus faible [126].	41
8	Processus de formation de deux jets issus de quarks. À cause du confinement des quarks et de la liberté asymptotique, une paire de quark-antiquark	

	sera produite si on essaie de séparer des quarks (gauche) [148]. Aux hautes énergies, ce processus se produit de façon successive, résultant en une cascade de quarks (droite) [136].	42
9	Production de jet boosté dans une collision proton-antiproton. Ici, la collision initiale produit un quark top qui se désintègre et produit divers jets. Dans le régime boosté à haute énergie, les jets vont parfois se fusionner pour former un seul grand jet [145].	42
10	Diagramme de boucle entre le boson de Higgs et le quark top [158].	46
11	Limites sur les masses de diverses particules supersymétriques par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [33].	49
12	Limites sur les masses des nouvelles particules par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [30].	51
13	Limites sur les temps de vie des nouvelles particules par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [30].	52
14	Une vue aérienne du complexe du CERN montrant l'anneau de 27 km de circonférence du Grand collisionneur de hadrons [51].	54
15	Un schéma du complexe d'accélération qui indique toutes les étapes d'injection des protons dans l'anneau final [92].	55
16	Luminosité intégrée de l'expérience ATLAS entre 2015 et 2018. Ici, la luminosité est exprimée en fb inverse, où $1 \text{ fb} = 10^{-39} \text{ cm}^2$ [26].	57
17	Coupe transversale du détecteur ATLAS. La Figure représente les trois systèmes différents de détection des particules : le détecteur interne, les calorimètres et le spectromètre à muons [130]. Les axes en rouge ont été ajoutés à l'image originale.	58
18	Représentation simplifiée de la variable η dans le détecteur ATLAS. Le détecteur ATLAS est imagé en bleu. L'axe \hat{z} correspond à la limite où $\eta = \infty$. L'axe \hat{y} correspond à une valeur de $\eta = 0$	59
19	Schéma du détecteur interne. Celui-ci est composé du détecteur à pixels, du trajectomètre semi-conducteur (SCT) et du trajectomètre à radiation de transition (TRT) [128].	61

20	Le système de calorimètres du détecteur ATLAS. Le calorimètre électromagnétique est formé d'un baril (<i>LAr electromagnetic barrel</i>) et de deux disques (EMEC). Le calorimètre hadronique possède deux barils (<i>Tile barrel</i> et <i>Tile extended barrel</i>), ainsi que deux disques (HEC) et une partie avant (FCal) [127].	63
21	Le spectromètre à muons du détecteur ATLAS. Le spectromètre regroupe trois aimants (<i>Barrel</i> et <i>End-cap toroid</i>) et quatre systèmes de détection et de déclenchement (MDT, TGC, CSC et RPC) [129].	65
22	Formation des amas topologiques dans le calorimètre hadronique avant (FCAL). Les trois images représentent les étapes de formation des amas [25]. . . .	69
23	Le comportement de divers algorithmes de reconstruction de jets. Les algorithmes de k_T , de Cambridge et Aachen, de cône et d'anti- k_T reconstruisent différemment une collision de partons générée par un logiciel [48].	71
24	Figures discutant de l'évolution de la fraction des femmes en sciences, technologies, ingénierie et mathématiques (STIM) et au CERN. À gauche, la fraction des membres de l'expérience ATLAS qui sont des femmes en date de 2019 [31]. À droite, la fraction de baccalauréat obtenus par les femmes aux États-Unis dans les domaines des STEM [13].	75
25	L'observation de la résonance du boson Z au collisionneur LEP par la collaboration ALEPH. La section efficace du processus $e^+e^- \rightarrow q\bar{q}$ est affichée pour les canaux avec un photon et un boson Z virtuel (ligne pleine), ainsi que pour le processus de la QED (ligne pointillée) [7].	79
26	Production d'une paire de top anti-top au LHC. Le quark top se désintègre en un quark bottom et un boson W . Ce dernier peut se désintégrer leptoniquement (neutrino-lepton), comme à gauche, ou hadroniquement (quarks), comme à droite [47].	80
27	Masse invariante du premier jet, ou <i>leading jet</i> (m^{jet}), pour le signal (bleu) et le bruit de fond (orange). Les masses sont obtenues par les données de simulation Monte Carlo de collision proton-proton à 13 TeV et ajustée à une luminosité intégrée de 36 fb^{-1}	81
28	Transport optimal entre un jet de quark top de $p_T = 503.8 \text{ GeV}/c$ et un jet QCD de $p_T = 480.3 \text{ GeV}/c$. La valeur de EMD correspondant au transport optimal entre les deux événements est indiquée.	84

29	Distribution moyenne du EMD entre l'ensemble d'entraînement et les jets QCD et top de l'ensemble test.	85
30	Courbe ROC obtenue en utilisant le EMD comme discriminant. La Figure affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l'efficacité du signal (ϵ_S). Plus la courbe se rapproche du coin supérieur droit, mieux est la performance.	85
31	Comparaison entre les neurones du cerveau et un réseau de neurones artificielles. À gauche, des neurones traditionnelles qui propage un signal électrique vers l'avant [72]. À droite, un réseau de neurones avec une seule couche cachée qui propage l'information de gauche à droite [123]. Les b représentent les biais de chaque neurone, a l'activation de la neurone et w le poids transmis d'une neurone à l'autre (ces informations ont été ajoutées à l'image originale). La fonction $f(x)$ est celle prédite par le réseau.....	88
32	Fonctions d'activation les plus communes. Les fonctions sigmoïde et <i>softmax</i> permettent d'envoyer les données en sortie sur un intervalle entre (0,1). La fonction ReLU prend la forme d'une fonction linéaire pour $x > 0$	89
33	Schéma d'un auto-encodeur. À gauche, l'encodeur réduit la dimension des données d'entrées x et les envoie dans la couche latente. Puis, le décodeur génère des sorties y à partir de l'information de la couche latente.	93
34	Schéma d'un auto-encodeur variationnel. À l'entraînement, nous demandons à l'espace latent de suivre une distribution normale. L'information encodée pourra être échantillonnée à partir de cette distribution.	93
35	Courbes ROC pour le TopoDNN intégré dans les logiciels d'ATLAS. La Figure affiche le rejet du bruit de fond ($1/\epsilon_{\text{bkg}}$) en fonction de l'efficacité du signal (ϵ_{sig}). Les jets considérés sont à grande impulsion transverse [29].....	96
36	Impulsion transverse du premier jet (p_T^{jet}) pour le signal et le bruit de fond. À gauche, nous voyons le spectre d'impulsion tombant du bruit de fond des divers échantillons JZW de dijets QCD. À droite, nous voyons le spectre d'impulsion du signal en comparaison au bruit de fond. Les jets sont obtenus par des simulations Monte Carlo de collisions proton-proton à $\sqrt{s} = 13$ TeV et ajustées à une luminosité intégrée de 36 fb^{-1}	104
37	Effet espéré de la sélection à partir de la métrique d'anomalie sur le spectre de masse invariant. À gauche, le spectre de masse invariante du	

	premier jet (m_j), ou <i>leading jet</i> , avant la sélection. À droite, le spectre de masse invariant après la sélection effectuée à partir de la métrique d'anomalie.	107
38	Courbes ROC (gauche) et de gain de signal (droite) pour le beta-VAE sur l'ensemble test. La Figure de gauche affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l'efficacité du signal des quark top boostés (ϵ_S). La Figure de droite affiche le gain en signal (ϵ_S/ϵ_B) en fonction de l'efficacité du signal (ϵ_S).	109
39	Allure des différentes erreur de reconstruction pour le bruit de fond (bleu) et le signal (orange). Les histogrammes correspondent aux distributions attendues dans 36 fb^{-1} de données.	110
40	Effet des sélections sur le spectre de masse invariante du premier jet du bruit de fond (m_j) dans 36 fb^{-1} pour le beta-VAE. Les distributions sont normalisées.	111
41	Corrélation entre les métriques d'anomalie de MSE, JSD et KLD (latent) et la masse des jets. La Figure affiche la masse moyenne des jets (m_j^{mean}) en fonction de la sélection effectuée sur le score d'anomalie.	112
42	Courbes ROC (gauche) et de gain de signal (droite) pour le KL-OE-VAE sur l'ensemble test. La Figure de gauche affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l'efficacité du signal des quark top boostés (ϵ_S). La Figure de droite le gain en signal (ϵ_S/ϵ_B) en fonction de l'efficacité du signal des quark top boostés (ϵ_S).	117
43	Allure des erreurs de reconstruction du bruit de fond (bleu) et le signal (orange) pour le MSE-OE-VAE (gauche) et le KL-OE-VAE (droite). . . .	118
44	Effet des sélections sur le spectre de masse invariante du premier jet (m_j) du bruit de fond dans 36 fb^{-1} pour le beta-VAE (gauche) et le MSE-OE-VAE (droite). Les distributions sont normalisées.	119
45	Effet des sélections sur le spectre de masse invariante du premier jet (m_j) du bruit de fond dans 36 fb^{-1} pour le beta-VAE (gauche) et le KL-OE-VAE (droite). Les distributions sont normalisées.	119
46	Corrélation entre les métriques d'anomalie de MSE et KLD (latent) et la masses des jets pour le MSE-OE-VAE (gauche) et le KL-OE-VAE (droite). La Figure affiche la masse moyenne des jets (m_j^{mean}) en fonction de la sélection effectuée sur la métrique d'anomalie.	120

Liste des sigles et des abréviations

ADAM	<i>Adaptative Moment Estimation</i>
AE	Auto-encodeur
ALICE	<i>A Large Ion Collider Experiment</i>
ATLAS	<i>A Toroidal LHC Apparatus</i>
AUC	<i>Area Under the Curve</i>
BDT	<i>Boosted Decision Tree</i>
CAST	<i>CERN Axion Solar Telescope</i>
CB	<i>Combined muon</i>
CDMS	<i>Cryogenic Dark Matter Search</i>
CERN	Organisation européenne pour la recherche nucléaire
CMS	<i>Compact Muon Solenoid</i>
CNN	<i>Convolutional Neural Network</i>
CP	Charge-parité
CSC	<i>Cathode Strip Chamber</i>
CT	<i>Calorimeter-tagged muon</i>
CV	Curriculum vitae
DNN	<i>Deep Neural Network</i>
DPJ	Direction de la protection de la jeunesse
D-PHY	Comité Diversité Physique
ÉDI	Équité, diversité et inclusion
EMD	<i>Earth Mover's Distance</i>
EMEC	<i>Electromagnetic Endcap</i>
FCal	<i>Forward Calorimeter</i>
HEC	<i>Hadronic Endcap</i>
IA	Intelligence artificielle

IVADO	Institut de valorisation des données
JSD	<i>Jensen-Shannon Distance</i>
KLD	<i>Kullback-Leibler Divergence</i>
LAr	<i>Liquid Argon</i>
LEP	<i>Large Electron-Positron Collider</i>
LGBTQ+	Lesbiennes, gays, bisexuelles, trans, queers, et autres
LHC	<i>Large Hadron Collider</i>
LHCb	<i>Large Hadron Collider beauty</i>
MC	Monte Carlo
MDT	<i>Monitored Drift Tube</i>
ME	Muon extrapolé
MS	Modèle standard
MSE	<i>Mean Squared Error</i>
NLO	<i>Next to Leading Order</i>
OE	<i>Outlier Exposure</i>
OSQAR	<i>Optical Search of QED vacuum magnetic birefringence, Axion and photon Regeneration</i>
PS	Proton-synchrotron
QCD	<i>Quantum Chromodynamics</i>
QED	<i>Quantum Electrodynamics</i>
ReLU	<i>Rectifying Linear Unit</i>
RF	Radio-fréquence
ROC	<i>Receiver Operating Characteristic</i>
RPC	<i>Resistive Plate Chamber</i>
SCT	<i>Semiconductor Tracker</i>
SNO	<i>Sudbury Neutrino Observatory</i>
SPS	Super proton-synchrotron
ST	<i>Segment-tagged muon</i>
STIM	Sciences, technologies, ingénierie et mathématiques
SUSY	Supersymétrie

TGC	<i>Thin Gap Chamber</i>
TRT	<i>Transition Radiation Tracker</i>
VAE	<i>Variational Autoencoder</i>
WIMP	<i>Weakly Interactive Massive Particle</i>

Remerciements

Je voudrais d'abord remercier Jean-François, mon directeur de recherche. Je suis choyée d'avoir pu travailler avec toi au cours des deux dernières. Ton ouverture d'esprit et ton soutien m'ont encouragé tout au long de ma maîtrise. Je remercie les différentes personnes avec qui j'ai travaillé au sein de notre équipe de recherche, le groupe ATLAS-UdeM, ainsi que nos collaborateurs à Genève, Arlington et Tokyo, qui m'ont procuré plusieurs conseils à travers la réalisation de mon projet.

Je remercie IVADO de m'avoir soutenu financièrement durant ces deux années. En plus de ce soutien financier, j'ai pu profiter de leurs conseils et de leur offre d'activités de formation et de sensibilisation sur l'ÉDI. Je salue leur désir de toujours chercher à mieux soutenir leurs étudiants et étudiantes.

Pour les discussions, les projets et les belles rencontres, je remercie l'ensemble de mes collègues du comité D-PHY avec qui j'ai pu travailler durant les cinq dernières années. Le comité a été pour moi un environnement d'ouverture et d'échange d'idées dont j'avais besoin. Continuez votre fabuleux travail et je vous souhaite le meilleur pour la suite!

Je salue l'association étudiante, la PHYSUM, avec qui j'ai travaillé durant plusieurs années. Sans vous, la vie étudiante au département ne serait pas la même. On oublie souvent que les gens qui s'impliquent le font bénévolement, en plus de leurs études. Continuez votre travail acharné et continuez toujours à vous améliorer.

Un merci à Mirjam d'avoir fait figure de mentore et d'amie durant mes études en physique. Merci pour nos nombreuses discussions sur l'ÉDI, pour ton grand soutien durant mes études supérieures, et de m'avoir lu et conseillé durant l'écriture de ce mémoire.

Merci à mes parents, André et Brigitte, qui nous ont toujours encouragés à poursuivre nos études. Merci de m'avoir soutenue, c'est grâce à vous que je suis rendue ici. Merci à mes grands frères (littéralement), dont je suis la plus fière! Merci à Dune, mon chien,

pour les câlins réconfortants. Je salue aussi ma belle-famille, qui m'encourage toujours dans ce que j'entreprends.

Je remercie mes chères amies d'avoir été là pour moi durant les moments difficiles - Jessie, Vicky et Sarah. Vous êtes les meilleures.

Enfin, je remercie mon journaliste préféré, mon réviseur linguistique par excellence et mon compagnon de vie – François - qui m'a épaulé durant mes nombreuses remises en question et mes périodes difficiles. Merci d'avoir lu ce mémoire beaucoup trop de fois.

Commentaires

Je tiens à débiter ce mémoire par un petit mot sur mon parcours. Mon désir de poursuivre des études en physique remonte au troisième secondaire. À cette époque, j'habitais mon village d'origine des Laurentides, Nominique, où la pollution visuelle est presque inexistante et où nous voyons les étoiles à l'année longue. À ce moment, je voulais devenir astrophysicienne. Dans mes temps libres, je m'amusais à lire la vieille version de mes parents du livre *Cosmos* de Carl Sagan, et je louais l'unique manuel d'astrophysique à la bibliothèque du village pour en faire les exercices. Je souhaitais comprendre comment fonctionne notre Univers, et rien n'allait m'arrêter.

Avant d'entrer au baccalauréat en physique à l'Université de Montréal, en 2016, je n'avais jamais questionné mes compétences en sciences, ni ma place dans la communauté en physique. Je n'avais pas conscience du portrait démographique actuel de la physique, majoritairement masculin, et je n'avais donc aucune attente en entrant au département. C'était pleine de volonté et d'étoiles dans les yeux que je commençais mon parcours. Cependant, à mesure que mon expérience en tant que physicienne s'enrichissait, et malgré mes bonnes notes, mes sentiments de compétence et d'appartenance à cette « étiquette » de physicienne diminuaient. Ce sentiment grandissant, a fini par m'habiter tout au long de mes études de 1^{er} cycle, et même ma maîtrise.

En 2017, lors d'un de mes cours, j'ai pris connaissance de l'existence du comité Diversité Physique (D-PHY) au département. Ce comité qui a pour objectifs « d'augmenter la visibilité des minorités (genre, origine ethnique, orientation sexuelle) en physique et en sciences », de « briser les stéréotypes », de « soutenir les physiciennes et physiciens lors de l'entrée sur le marché de l'emploi » et de « favoriser une meilleure conciliation travail-famille en physique » m'a interpellé. Ayant déjà, à cette époque, vécu des situations difficiles au département, je me rappelle être allée dans le bureau d'une professeure du comité pour lui partager mon inconfort et mon désir de faire partie du comité. Je souhaitais, en m'impliquant, participer à un changement de culture et initier des discussions au département. J'ai été accueillie à bras ouverts.

Au sein du comité D-PHY, j'ai compris que d'autres femmes vivaient la même situation que moi et consacraient une partie de leur énergie à se convaincre qu'elles avaient une place en physique. Ce n'est qu'après plusieurs années de réflexions et de discussions avec mes collègues que j'ai pu attribuer des mots précis à ces sentiments. Un manque d'identification à l'identité stéréotypée du physicien « génie » et un syndrome de l'imposteur avaient réussi à me faire douter de mes qualifications de physicienne.

Pour faire sens de mon expérience, je me suis donc activement impliquée dans le Comité D-PHY. Cette implication a profondément marqué mon parcours au baccalauréat et à la maîtrise. Mes années au comité m'ont fait réfléchir sur plusieurs aspects de la communauté en physique. Nous avons souvent tendance à détacher la physique de tous enjeux sociaux, mais la réalité est que nos projets de recherche, à première vue sans incidence, ont des implications sociales et environnementales. De ce fait, pour refléter mes réflexions et mon parcours, j'ai tenu à commenter mon projet de recherche, et par le fait même humaniser un peu la recherche en physique. Ainsi, au long de ce mémoire, je commente divers aspects de mon projet en lien avec l'équité, la diversité et l'inclusion (ÉDI) [100]. Ces commentaires sont issus d'une réflexion personnelle, de nombreuses discussions avec mes collègues du comité D-PHY et d'une revue de littérature sur le sujet. Ces commentaires ont trois objectifs. Premièrement, ils visent à sensibiliser sur des enjeux qui me sont absolument essentiels à la bonne pratique de la recherche scientifique. Deuxièmement, ils reflètent mes questionnements sur nos pratiques et notre responsabilité en tant que physiciens et physiciennes. Troisièmement, ils visent à faire le point sur la place des femmes et des minorités en physique en 2021. Par cette démarche, j'espère inspirer mes collègues en physique d'effectuer une réflexion similaire en lien avec leurs propres projets.

Enfin, j'ai choisi d'utiliser, dans la mesure du possible, l'écriture inclusive à travers mon mémoire. Comme l'indique l'Université de Montréal dans son *Guide d'écriture inclusive pour tous et toutes*, « traditionnellement, en français, au masculin était attribuée une universalité de fait, voire une certaine neutralité, mais ce n'est plus le cas. [...] [L]'égalité entre les femmes et les hommes passe aussi par la mobilisation des ressources de la langue française » [45]. La forme en doublon (physicien et physicienne, chercheur et chercheure) m'a semblée la plus appropriée, car elle explicite la présence des femmes à part entière en physique. Lorsque possible, j'utilise également la forme neutre pour inclure toute personne qui ne se sent pas rejointe par la forme masculine ou féminine traditionnelle.

Introduction

L'intelligence artificielle (IA) séduit de plus en plus la communauté scientifique. Des algorithmes de reconnaissance vocale aux algorithmes de détection par images, les possibilités d'application sont infinies. La recherche en physique profite de ses nombreuses possibilités. L'apprentissage machine est de plus en plus intégré au Grand collisionneur de hadrons (LHC), que ce soit pour la reconstruction ou pour l'identification d'événements, et nous permet d'améliorer nos résultats actuels confirmant le modèle standard de la physique des particules. Alors que les expériences ATLAS et CMS situées au LHC ont conjointement annoncé la découverte du boson de Higgs en 2012, nous nous trouvons dans une nouvelle ère pour la physique des particules : celle de la recherche de nouvelle physique. En effet, de nombreux mystères demeurent non résolus, dont la matière sombre, l'absence de la gravité dans le modèle standard et l'oscillation des neutrinos. Ces mystères impliquent l'existence de nouvelles particules qui pourraient potentiellement être produites et découvertes au LHC et motivent nos recherches de nouvelle physique.

Les physiciens et physiciennes ne sont pas sans outils pour faire face à ces nouveaux défis. De nombreux modèles théoriques existent pour tenter d'expliquer ces phénomènes. La supersymétrie est l'un des modèles de nouvelle physique les plus populaires qui fait toujours l'objet de recherches au CERN. Cependant, ces nombreux modèles, ingénieux sur papier, sont actuellement introuvables dans nos données après des années de recherche qui nous ont laissé bredouilles. Avons-nous atteint les limites de nos connaissances ? Ou bien n'avons-nous simplement pas encore trouvé le « bon » modèle théorique de physique au-delà du modèle standard qui répondrait à toutes nos questions ? Heureusement, l'apprentissage machine pourrait être la lueur d'espoir dans notre quête de nouvelle physique, et les nombreux développements récents dans le domaine de l'intelligence artificielle pourraient s'avérer bénéfiques dans nos recherches.

Dans ce mémoire, nous tentons de procéder à une recherche de nouvelle physique dans l'expérience ATLAS indépendante des modèles théoriques. Ainsi, nous essayerons de trouver des anomalies dans les données du LHC sans hypothèse théorique initiale sur

leur nature. Pour ce faire, nous développons un réseau de neurones, soit un auto-encodeur variationnel, et nous appliquons ce dernier à un espace restreint des données d'ATLAS. Les jets boostés, qui sont un des suspects favoris dans les recherches de nouvelle physique, sont notre terrain de jeu. Nous entraînons le réseau sur ces jets du modèle standard, en espérant qu'il en apprenne la structure. Si le réseau rencontre des jets anormaux issus de nouvelle physique, celui-ci devrait alors être capable de les identifier par des techniques qui seront développées dans ce mémoire.

Toutefois, ce projet novateur requiert de nombreuses étapes avant d'effectuer la recherche générale elle-même. Pour tester notre approche, nous entraînons d'abord notre réseau de neurones sur un ensemble simulé de jets boostés du modèle standard provenant de gluons et de quarks légers. Nous appliquons ensuite notre réseau dans un ensemble d'évaluation incluant ces jets et un signal « test », pour tenter de retrouver dans nos données, ou de « redécouvrir », ce signal correspondant à une particule bien connue du modèle standard : le quark top.

Le premier chapitre de ce mémoire présente les notions théoriques du modèle standard de la physique des particules. Nous y explorons les forces, les particules, mais aussi les limites de ce modèle et des exemples de recherches de nouvelle physique effectuées au CERN. Dans le deuxième chapitre, nous décrivons plus en détails le fonctionnement technique de l'expérience ATLAS. Nous expliquons comment les jets sont formés et reconstruits dans ce détecteur et comment nous en extrayons les données. Le troisième chapitre aborde les bases de l'apprentissage machine nécessaires à la compréhension de ce projet. Nous y expliquons notre choix de signal à retrouver (le quark top) et notre choix de bruit de fond. Puis, nous détaillons les premiers résultats de notre recherche au quatrième chapitre. Nous y abordons également une nouvelle technique nous permettant de régler un problème de corrélation entre nos données et notre réseau de neurones.

Enfin, à travers ce mémoire, nous explorerons mes contributions à ce projet de développement de recherche de nouvelle physique. Durant mes études graduées, j'ai travaillé au développement de la méthode utilisant la distance de Wasserstein, ou *Earth Mover's Distance* (EMD), comme détection d'anomalies (section 3.2) et à l'application d'un auto-encodeur variationnel dans des données de simulation ATLAS (résultats du chapitre 4). Bien que ce mémoire reflète mon travail, il représente aussi le fruit de la collaboration ATLAS qui a produit les échantillons de simulation ATLAS et du groupe de recherche de l'UdeM, qui a travaillé à l'élaboration du réseau de neurones en utilisant la technique d'exposition aux anomalies et à la production des échantillons Delphes dans la référence [58].

Chapitre 1

Le Modèle standard

Le présent mémoire s'inscrit dans le domaine de la physique des particules. Cette branche de la physique s'intéresse aux constituants élémentaires de la matière et à leurs interactions. C'est un domaine qui s'est particulièrement développé durant les années 1960 et 1970 avec l'essor des accélérateurs de particules [88]. Depuis, les technologies modernes ont permis d'atteindre des énergies de centre de masse de plus en plus élevées dans les accélérateurs et d'effectuer des mesures d'une précision inégalée confirmant nos connaissances actuelles.

Le modèle standard (MS) est la théorie principale en physique des particules. Elle décrit les particules élémentaires et leurs interactions par trois des quatre forces fondamentales en physique : la force électromagnétique, la force nucléaire forte et la force nucléaire faible. De nombreuses expériences ont confirmé cette théorie, dont plusieurs réalisées au Grand collisionneur de hadrons (LHC) en Suisse [32]. D'ailleurs, le dernier élément prédit par le modèle standard, le boson de Higgs, a finalement été observé au LHC en 2012 [18, 60]. Cependant, la nouvelle physique se cache toujours, et de nombreuses questions demeurent en suspens. Alors que le MS décrit adéquatement les interactions de la matière ordinaire, il possède certaines limites. Par exemple, cette théorie n'inclut ni la gravité, quatrième force fondamentale, ni la matière sombre qui compose près de 25% de l'Univers [65]. Un des principaux objectifs de la physique des particules moderne consiste à combler ces lacunes et à trouver des signes de physique au-delà du modèle standard [49]. Les récentes expériences au LHC n'ont toujours pas détecté ces signes, et les recherches se poursuivent.

Durant le présent chapitre, nous détaillerons la théorie du MS. D'abord, nous en poserons les bases mathématiques. Puis, nous présenterons les particules fondamentales décrites par le MS. Nous aborderons ensuite les trois forces fondamentales incluses dans cette théorie, de même que la théorie électrofaible et le mécanisme de Higgs. Finalement, nous concluons ce chapitre par une description des limites de cette théorie et par un survol

des recherches de nouvelle physique effectuées jusqu'à présent pour mettre en relief les motivations derrière cette recherche.

1.1. Le Modèle standard

Le MS ne consiste pas seulement en un registre des particules fondamentales ; il incarne aussi une théorie des champs quantiques qui régit leurs interactions. Le MS est aujourd'hui composé de deux théories principales : la chromodynamique quantique (QCD) et la théorie électrofaible [136]. Pour bien comprendre le comportement des particules dans ces théories, il nous faut d'abord établir les fondements de cette théorie des champs. Dans la présente section, nous explorerons d'abord le formalisme du MS. Ensuite, nous pourrions découvrir les particules décrites par ce modèle et explorer les différentes forces par lesquelles elles interagissent. Nous commencerons par décrire les forces électromagnétique et nucléaire faible, puis nous décrirons la chromodynamique quantique (QCD), soit la théorie de la force forte. Finalement, nous développerons la théorie électrofaible qui unit les deux autres forces, ainsi que le mécanisme de Higgs, qui joue un rôle particulier dans le MS.

1.1.1. Invariance de jauge locale

D'un point de vue mathématique, le modèle standard pose ses fondations sur les concepts de symétrie et d'invariance de jauge [136]. D'abord, les interactions des particules sont définies par un Lagrangien, c'est-à-dire une équation du système qui décrit le mouvement des particules. Cette équation nous permet de comprendre le comportement des particules libres, un peu comme des billes flottant dans l'espace. Toutefois, pour pouvoir décrire les particules sous l'influence des différentes forces fondamentales, il est nécessaire d'effectuer diverses transformations de jauge sur le Lagrangien des particules libres et d'exiger que celui-ci demeure invariant. Cette invariance impose alors une symétrie au système représentée par un groupe de symétrie. Ce groupe inclut toutes les transformations sous lesquelles le Lagrangien est invariant.

Nous introduisons ensuite deux types d'invariance : les invariances de jauge globale et locale. L'invariance globale concerne un Lagrangien dont les paramètres sont constants dans l'espace-temps, un peu comme si nos billes étaient affectées de la même façon par la gravité partout dans l'espace. L'invariance locale concerne plutôt un Lagrangien pour lequel les paramètres dépendent de la position. Selon le théorème de Emmy Noether, mathématicienne allemande ayant largement contribué aux travaux de la Théorie des

groupes de symétrie : une invariance du Lagrangien sous une transformation correspond aussi à une quantité conservée. Ainsi, soit une transformation locale d'un champ $\phi(x)$:

$$\phi(x) \longrightarrow e^{iq\alpha(x)}\phi(x) \quad ; \quad (1)$$

avec α une quantité infinitésimale et q une charge. Cette transformation correspond à une symétrie si le Lagrangien est modifié, en suivant la notation tensorielle, tel que :

$$\mathcal{L}(x) \longrightarrow \mathcal{L}(x) + \alpha(x)\partial_\mu J^\mu(x) \quad ; \quad (2)$$

avec $J^\mu(x)$ un courant. Cette symétrie impliquera alors une quantité conservée qui correspond au courant $J^\mu(x)$. Dans la théorie du MS, on parle toujours en termes d'invariance de jauge locale. En effet, l'apparition des bosons de jauge est une conséquence de la condition d'invariance du Lagrangien sous une transformation de jauge locale.

Dans le cas du MS, les quantités (ou courants) conservées correspondent à différentes charges. La symétrie de jauge du MS correspond au groupe $SU_c(3) \times SU(2)_L \times U(1)_Y$. La symétrie $SU_c(3)$ décrit la chromodynamique quantique, soit les interactions par la force forte, avec la charge de couleur c . Le groupe $SU(2)_L \times U(1)_Y$ décrit la théorie électrofaible qui unit les forces électromagnétique et nucléaire faible. Dans le cadre de cette théorie, l'hypercharge (Y) est conservée et le L fait référence aux particules gauchères. Finalement, chacun de ces groupes de symétrie possède des générateurs qui correspondent à divers bosons de jauge.

Ce formalisme nous permet de comprendre les interactions des 17 particules fondamentales qui composent le MS. Les forces par lesquelles elles interagissent sont alors décrites par diverses symétries de jauge. Avec ces bases mathématiques, nous pourrons maintenant explorer dans la prochaine section les différentes particules fondamentales, avant de se pencher sur les différentes forces décrites par ces symétries de jauge dans le MS.

1.1.2. Les particules

Le MS contemporain est généralement illustré tel qu'à la Figure 1. Ici, les différentes particules y sont regroupées par propriétés, incluant la masse, la charge électrique et le spin.

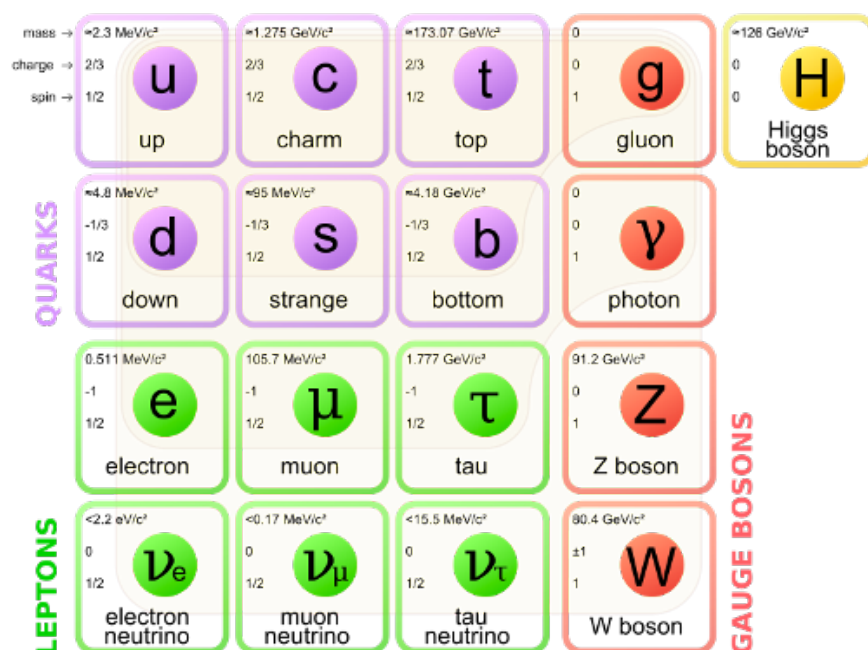


Figure 1. Le tableau du modèle standard. Cette visualisation regroupe les particules par leurs caractéristiques : la masse, le spin et la charge électrique. Les masses sont mesurées en méga électronvolts par c^2 ($1 \text{ MeV} = 1.602 \times 10^{-13} \text{ J}$ et $c = 3 \times 10^8 \text{ m/s}$). La charge est en unité de charge élémentaire ($e = 1.602177 \times 10^{-19} \text{ C}$) [157].

Le MS sépare les particules en deux grandes catégories : les bosons de jauge et les fermions. D'abord, les bosons de jauge, situés dans la quatrième colonne de la Figure 1, possèdent des spins entiers ($S = 1$). Ils sont aussi nommés vecteurs de force, car ils sont responsables des interactions entre les différentes particules. En effet, c'est par l'échange d'un boson de jauge que les particules interagissent. Les bosons couplent également à différents types de charges. Le photon, qui a une masse et une charge électrique nulles, interagit avec toutes les particules chargées électriquement. Il est effectivement responsable des interactions liées à la force électromagnétique. Les gluons – on en compte huit – possèdent également une masse et une charge électrique nulles. Ils couplent plutôt à ce qu'on appelle la charge de couleur. Ces bosons portent en effet une couleur : bleu, rouge ou vert (ou une combinaison linéaire non neutre) et vont interagir avec les particules qui en portent également. Ils sont responsables des interactions par la force nucléaire forte, la force responsable de la cohésion du noyau atomique. Les bosons W (W^+ et W^-) et le boson Z (charge électrique neutre),

quant à eux, interagissent par la force nucléaire faible. Ils forment ce qu'on appelle les bosons intermédiaires et sont tous les trois massifs.

Les fermions, situés dans les trois premières colonnes de la Figure 1, possèdent un spin demi-entier ($S = 1/2$). Ces particules sont présentes en trois générations, représentées par les trois colonnes : $\{u, d, e, \nu_e\}$, $\{c, s, \mu, \nu_\mu\}$ et $\{b, t, \tau, \nu_\tau\}$, placées en ordre croissant de masse de gauche à droite. La première génération est stable, ce qui veut dire que ses particules ont un très long temps de vie. Contrairement à cette dernière, les deuxième et troisième générations sont très massives et sont donc très instables. Les particules de ces générations se désintègrent alors rapidement vers des fermions d'une génération précédente et ne se retrouvent pas dans notre univers courant. Aussi, pour chacun de ces fermions, il existe une antiparticule correspondante de même masse, de même spin et de nombres quantiques opposés. Par exemple, l'antiparticule de l'électron (e^-) est le positron (e^+) qui a une charge électrique opposée ($Q = +1$). L'antiparticule du neutrino électronique (ν_e) sera l'antineutrino électronique ($\bar{\nu}_e$) qui est également électriquement neutre ($Q = 0$).

Les fermions se divisent également en deux sous-catégories – les leptons et les quarks. Les six leptons du MS sont l'électron (e^-), le muon (μ^-), le tau (τ^-) et les neutrinos électronique (ν_e), muonique (ν_μ) et taunique (ν_τ). Les neutrinos sont très légers et possèdent une charge électrique nulle, alors que l'électron, le muon et le tau sont chargés négativement et possèdent une masse non-nulle. Le trio (e, μ, τ) interagit par les forces électromagnétique et nucléaire faible en échangeant un photon ou un des bosons W et Z . Les neutrinos, quant à eux, n'interagissent uniquement que par la force nucléaire faible.

Puis, nous avons les six quarks, soient les quarks up (u), down (d), strange (s), charm (c), bottom (b) et top (t), en ordre croissant de masse. Ceux-ci ont des charges électriques fractionnaires de $Q = +2/3$ ou $Q = -1/3$. En plus de posséder une charge électrique, les quarks portent aussi une charge de couleur, comme les gluons. Ainsi, les quarks peuvent interagir par la force forte, en plus de la force électromagnétique et de la force faible. De plus, on ne retrouve pas de quarks libres dans la nature, contrairement aux leptons. Ceux-ci se retrouvent plutôt sous forme combinée pour former d'autres particules de couleur neutre que l'on nomme les hadrons. Ces hadrons existent sous deux formes. D'un côté, les mésons sont formés d'un quark et d'un antiquark, tel que le pion chargé ($\pi^+ = u\bar{d}$). D'un autre côté, se trouvent les baryons, formés de trois quarks ou trois antiquarks, tel que le proton ($p = uud$) qui est une des particules située dans le noyau atomique.

Finalement, le boson de Higgs, situé tout à droite de la Figure 1, est le seul boson scalaire du MS, c'est-à-dire qu'il possède un spin nul ($S = 0$) et une parité P positive

(+1). Ici, la parité fait référence au comportement de la particule sous l'opérateur de parité P qui inverse l'espace, telle une réflexion dans un miroir. C'est l'interaction entre le champ du Higgs scalaire et les fermions et les bosons intermédiaires qui permet à ces particules d'acquiescer une masse. Ce processus sera décrit plus en détails à la Section 1.1.6.

En bref, les fermions et les bosons représentent les ingrédients du MS. Toutefois, les étapes de la recette, soit leurs interactions, doivent aussi être définies. Pourquoi, par exemple, le gluon n'interagit-il pas avec l'électron? Que se passe-t-il si on essaie de faire éclater le proton pour observer les quarks librement? Pour répondre à ces questions, nous allons explorer les différentes forces qui décrivent ces interactions.

1.1.3. La force électromagnétique

La force électromagnétique a d'abord été décrite par l'électrodynamique quantique, ou QED. C'est la première théorie de jauge à avoir été introduite dans le MS, et ses débuts remontent aux années 1920. Le développement de l'électrodynamique quantique a fortement inspiré les physiciens et physiciennes à développer un formalisme semblable pour les autres forces du modèle standard [88].

La QED est décrite par le groupe de symétrie $U(1)$, soit le groupe des transformations de phase unitaires. Le Lagrangien pour la QED se développe en partant d'un fermion libre, par exemple un électron. Comme énoncé précédemment, on effectue alors une transformation de jauge locale sur le Lagrangien du fermion. Afin de préserver l'invariance, il est nécessaire de faire apparaître un boson de jauge, ici le A_μ , qui sera le générateur du groupe $U(1)$. Ce boson de jauge représente ici le photon sans masse. L'invariance de jauge implique alors une quantité conservée. Dans le cadre de la QED, celle-ci est la charge électromagnétique. Le photon interagit donc avec toutes les particules qui possèdent une charge électromagnétique.

Il existe un seul vertex de base pour la QED. Celui-ci est présenté à la Figure 2. C'est à partir de ce vertex que toutes les autres interactions de la QED sont décrites. Ces interactions sont incluses dans le Lagrangien de la QED à l'équation 3. Le premier terme correspond au terme cinétique, le deuxième terme correspond à l'interaction entre le photon et les fermions et le dernier terme représente les champs (potentiels) électriques et magnétiques issus des équations de Maxwell.

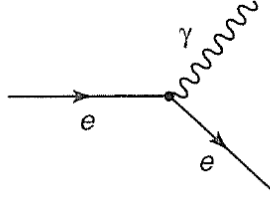


Figure 2. Le vertex principal de l'électrodynamique quantique (QED). Ce vertex permet les interactions entre fermion(chargé)-fermion(chargé)-photon. Dans le cas présent, deux électrons sont impliqués, mais ceux-ci peuvent également être des muons, des taus ou des quarks [88].

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad ; \quad (3)$$

La QED permet de décrire plusieurs types d'interactions électromagnétiques. Parmi celles-ci se trouvent l'émission ou l'absorption d'un photon. Dans ce processus, un électron émet ou absorbe un photon pour diminuer ou augmenter son énergie. L'annihilation ou la production de paires fermion-antifermion et la diffusion Compton, telles qu'illustrées à la Figure 3, sont aussi des interactions communes.

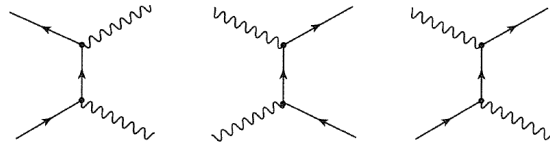


Figure 3. Diagrammes de Feynman d'ordre inférieur d'interactions de la QED. L'annihilation de paires (gauche), la production de paires (centre) et la Diffusion Compton (droite) [88].

Le rayonnement Bremsstrahlung est une autre interaction importante en QED et très commune au LHC. Dans ce processus, un électron qui passe près d'un noyau atomique est influencé par le champ électrique de celui-ci. Sa trajectoire est alors déviée et, tel que décrit par les équations Maxwell, il doit émettre un photon. Dans un contexte de hautes énergies, les photons produits pourront produire une paire fermion-antifermion, qui à son tour va émettre un photon par rayonnement Bremsstrahlung. Ce cycle se répète et une cascade électromagnétique de photons et de fermions se crée.

1.1.4. La force nucléaire faible

L'interaction faible est responsable de la radioactivité, tel que mentionné plus tôt. L'étude de la radioactivité a certainement été charnière dans la compréhension de la force faible. Malheureusement, les contributions des femmes dans ce domaine ont longtemps été occultées. Parmi ces scientifiques se trouvent Harriet Brooks, Marie Curie, Irène Curie, Stefanie Hoont, Lise Meitner et Jesse Slater, pour n'en nommer que quelques-unes, et qui ont joué des rôles importants dans la compréhension des désintégrations radioactives [138].

La force nucléaire faible a d'abord été décrite par le groupe de symétrie $SU(2)$, soit le groupe des matrices unitaires sous rotations en deux dimensions. Le groupe $SU(2)$ possède trois générateurs et sont représentés par les bosons massifs W^+ , W^- et Z^0 . La force faible affecte l'ensemble des fermions, incluant les neutrinos, et conserve la charge que l'on nomme l'isospin faible (I_W).

Les interactions possibles de la force faible sont divisées en deux catégories : les courants chargés et les courants neutres. Les vertex de base de ces deux courants sont présentés à la Figure 4. Comme les bosons W^\pm et Z^0 sont massifs, ils peuvent aussi interagir entre eux.

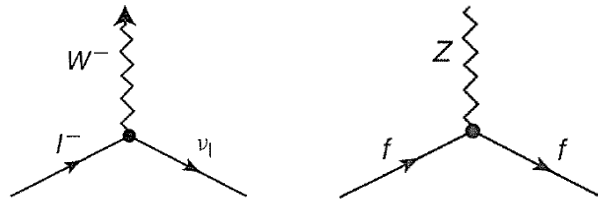


Figure 4. Les vertex principaux de la force nucléaire faible. À gauche se trouve le vertex de base pour les courants chargés et à droite celui-ci pour les courants neutres [88]. Le courant chargé peut modifier la saveur des quarks et changer un lepton en son neutrino associé.

Il existe différents types d'interactions possibles définies à partir de ces vertex principaux. Par exemple, le boson W peut interagir avec un hadron d'un côté du vertex et avec des leptons de l'autre, comme à la gauche de la Figure 5. On peut aussi avoir un canal impliquant que des leptons, comme à la droite de la Figure 5 dans le cas de la désintégration du τ .

La force faible est également la seule force du MS qui brise la symétrie CP (charge-parité). L'opérateur de charge (C) permet d'inverser la charge d'une particule, alors que l'opérateur

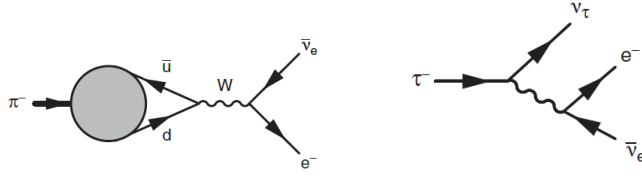


Figure 5. Exemples de diagrammes de Feynman pour des interactions sous la force faible. À gauche, un courant chargé impliquant des hadrons et des leptons lors d’une désintégration d’un pion (π). À droite, un canal impliquant que des leptons lors de la désintégration du tau (τ) par le W [148].

de parité (P) permet d’inverser l’espace, telle une réflexion dans un miroir. Les forces électromagnétique et forte conservent ces deux symétries individuelles et combinées, alors que la faible les brise. Ce phénomène reste toujours inexplicé et est un des nombreux mystères qui demeurent à résoudre dans le MS. C’est d’ailleurs la physicienne chino-américaine Chien-Shiung Wu qui a démontré expérimentalement que la parité n’était pas conservée dans les interactions faibles lors de l’expérience Wu [122]. Elle reçut plus tard le *Wolf Prize in Physics* pour ses contributions à l’élaboration du MS.

1.1.5. La force nucléaire forte

La théorie qui décrit les interactions de la force nucléaire forte se nomme la chromodynamique quantique (QCD). Elle est décrite par le groupe de symétrie $SU(3)$, soient les matrices unitaires sous rotations en trois dimensions. Alors que la force électromagnétique est liée à la charge électrique, la force forte est plutôt liée à la charge de couleur. Les trois composantes de la charge $SU(3)$ sont en effet des couleurs : bleu, rouge et vert. De plus, les bosons de jauge associés à ce groupe, les gluons, sont au nombre de huit. Ils représentent les générateurs du groupe $SU(3)$. Ainsi, la charge de couleur sera conservée dans les interactions de la force forte entre gluons et quarks. Comme les fermions et les autres bosons ne ressentent pas cette force, ils doivent conséquemment être de couleur « neutre ». La force forte existe aussi dans les interactions avec des hadrons et des mésons, qui sont formés de quarks et d’antiquarks. Par exemple, le proton, principale particule utilisée dans les collisions au LHC, peut interagir via la force forte.

Il existe trois vertex de base pour la QCD, tel qu’illustré à la Figure 6. Le premier permet les interactions entre quark-quark-gluon, et les deux autres permettent les auto-interactions à trois et quatre gluons. Le vertex impliquant les quarks permet aussi à ces derniers de changer de couleur. Par exemple, un quark rouge pourrait, en échangeant un

gluon, devenir un quark bleu. Ces vertex sont décrits par le Lagrangien de la QCD des équations 4 qui suivent.

$$\mathcal{L}_{QCD} = \sum_q \bar{\psi}_{q,a} (i\gamma^\mu \partial_\mu \delta_{ab} - g_s \gamma^\mu t_{ab}^C \mathcal{A}_\mu^C - m_q \delta_{ab}) \psi_{q,b} - \frac{1}{4} \mathcal{F}_{\mu\nu}^C \mathcal{F}_C^{\mu\nu} \quad , \quad (4a)$$

$$\mathcal{F}_C^{\mu\nu} = \partial^\mu F_C^\nu - \partial^\nu F_C^\mu - gf^{C\beta\gamma} F_\beta^\mu F_\gamma^\nu \quad ; \quad (4b)$$

Ici, les indices répétés sont sommés. L'indice q représente les saveurs de quark, les indices a les trois couleurs, m_q la masse du quark, \mathcal{A}_μ^C les champs des gluons, C les huit gluons, ψ les champs des quarks, les γ^μ les matrices de Dirac et les t_{ab}^C les matrices génératrices du groupe SU(3). Le terme au centre de la parenthèse représente l'interaction quark-quark-gluon, alors que les deux autres incarnent le terme cinétique. Finalement, le terme $\mathcal{F}_{\mu\nu}^C \mathcal{F}_C^{\mu\nu}$ de l'équation 4a est responsable des auto-interactions à trois et quatre gluons et est détaillé à l'équation 4b [126], où il exprime les tenseurs de couleur et où les $f^{\alpha\beta\gamma}$ sont les constantes de structures fines et les α et β sont sommés sur les huit gluons.

Dans le Lagrangien de la QCD (équation 4a), le terme g_s apparaît. Ce g_s s'exprime aussi en fonction d'une quantité $\alpha_s = \frac{g_s^2}{4\pi}$. Ce terme correspond à la constante de couplage de la QCD, un des paramètres fondamentaux du MS. Une particularité de la constante de couplage α_s est qu'elle diminue à hautes énergies (aux petites distances). Ce comportement particulier de la force forte se nomme « liberté asymptotique » et est illustré à la Figure 7. Ainsi, aux petites distances, comme à l'intérieur des noyaux atomiques, les interactions de la force forte sont très faibles.

Voyons un autre comportement particulier de la force forte qui est observable à basse énergie. Si nous essayons de séparer des quarks à l'intérieur d'un proton, il en résulte une production de paires par l'échange d'un gluon, tel qu'illustré à gauche de la Figure 8. En effet, on ne trouve jamais de quarks libres (ni de gluons libres) dans la nature. Ceux-ci sont

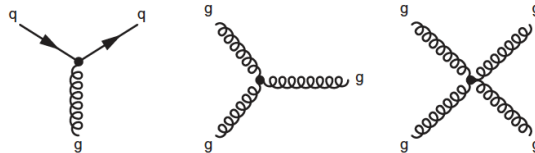


Figure 6. Les vertex principaux de la chromodynamique quantique (QCD). À gauche, le vertex qui permet les interactions entre quark-quark-gluon. Au centre, les interactions à trois et à droite, à quatre gluons [148].

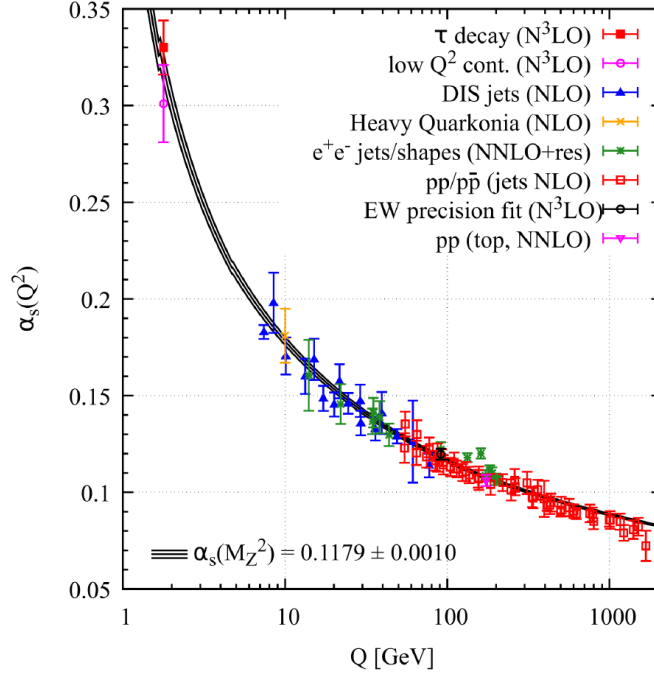


Figure 7. Mesure de la quantité α_s en fonction de l'échelle d'énergie Q (GeV). Cette figure illustre la liberté asymptotique, caractéristique de la force forte. À mesure que l'énergie augmente, la constante de couplage diminue et la force forte devient de plus en plus faible [126].

toujours présents en groupes de deux ou plus. Si nous essayons de les séparer, il en résulte toujours une production de paires, car cette production est plus bénéfique énergétiquement que de continuer à essayer de séparer les quarks. C'est ce qu'on appelle le *confinement*. Cette particularité, combinée au comportement de la constante α_s , fait en sorte qu'en collisionnant des hadrons (protons) à de très hautes énergies, il se produit une production de paires successives de quark-antiquark, tel que vu à la droite de la Figure 8. Ces quarks finissent par former des hadrons pour respecter le confinement et neutraliser la couleur. Ce processus est ce qu'on appelle l'hadronisation [148]. De plus, il arrive que les particules produites par les quarks et gluons initiaux soient situées dans un flot collimaté que l'on nomme un jet. Ces jets forment des cônes restreints de particules de couleur neutre créées dans la même direction que les partons (quarks et gluons) initiaux.

Au LHC, les collisions effectuées impliquent généralement des protons. Nous savons qu'il est toutefois impossible d'observer les partons les composant dans un état libre. Il faut alors mesurer les hadrons et les produits des désintégrations suite aux collisions des protons. Dans le cas de la production de jets, des algorithmes sont utilisés afin d'en reconstruire l'origine. Lorsque le quark à l'origine du jet n'est pas très énergétique, les produits de désintégration

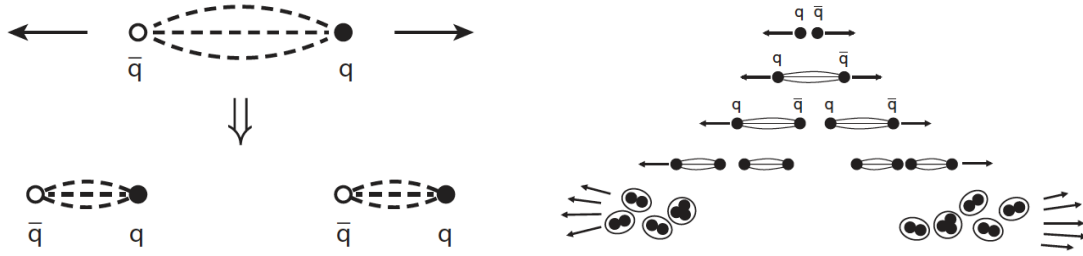


Figure 8. Processus de formation de deux jets issus de quarks. À cause du confinement des quarks et de la liberté asymptotique, une paire de quark-antiquark sera produite si on essaie de séparer des quarks (gauche) [148]. Aux hautes énergies, ce processus se produit de façon successive, résultant en une cascade de quarks (droite) [136].

auront tendance à être plus étendus dans l'espace. Les différentes composantes du jets seront séparées en plusieurs cônes de petits rayons R . Toutefois, l'énergie du centre de masse au LHC étant très élevée, les particules massives vont parfois être produites avec un « boost de Lorentz » dans le référentiel du laboratoire. Les produits de désintégration vont subir un « boost de Lorentz » dans la même direction et vont commencer à se fusionner pour former un seul jet collimaté. Ces produits sont alors reconstruits dans un seul jet ou cône de grand rayon R qu'on qualifie de « boosté ».

Les jets boostés peuvent se produire par la désintégration d'une particule massive, comme un boson W ou un quark top, tel qu'indiqué à la Figure 9. Ces particules massives ont un temps de vie très court et vont donc se désintégrer en deux ou trois autres particules. Si la particule initiale a une impulsion suffisamment élevée, les produits de sa désintégration sont alors reconstruits dans un seul jet à grand rayon R [140]. Pour bien identifier la particule à l'origine du jet, il faut alors inspecter sa sous-structure. Ce sujet sera développé plus en détails dans le Chapitre 3.



Figure 9. Production de jet boosté dans une collision proton-antiproton. Ici, la collision initiale produit un quark top qui se désintègre et produit divers jets. Dans le régime boosté à haute énergie, les jets vont parfois se fusionner pour former un seul grand jet [145].

L'étude des jets boostés s'avère particulièrement importante dans l'optique d'une recherche de nouvelle physique. Plusieurs modèles de nouvelle physique prédisent effectivement des

résonances lourdes qui pourraient se manifester dans les jets boostés [156]. Ces jets sont des événements complexes qui sont observés très souvent au LHC. Il est donc fort probable que des signes de nouvelle physique s'y cachent, d'où l'intérêt de leur étude.

1.1.6. La théorie électrofaible et le mécanisme de Higgs

La théorie électrofaible, développée majoritairement dans les années 1960 et fortement inspirée des travaux sur l'électromagnétisme, unit les interactions électromagnétique et nucléaire faible dans le modèle standard. Cette théorie part de l'hypothèse que ces deux interactions sont plutôt les symptômes d'une seule et même force [88]. Pour relier les interactions de la QED à la force faible et unir ces forces, il faut introduire de nouveaux bosons de jauge, le boson B_μ et les bosons W_μ^i ($i = 1,2,3$).

Le groupe de symétrie représentant la théorie électrofaible est le groupe $SU(2)_L \times U(1)_Y$. Ce groupe possède quatre générateurs, qui sont représentés par les bosons de jauge sans masse W_μ^i et B_μ . Le groupe $SU(2)_L$ possède 3 générateurs, qui sont associés aux W_μ^i . Le suffixe L fait référence aux fermions gauchers : les W_μ^i ne couplent qu'à ces derniers. La charge liée au groupe $SU(2)$ est l'isospin faible introduit précédemment, I_W . Quant à lui, le groupe $U(1)_Y$ est généré par le boson B_μ et conserve l'hypercharge, qui est définie comme : $Y = 2(Q_{em} + I_W^3)$, avec I_W^3 la troisième composante de l'isospin faible.

Cependant, dans les années 1960, seulement deux bosons physiques avaient alors été observés : le W^+ et le W^- avec les vertex de base des courants chargés définis à la section 1.1.4. C'est qu'il manquait un élément important dans le Lagrangien pour que la théorie soit complète. En effet, dans la nature, il n'existe qu'un seul boson sans masse pour les interactions électrofaibles, soit le photon. Il était donc nécessaire de faire acquérir une masse aux bosons W et Z , ainsi qu'aux fermions, dans le nouveau Lagrangien. Pour faire apparaître cette masse, on introduit un champ scalaire : le champ de Higgs. Une nouvelle interaction apparaît alors entre ce champ scalaire et les fermions, qu'on appelle « potentiel de Yukawa ». Ce potentiel permet le couplage des fermions à ce nouveau champ scalaire. La clé pour faire acquérir une masse aux particules est alors de demander une brisure de symétrie spontanée. Par cette brisure, le champ de Higgs acquiert une valeur énergétique du vide v , ou « vev ». C'est sous l'action de ce « vev » que les fermions et les bosons W et Z acquièrent une masse et que le photon demeure non massif. Ces masses sont définies aux équations 5 en fonction des constantes de couplages g et g' des groupes $SU(2)$ et $U(1)$ et de la valeur du « vev » [126].

$$M_W = \frac{gv}{2} \quad , \quad (5a)$$

$$M_Z = \frac{\sqrt{g^2 + g'^2}v}{2} \quad , \quad (5b)$$

$$M_\gamma = 0 \quad ; \quad (5c)$$

En demandant l'invariance de jauge du groupe $SU(2)_L \times U(1)_Y$, ceci introduit le boson de jauge B_μ et sa relation à deux bosons physiques : le A_μ , qui représente le photon, et le Z_μ , qui représente le boson Z^0 responsable des courants neutres dans l'interaction faible. Les relations entre ces bosons de jauge et les bosons physiques sont décrites aux équations 6. Le développement du Lagrangien de cette théorie permet aussi l'interaction entre les bosons intermédiaires et introduit les vertex à trois bosons ($W^+W^- \rightarrow \gamma/Z^0$) et d'autres à quatre bosons. Cette manoeuvre fait également apparaître un nouveau paramètre du MS : l'angle faible de mixage θ_W [126].

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad , \quad (6a)$$

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad , \quad (6b)$$

$$Z_\mu = -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \quad ; \quad (6c)$$

Finalemnt, le Lagrangien de la théorie électrofaible est décrit dans l'équation 7. Les ψ_i représentent les champs des leptons, les σ les matrices de Pauli, le Ψ le champ scalaire du Higgs, m la masse non-renormalisée (ou observable) du Higgs et $D_\mu = \partial^\mu + ig\sigma^i W_\mu^i/2 + ig'Y B_\mu/2$. En y ajoutant le Lagrangien de la QCD de l'équation 4a, on obtient alors le Lagrangien complet du modèle standard.

$$\begin{aligned} \mathcal{L}_{leptons} + \mathcal{L}_{EWK} + \mathcal{L}_{Higgs} = & \sum_i \bar{\psi}_i (i\gamma^\mu \partial_\mu - m_i) \psi_i \\ & + \frac{g}{2} \sum_{\phi_L} \bar{\phi}_L (\sigma_{1,2,3} W_\mu^{1,2,3}) \phi_L - \frac{g'}{2} \sum_i \bar{\psi}_i (Y_W B_\mu) \psi_i \\ & + (D_\mu \Psi)^\dagger (D^\mu \Psi) - m^2 \Psi^\dagger \Psi + \lambda (\Psi^\dagger \Psi)^2 \quad ; \end{aligned} \quad (7)$$

La découverte du boson de Higgs en 2012 par les expériences ATLAS et CMS a permis de confirmer la théorie électrofaible [18, 60]. En plus de mener à l'obtention du

prix Nobel de physique en 2013, cette découverte a notamment mené à l'obtention du *Breakthrough Prize in Fundamental Physics* décerné à Fabiola Gianotti, ainsi qu'à d'autres porte-paroles du CERN, pour son grand rôle de leadership lié à la découverte du Higgs [79]. Fabiola Gianotti a depuis été nommée directrice générale du CERN. Elle est d'ailleurs la première femme à occuper ce poste. La découverte du Higgs a certainement marqué l'imaginaire collectif et propulsé la physique des particules dans la culture générale. Le Higgs est aussi devenu un outil puissant pour étudier le modèle standard au LHC et pour explorer les possibles manifestations de physique au-delà de ce modèle.

1.2. La physique au-delà du modèle standard

Dans la section précédente, nous avons peint le tableau complet du modèle standard. Toutefois, derrière ces apparences lisses se dissimulent quelques défauts qui sont impossibles à ignorer. En effet, bien que le MS décrive d'une façon juste les particules connues et qu'il ait été confirmé expérimentalement [32], il ne décrit que 5% de la densité d'énergie de l'Univers. Ceci signifie que la quasi-totalité de ce qui compose notre Univers nous est toujours inconnue. Toutefois, nous avons plusieurs pistes de solutions pour résoudre ce problème. Dans la prochaine section, nous toucherons d'abord aux principales lacunes que possède le MS. Puis, nous explorerons quelques-uns des nombreux modèles de nouvelle physique qui existent et regarderons quel est l'état actuel des recherches.

1.2.1. Les limites du modèle standard

D'abord, les dernières années de recherche en physique des particules et en astrophysique nous ont amené les preuves de l'existence d'un différent type de matière qui ne serait pas décrit par le MS. Les premiers soupçons sont apparus dans les années 1930. À cette époque, les scientifiques étudiaient les courbes de rotation des galaxies pour mesurer la masse des amas de galaxies. L'estimation de la masse était effectuée en observant la quantité de lumière émise par l'amas. Toutefois, les mesures indiquaient qu'il manquait une énorme quantité de matière pour permettre aux galaxies de ne pas s'effondrer sur elles-mêmes [65]. L'hypothèse de la présence d'une matière dans les amas qui ne produit pas de lumière a alors été formulée, d'où l'appellation « matière sombre ». C'est l'astronome américaine Vera Rubin qui consolida cette hypothèse grâce à ses mesures de vitesse de rotation de galaxies. Ses travaux réussirent à convaincre la communauté scientifique de l'époque de l'existence de la matière sombre [80]. Les plus récentes estimations évaluent maintenant la quantité de matière sombre dans l'Univers à 27% [80].

Si la matière ordinaire et la matière sombre comptent pour seulement 32% de la densité de l'Univers, qu'en est-il du 68% restant ? La réponse à cette question est arrivée beaucoup plus tard, à la fin des années 1990 lorsque les astronomes ont mesuré la vitesse d'éloignement des galaxies. Leurs mesures ont montré que les galaxies s'éloignaient de plus en plus rapidement, ce qui signifie que notre univers est en expansion [80]. Pour que cette expansion se produise et s'accélère, il faut alors qu'une force quelconque agisse. C'est cette force invisible que nous avons nommé énergie sombre. Cette énergie représente le 68% de la densité restante et n'est toujours pas bien comprise aujourd'hui [120]. Ainsi, seulement 5% de la densité d'énergie totale de l'univers serait actuellement décrite par le MS.

De plus, l'ajout du boson de Higgs dans la théorie du MS n'a pas seulement résolu la question des masses. Il a aussi créé certains problèmes. Un de ces problèmes est généré par les interactions du boson de Higgs avec, par exemple, le quark top comme à la Figure 10. Ces interactions font apparaître plusieurs boucles possibles dans les diagrammes d'interactions qui font diverger la masse du Higgs à l'échelle de Planck, soit autour de 10^{16} GeV. Pourtant, la masse mesurée du boson de Higgs se situe plutôt autour de 125 GeV [23]. C'est ce qu'on appelle le problème de la hiérarchie. Ce problème causerait également le grand écart d'intensité entre la force gravitationnelle et la force nucléaire faible.

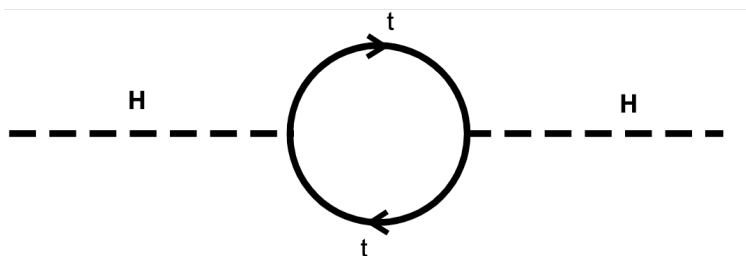


Figure 10. Diagramme de boucle entre le boson de Higgs et le quark top [158].

Une autre observation récente a également chamboulé notre compréhension des neutrinos. Plusieurs expériences, dont la collaboration SNO située à Sudbury, ont mené à la découverte de l'oscillation des neutrinos. En effet, les observations montrent qu'entre le moment où les neutrinos sont émis par le Soleil et qu'ils arrivent sur Terre, leur saveur leptonique initiale peut changer : les neutrinos « oscillent » de saveur [144]. Par exemple, un neutrino électronique (ν_e) émis par le Soleil pourrait être détecté comme un neutrino muonique (ν_μ) sur Terre. Cette découverte s'est avérée cruciale, car ce comportement particulier implique que les neutrinos sont massifs et non sans masse tel que le MS le prédit. Depuis, les physiciens et physiciennes s'évertuent à comprendre l'origine de la masse des neutrinos

et à mesurer ces dernières.

Un mot sur les paramètres libres du MS s'avère de mise. Nous avons touché à quelques-uns de ces paramètres à la section précédente, comme l'angle électrofaible θ_W et la constante de couplage de la force forte α_s . Au total, le MS possède en fait 19 de ces paramètres libres, qui incluent entre autres les masses des fermions et les constantes de couplage des trois différentes forces. Ces paramètres libres ne peuvent être déterminés sans passer par l'expérience. Ce fait ne représente pas un problème en soit, mais rend la théorie du MS un peu moins robuste comme le choix de ces paramètres demeure plus ou moins arbitraire.

Parmi les nombreux autres mystères non résolus, mentionnons rapidement que la gravité n'est pas décrite par le MS. Une théorie qui n'inclut pas les quatre forces fondamentales ne peut prétendre être complète, car elle ne peut décrire l'ensemble des interactions possibles entre les particules. Les scientifiques sont toujours à la recherche d'une « théorie du tout » qui permettrait d'unifier toutes les forces fondamentales, à la manière de la théorie électrofaible [94]. Cette théorie demeure toutefois indéfinie et son élaboration sera sûrement le Graal de la physique moderne.

Tous ces questionnements ont suscité l'intérêt des équipes de recherche en expérimentation et théorie en physique des particules durant les dernières années. Plusieurs ont tenté et tentent toujours d'élaborer des théories au-delà du MS et des expériences pour essayer de détecter des manifestations de cette nouvelle physique.

1.2.2. Les théories au-delà du modèle standard

Il existe un nombre impressionnant de théories de nouvelle physique. Devant les nombreux problèmes qui demeurent à résoudre, les scientifiques des particules font preuve d'une créativité sans limites pour élaborer de nouvelles théories. Certaines se contentent de modifier légèrement un aspect du MS, alors que d'autres ajoutent un tout nouvel ensemble de particules. Ces théories prétendent même que le MS serait une limite de basse énergie d'un modèle plus grand qui demeure à découvrir [102]. Nous explorerons quelques-uns de ces modèles puis, nous résumerons les limites des recherches de nouvelle physique réalisées au LHC.

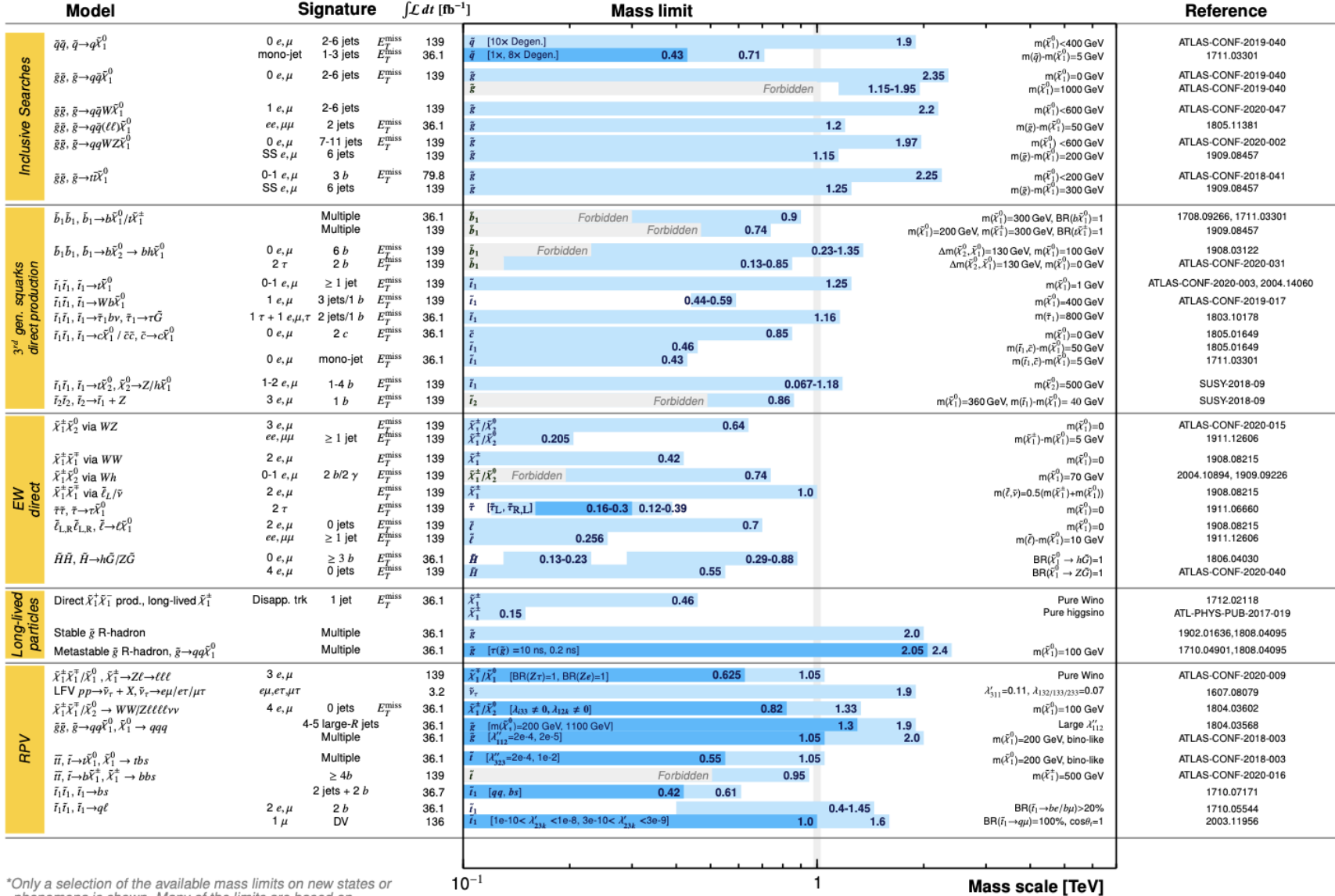
Une des théories principales qui a été étudiée durant les dernières années est la supersymétrie (SUSY). Cette théorie introduit une symétrie supplémentaire entre les fermions et les bosons en ajoutant des partenaires supersymétriques à chaque particule.

Ces partenaires auraient les mêmes caractéristiques que leurs particules associées, mais avec un spin qui diffère de $1/2$. Le modèle supersymétrique permettrait d'ajouter des corrections radiatives grâce, par exemple, aux interactions Higgs-stop (superpartenaire du quark top) qui balanceraient les autres diagrammes qui font diverger la masse du Higgs comme à la Figure 10 [39]. Il permettrait aussi de faire converger les constantes de couplage des différentes forces à l'échelle de Planck et serait donc un pas vers la grande unification des forces fondamentales [71]. De nombreuses recherches au LHC sont dédiées à la poursuite de preuves du modèle supersymétrique. Ces expériences ont réussi à établir des limites de plus en plus précises sur les masses des particules supersymétriques, sans toutefois en prouver l'existence [33]. Les limites sur les recherches de SUSY sont affichées à la Figure 11.

Il existe également une multitude de modèles de nouvelle physique s'intéressant à la matière sombre. Parmi les deux modèles les plus populaires se trouvent les axions et les WIMP (*Weakly Interactive Massive Particle*). Les axions seraient des bosons électriquement neutres et de très petite masse. Ils interagiraient avec les champs magnétiques élevés, leur permettant de se transformer en photon (voir le chapitre 91 de la référence [126]). Ceci les rend détectables par l'effet Primakoff, c'est-à-dire la production de mésons par des photons très énergétiques. Plusieurs expériences du LHC tentent de découvrir les axions, dont les collaborations CAST et OSQAR [52] [124]. Tout comme pour les particules supersymétriques, les axions demeurent cependant introuvables.

L'autre modèle très populaire des WIMP prédit des particules massives qui n'interagissent que très rarement avec la matière ordinaire [83]. Ce comportement les rend d'autant plus difficiles à détecter. La façon la plus commune de procéder est donc de construire de grands détecteurs qui vont parfois jusqu'à plusieurs tonnes de matériaux pour maximiser les chances d'interaction entre ce type de matière sombre et la matière ordinaire. Il est aussi primordial de placer ces détecteurs à des endroits où se trouve le moins de bruit de fond possible. Par exemple, le détecteur PICO [133] se trouve 2km sous Terre dans une mine en Ontario. Il existe plusieurs expériences qui se penchent sur le cas des WIMP autre que PICO, dont les expériences Xenon 1T [163] et CDMS [146].

Finalement, un des autres candidats à la fois pour la matière sombre que pour la supersymétrie sont les particules de Majorana. L'attribut de ces nouvelles particules est qu'elles sont leur propre antiparticule, à la différence qu'elles se doivent aussi d'être des fermions. Ceci implique alors que ces particules sont neutres électriquement. Un des suspects qui répond à ces critères est une des particules déjà présentes



*Only a selection of the available mass limits on new states or phenomena is shown. Many of the limits are based on simplified models, c.f. refs. for the assumptions made.

Figure 11. Limites sur les masses de diverses particules supersymétriques par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [33].

dans le MS : le neutrino [5]. Les neutrinos Majorana pourraient possiblement expliquer les masses non-nulles des neutrinos et leurs oscillations de saveur. La preuve de l'existence des neutrinos Majorana se trouverait d'ailleurs dans l'observation de la double désintégration bêta sans neutrinos, quoique qu'il ne s'agirait pas nécessairement d'une garantie [91]. Un autre des candidats possibles pour les particules Majorana serait le neutralino, le superpartenaire des neutrinos.

Les groupes de recherches de l'expérience ATLAS au LHC sur les particules exotiques ont établi les limites sur ces nouveaux modèles, ainsi que plusieurs autres qui pourraient résoudre les problèmes du MS présentés à la section précédente. Les Figures 12 et 13 résument les limites obtenues sur ces recherches de nouvelle physique. Les équipes de recherche du LHC redoublent incessamment de créativité et d'ingéniosité pour confectionner de nouveaux modèles de nouvelle physique, mais ces efforts n'ont toujours pas porté fruit. La plupart des recherches de nouvelle physique requiert d'ailleurs d'atteindre de très hautes énergies pour en observer les manifestations. Bien que l'énergie actuelle du centre de masse au LHC soit élevée, elle n'atteint vraisemblablement pas les échelles de la physique au-delà du MS, puisque nous n'avons toujours pas observé cette nouvelle physique. Le LHC a donc entrepris une nouvelle ère de rénovations pour augmenter sa limite d'énergie. Les prochaines années risquent d'être décisives pour la recherche de nouvelle physique et permettront d'établir de nouvelles limites sur ces modèles, et peut-être même d'en découvrir les signes.

1.2.3. Conclusion

Nous avons vu les problèmes du MS ainsi que quelques exemples de nouvelle physique hypothétique formulée par la communauté physicienne. Mais une seule chose est commune à tous ces modèles ingénieux : peu importe les efforts déployés, ils demeurent introuvables. Ce fait irréfutable nous oblige à nous remettre en question. Notre façon d'approcher la recherche de nouvelle physique fonctionne-t-elle ? Si un modèle n'existe pas, suffit-il de créer nouveau modèle après nouveau modèle en espérant trouver le bon ? Nos outils actuels sont-ils suffisants pour ces recherches ? Heureusement, les récents développements en intelligence artificielle pourraient nous aider dans cette quête. L'expérience ATLAS au LHC intègre de plus en plus cet outil à sa recherche de physique au-delà du MS. Le Chapitre 2 s'intéressera à l'expérience ATLAS pour introduire cette expérience de pointe, puis les Chapitres 3 et 4 développeront sur le sujet de ce mémoire, qui s'inscrit dans cette nouvelle approche de recherche de nouvelle physique.

ATLAS Exotics Searches* - 95% CL Upper Exclusion Limits

Status: May 2019

ATLAS Preliminary

$$\int \mathcal{L} dt = (3.2 - 139) \text{ fb}^{-1}$$

$$\sqrt{s} = 8, 13 \text{ TeV}$$

Model	ℓ, γ	Jets [†]	E_T^{miss}	$\int \mathcal{L} dt [\text{fb}^{-1}]$	Limit	Reference	
Extra dimensions	ADD $G_{KK} + g/q$	$0 e, \mu$	1-4 j	Yes	36.1	M_D 7.7 TeV	$n = 2$
	ADD non-resonant $\gamma\gamma$	2γ	-	-	36.7	M_S 8.6 TeV	$n = 3$ HLZ NLO
	ADD QBH	-	2 j	-	37.0	M_{th} 8.9 TeV	$n = 6$
	ADD BH high Σp_T	$\geq 1 e, \mu$	$\geq 2 j$	-	3.2	M_{th} 8.2 TeV	$n = 6, M_D = 3 \text{ TeV, rot BH}$
	ADD BH multijet	-	$\geq 3 j$	-	3.6	M_{th} 9.55 TeV	$n = 6, M_D = 3 \text{ TeV, rot BH}$
	RS1 $G_{KK} \rightarrow \gamma\gamma$	2γ	-	-	36.7	G_{KK} mass 4.1 TeV	$k/\overline{M}_{Pl} = 0.1$
	Bulk RS $G_{KK} \rightarrow WW/ZZ$	multi-channel	-	-	36.1	G_{KK} mass 2.3 TeV	$k/\overline{M}_{Pl} = 1.0$
	Bulk RS $G_{KK} \rightarrow WW \rightarrow qqqq$	$0 e, \mu$	2 J	-	139	G_{KK} mass 1.6 TeV	$k/\overline{M}_{Pl} = 1.0$
	Bulk RS $g_{KK} \rightarrow tt$	$1 e, \mu$	$\geq 1 b, \geq 1 J/2 j$	Yes	36.1	g_{KK} mass 3.8 TeV	$\Gamma/m = 15\%$
	2UED / RPP	$1 e, \mu$	$\geq 2 b, \geq 3 j$	Yes	36.1	KK mass 1.8 TeV	Tier (1,1), $\mathcal{B}(A^{(1,1)} \rightarrow \tau\tau) = 1$
Gauge bosons	SSM $Z' \rightarrow \ell\ell$	$2 e, \mu$	-	-	139	Z' mass 5.1 TeV	
	SSM $Z' \rightarrow \tau\tau$	2τ	-	-	36.1	Z' mass 2.42 TeV	
	Leptophobic $Z' \rightarrow bb$	-	2 b	-	36.1	Z' mass 2.1 TeV	
	Leptophobic $Z' \rightarrow tt$	$1 e, \mu$	$\geq 1 b, \geq 1 J/2 j$	Yes	36.1	Z' mass 3.0 TeV	$\Gamma/m = 1\%$
	SSM $W' \rightarrow \ell\nu$	$1 e, \mu$	-	Yes	139	W' mass 6.0 TeV	
	SSM $W' \rightarrow \tau\nu$	1τ	-	Yes	36.1	W' mass 3.7 TeV	
	HVT $V' \rightarrow WZ \rightarrow qqqq$ model B	$0 e, \mu$	2 J	-	139	V' mass 3.6 TeV	$g_V = 3$
	HVT $V' \rightarrow WH/ZH$ model B	multi-channel	-	-	36.1	V' mass 2.93 TeV	$g_V = 3$
	LRSB $W_R \rightarrow tb$	multi-channel	-	-	36.1	W_R mass 3.25 TeV	
	LRSB $W_R \rightarrow \mu N_R$	2μ	1 J	-	80	W_R mass 5.0 TeV	$m(N_R) = 0.5 \text{ TeV, } g_L = g_R$
CI	CI $qqqq$	-	2 j	-	37.0	Λ 21.8 TeV	η_{LL}
	CI $\ell\ell qq$	$2 e, \mu$	-	-	36.1	Λ 40.0 TeV	η_{LL}
	CI $tttt$	$\geq 1 e, \mu$	$\geq 1 b, \geq 1 j$	Yes	36.1	Λ 2.57 TeV	$ C_{4t} = 4\pi$
DM	Axial-vector mediator (Dirac DM)	$0 e, \mu$	1-4 j	Yes	36.1	m_{med} 1.55 TeV	$g_a = 0.25, g_s = 1.0, m(\chi) = 1 \text{ GeV}$
	Colored scalar mediator (Dirac DM)	$0 e, \mu$	1-4 j	Yes	36.1	m_{med} 1.67 TeV	$g = -1.0, m(\chi) = 1 \text{ GeV}$
	$VV\chi\chi$ EFT (Dirac DM)	$0 e, \mu$	1 J, $\leq 1 j$	Yes	3.2	M_* 700 GeV	$m(\chi) < 150 \text{ GeV}$
	Scalar reson. $\phi \rightarrow \chi\chi$ (Dirac DM)	$0-1 e, \mu$	1 b, 0-1 J	Yes	36.1	m_ϕ 3.4 TeV	$\gamma = 0.4, \lambda = 0.2, m(\chi) = 10 \text{ GeV}$
LQ	Scalar LQ 1 st gen	$1, 2 e$	$\geq 2 j$	Yes	36.1	LQ mass 1.4 TeV	$\beta = 1$
	Scalar LQ 2 nd gen	$1, 2 \mu$	$\geq 2 j$	Yes	36.1	LQ mass 1.56 TeV	$\beta = 1$
	Scalar LQ 3 rd gen	2τ	2 b	-	36.1	LQ_3^* mass 1.03 TeV	$\mathcal{B}(LQ_3^* \rightarrow b\tau) = 1$
	Scalar LQ 3 rd gen	$0-1 e, \mu$	2 b	Yes	36.1	LQ_3^* mass 970 GeV	$\mathcal{B}(LQ_3^* \rightarrow \tau\tau) = 0$
	Heavy quarks	VLQ $TT \rightarrow Ht/Zt/Wb + X$	multi-channel	-	-	36.1	T mass 1.37 TeV
VLQ $BB \rightarrow Wt/Zb + X$		multi-channel	-	-	36.1	B mass 1.34 TeV	SU(2) doublet
VLQ $T_{5/3} T_{5/3} / T_{5/3} \rightarrow Wt + X$		$2(SS) \geq 3 e, \mu \geq 1 b, \geq 1 j$	Yes	36.1	$T_{5/3}$ mass 1.64 TeV	$\mathcal{B}(T_{5/3} \rightarrow Wt) = 1, c(T_{5/3} Wt) = 1$	
VLQ $Y \rightarrow Wb + X$		$1 e, \mu$	$\geq 1 b, \geq 1 j$	Yes	36.1	Y mass 1.85 TeV	$\mathcal{B}(Y \rightarrow Wb) = 1, c_R(Wb) = 1$
VLQ $B \rightarrow Hb + X$		$0 e, \mu, 2 \gamma$	$\geq 1 b, \geq 1 j$	Yes	79.8	B mass 1.21 TeV	$\kappa_B = 0.5$
VLQ $QQ \rightarrow WqWq$		$1 e, \mu$	$\geq 4 j$	Yes	20.3	Q mass 690 GeV	
Excited fermions	Excited quark $q^* \rightarrow qg$	-	2 j	-	139	q^* mass 6.7 TeV	only u^* and d^* , $\Lambda = m(q^*)$
	Excited quark $q^* \rightarrow q\gamma$	1γ	1 j	-	36.7	q^* mass 5.3 TeV	only u^* and d^* , $\Lambda = m(q^*)$
	Excited quark $b^* \rightarrow bg$	-	1 b, 1 j	-	36.1	b^* mass 2.6 TeV	
	Excited lepton ℓ^*	$3 e, \mu$	-	-	20.3	ℓ^* mass 3.0 TeV	$\Lambda = 3.0 \text{ TeV}$
	Excited lepton ν^*	$3 e, \mu, \tau$	-	-	20.3	ν^* mass 1.6 TeV	$\Lambda = 1.6 \text{ TeV}$
Other	Type III Seesaw	$1 e, \mu$	$\geq 2 j$	Yes	79.8	N^0 mass 560 GeV	$m(W_R) = 4.1 \text{ TeV, } g_L = g_R$
	LRSB Majorana ν	2μ	2 j	-	36.1	N_R mass 3.2 TeV	DY production
	Higgs triplet $H^{\pm\pm} \rightarrow \ell\ell$	$2, 3, 4 e, \mu$ (SS)	-	-	36.1	$H^{\pm\pm}$ mass 870 GeV	$\mathcal{B}(H^{\pm\pm} \rightarrow \ell\tau) = 1$
	Higgs triplet $H^{\pm\pm} \rightarrow \ell\tau$	$3 e, \mu, \tau$	-	-	20.3	$H^{\pm\pm}$ mass 400 GeV	DY production, $ q = 5e$
	Multi-charged particles	-	-	-	36.1	multi-charged particle mass 1.22 TeV	DY production, $ g = 1g_D, \text{spin } 1/2$
	Magnetic monopoles	-	-	-	34.4	monopole mass 2.37 TeV	

*Only a selection of the available mass limits on new states or phenomena is shown.

†Small-radius (large-radius) jets are denoted by the letter j (J).

$\sqrt{s} = 8 \text{ TeV}$ $\sqrt{s} = 13 \text{ TeV}$ partial data $\sqrt{s} = 13 \text{ TeV}$ full data

10⁻¹ 1 10 Mass scale [TeV]

Figure 12. Limites sur les masses des nouvelles particules par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [30].

ATLAS Long-lived Particle Searches* - 95% CL Exclusion

Status: July 2019

ATLAS Preliminary

$$\int \mathcal{L} dt = (18.4 - 36.1) \text{ fb}^{-1} \sqrt{s} = 8, 13 \text{ TeV}$$

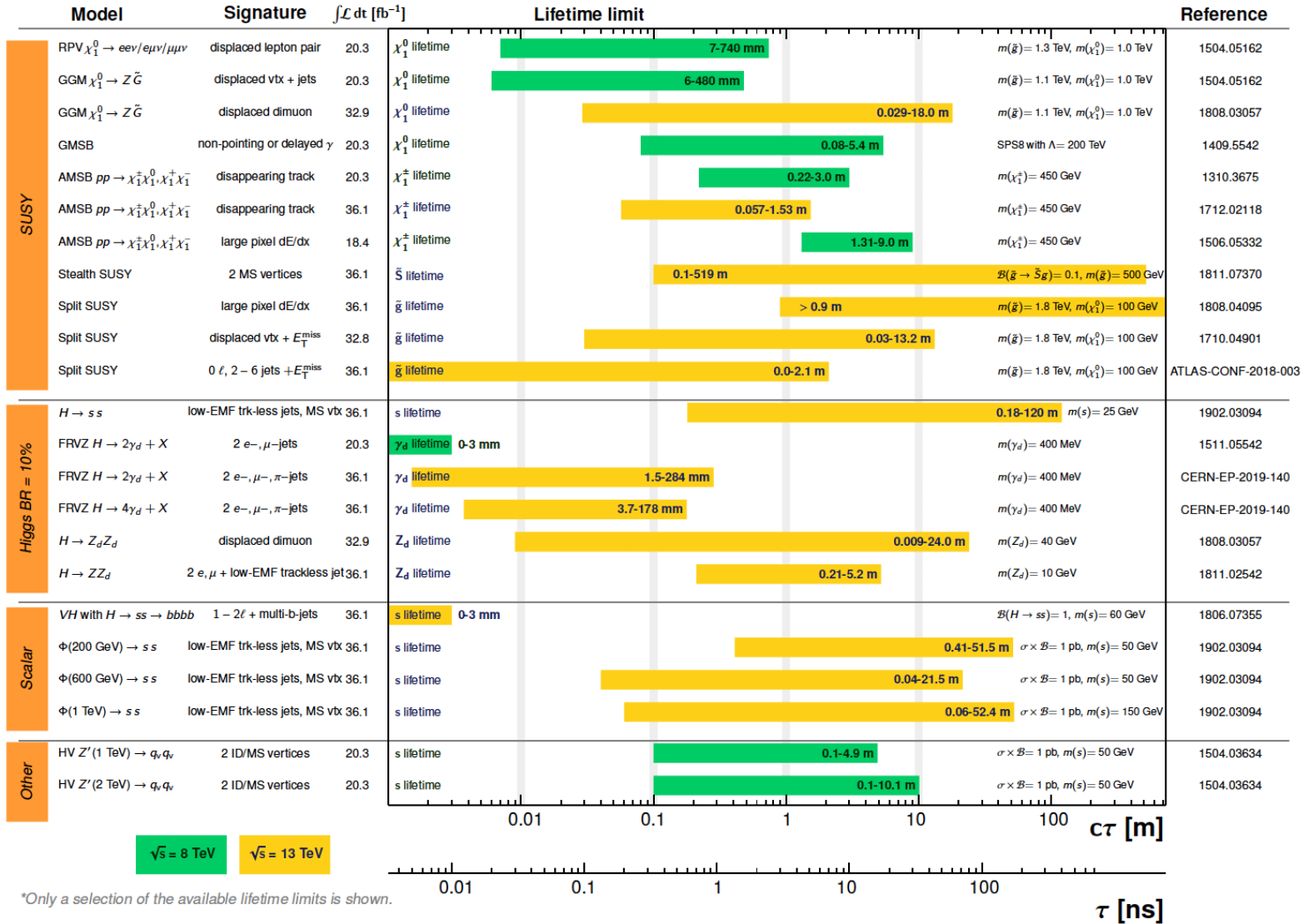


Figure 13. Limites sur les temps de vie des nouvelles particules par l'expérience ATLAS. Les valeurs sont exclues à un degré de confiance de 95% [30].

Chapitre 2

Le Grand collisionneur de hadrons et l'expérience ATLAS

Si le modèle standard a été confirmé avec une aussi bonne précision, c'est que les physiciens et physiciennes ont su allier la théorie à l'expérimentation. Un des outils expérimentaux qui a révolutionné l'étude des particules subatomiques est l'accélérateur de particules. Cet instrument combine les effets de champs magnétiques et électriques pour accélérer des particules à une vitesse près de celle de la lumière et pour les faire entrer en collision ensemble ou avec une cible [80]. Un de ces appareils est situé au complexe de l'Organisation européenne pour la recherche nucléaire (CERN). On nomme cet accélérateur le Grand collisionneur de hadrons. Le CERN est issu d'une collaboration internationale qui regroupe environ 60 pays et 175 institutions [56]. Sa mission est de comprendre la matière qui forme l'Univers, mais la recherche au CERN a aussi mené à plusieurs avancées technologiques utiles à la société. C'est d'ailleurs à cet endroit que le *World Wide Web* a été développé [55]. Dans le chapitre qui suit, nous décrirons cette expérience de grande envergure. Nous étudierons également le fonctionnement du détecteur ATLAS, une des quatre principales expériences du LHC et celle sur laquelle le présent mémoire se concentre. Puis, nous verrons comment les événements sont reconstruits à partir des données enregistrées dans ce détecteur.

2.1. Le Grand collisionneur de hadrons

Le LHC est un énorme accélérateur circulaire situé entre la Suisse et la France. Son anneau principal, dans lequel les particules sont accélérées à leur vitesse finale, est situé dans un tunnel 100 m sous terre [54] et fait 27 km de circonférence [73]. Le LHC a été construit au CERN dans l'ancien tunnel du Grand collisionneur électron-positron (LEP), qui a terminé ses activités en 2000 [15]. La Figure 14 présente une vue aérienne du LHC et la Figure

15 présente un schéma de son complexe d'accélération. Les scientifiques du LHC étudient plusieurs types de collisions. Les collisions proton-proton sont les plus communes, mais nous étudions aussi les collisions proton-plomb, plomb-plomb et xénon-xénon.



Figure 14. Une vue aérienne du complexe du CERN montrant l'anneau de 27 km de circonférence du Grand collisionneur de hadrons [51].

Le LHC exploite une des lois les plus célèbres de la physique, soit $E = mc^2$. En collisionnant des particules à une certaine énergie E , nous pouvons créer des particules de masse m . Pour produire des particules de masses élevées de l'ordre du TeV/c^2 ($\text{TeV} = 10^{12}$ eV), il faut alors atteindre des énergies très élevées. L'énergie actuelle dans le système du centre de masse \sqrt{s} au LHC, soit l'énergie dans le référentiel du laboratoire des deux faisceaux qui entrent en collision, est de 13 TeV. Ceci équivaut à 6.5 TeV par faisceau. Cette énergie de collision correspond à une vitesse des protons presque égale à celle de la lumière.

Au tout début de ses activités, en 2011-2012, le LHC opérait à une énergie \sqrt{s} de 7 et 8 TeV. Nous appelons cette période la Période 1, ou *Run 1*. C'est de 2015 à 2018, durant la *Run 2*, que le LHC opérait à 13 TeV [26]. Après les mises à niveau en cours, le LHC pourrait atteindre 14 TeV lors de sa reprise d'activités prévue pour 2022 [44].

2.1.1. L'accélération

Différentes étapes de préparation sont nécessaires afin d'amener les protons à leur vitesse finale de collision. Premièrement, les protons sont extraits d'atomes d'hydrogène ionisés, puis envoyés dans un premier accélérateur linéaire (Linac 2 dans la Figure 15) pour leur faire atteindre une énergie de 50 MeV. Ensuite, les protons passent dans deux proton-synchrotrons (Booster et PS), qui amènent respectivement leur énergie à 1.4 et 25 GeV. Finalement, les protons transitent dans le super proton-synchrotron (SPS) pour atteindre une énergie de 450 GeV, avant d'être envoyés dans l'anneau principal pour subir leur accélération finale [73].

CERN's Accelerator Complex

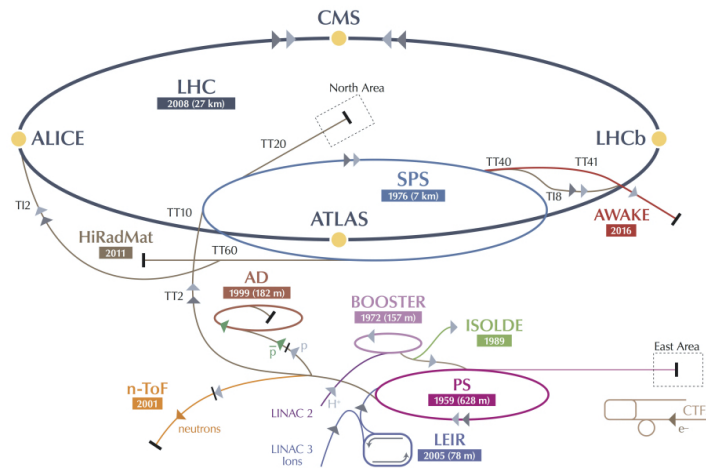


Figure 15. Un schéma du complexe d'accélération qui indique toutes les étapes d'injection des protons dans l'anneau final [92].

Le LHC utilise deux systèmes pour accélérer les protons à 6.5 TeV : les cavités radiofréquences (RF) et les aimants supraconducteurs. Les cavités RF sont des cavités dans lesquelles un champ électrique résonant oscille. Les protons sont envoyés à l'intérieur de ces cavités par paquets lorsque le champ électrique pointe dans leur direction du mouvement. Pour maximiser le nombre de paquets envoyés, nous pouvons en envoyer en déphasage avec le champ. Ces paquets subissent alors une accélération ajustée selon leur déphasage pour qu'ils soient synchronisés à la vitesse d'injection désirée [73].

Les aimants supraconducteurs, formés de dipôles magnétiques, conservent les protons dans une trajectoire circulaire. Comme les protons ont une très grande vitesse, les aimants doivent être très puissants. Ces aimants produisent un champ électromagnétique de 8.33 T, et pour atteindre cette valeur, ceux-ci sont refroidis à une température de 1.4 K, ou $-271.75\text{ }^{\circ}\text{C}$ [139]. Au final, le système de production et d'accélération des protons produit et collisionne 2808 paquets de protons, et ce à toutes les 25 nanosecondes [73].

La principale limite technologique du LHC est liée au champ magnétique et à l'impulsion des particules. En effet, l'impulsion des particules est limitée par l'équation 8 [131]. Comme le champ magnétique B est limité par des contraintes technologiques, la seule façon d'atteindre des impulsions élevées est d'augmenter R , le rayon de l'anneau. C'est pourquoi le LHC a une si grande circonférence. Aussi, l'énergie du centre de masse des collisions sera également limitée à deux fois l'énergie du faisceau.

$$p = 0.3BR \quad . \quad (8)$$

2.1.2. La luminosité

La luminosité nous permet de déterminer le nombre de collisions se produisant en un point d'interaction. Le nombre de collision produit est décrit par l'équation 9 [126], en nombre de collisions par seconde, par cm^2 . Dans cette équation, f représente la fréquence de collision entre paquets (25 ns^{-1}), N_1 et N_2 sont les quantités de protons dans chaque paquet (10^{11} protons par paquet), qui au LHC sont égales, et σ_x et σ_y sont représentatifs de la taille du faisceau en x - y au point d'interaction (de l'ordre du μm). Comme la fréquence et le nombre de protons dans les collisions sont optimisés pour augmenter la luminosité, nous devons réduire la taille du faisceau. Nous y arrivons en focalisant les faisceaux au point d'interaction à l'aide d'aimants quadripolaires. Le LHC est construit pour atteindre une luminosité nominale de 10^{34} interactions/ cm^2s [73].

$$L = \frac{fN_1N_2}{4\pi\sigma_x\sigma_y} \quad ; \quad (9)$$

Si nous nous intéressons au nombre d'interactions d'un processus en particulier, il suffit de multiplier la luminosité par la section efficace du processus, en suivant l'équation 10 [73]. Ici, σ représente la section efficace, qu'on peut interpréter comme la probabilité d'interaction d'une particule par le processus en question.

$$N_{\text{événements}} = L\sigma_{\text{processus}} \quad ; \quad (10)$$

Finalement, si on veut connaître la luminosité durant une période donnée, par exemple le nombre d'événements total sur une année, il suffit d'intégrer l'équation 9 sur ladite période, tel que $\mathcal{L} = \int Ldt$ [126]. On appelle cette quantité la luminosité intégrée. La Figure 16 indique la luminosité intégrée de l'expérience ATLAS jusqu'à la fin de la Période 2 (*Run 2*). Depuis 2015, avec l'augmentation de l'énergie dans le système du centre de masse, le LHC a grandement augmenté sa luminosité. Le détecteur ATLAS ne peut toutefois enregistrer toutes les données produites, pour des raisons qui seront expliquées plus bas.

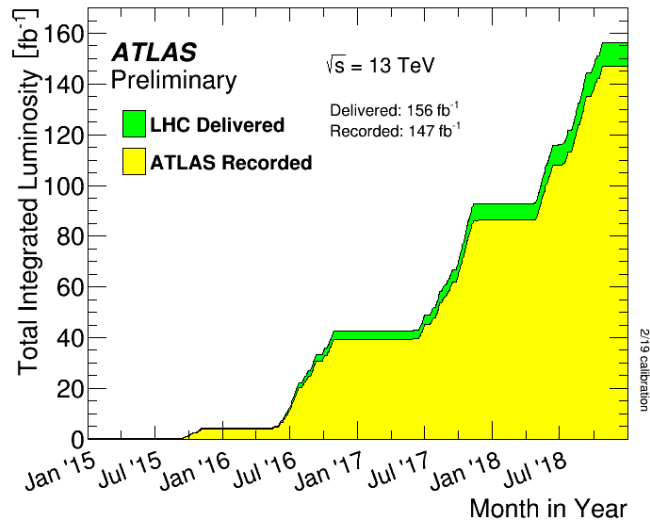


Figure 16. Luminosité intégrée de l’expérience ATLAS entre 2015 et 2018. Ici, la luminosité est exprimée en fb inverse, où $1 \text{ fb} = 10^{-39} \text{ cm}^2$ [26].

2.1.3. Les expériences

Le complexe du CERN abrite plusieurs expériences. Bien que le LHC soit l’anneau principal, de multiples détecteurs sont situés sur sa circonférence afin d’enregistrer les produits des collisions et réaliser diverses études. Les quatre expériences principales se nomment : *The Compact Muon Solenoid* (CMS) [59], *A Large Ion Collider Experiment* (ALICE) [9], *The Large Hadron Collider beauty* (LHCb) [111] et *A Toroidal LHC ApparatuS* (ATLAS) [17]. CMS se concentre sur l’étude du boson de Higgs, de la matière sombre et de la nouvelle physique, ALICE s’intéresse aux ions lourds pour comprendre la force forte et LHCb étudie la physique du quark bottom. ATLAS possède sensiblement les mêmes objectifs que CMS, mais diffère au niveau technique. Comme il s’agit du détecteur utilisé dans ce mémoire, nous détaillerons ce dernier dans les prochaines sections pour comprendre son fonctionnement.

2.2. Le détecteur ATLAS

Le détecteur ATLAS, situé en un point du pourtour du LHC, est tout aussi imposant que ce dernier. ATLAS mesure 44 mètres de long par 25 mètres de haut et pèse 7000 tonnes, soit l’équivalent du poids de la tour Eiffel [93]. Sa coupe transversale montrée à la Figure 17 laisse voir ses trois parties principales : le détecteur interne, les calorimètres et le spectromètre à muons. Ces trois systèmes utilisent des techniques différentes pour détecter les diverses particules produites lors des collisions. Nous explorerons d’abord le système

de coordonnées du détecteur, puis nous étudierons ces trois dispositifs de détection et le système d'acquisition de données.

2.2.1. Le système de coordonnées

Le détecteur ATLAS utilise une géométrie cylindrique : il est formé de couches en cylindres concentriques situées autour du faisceau. Son point central correspond au point d'interaction entre les paquets de protons. Dans le système de coordonnées en trois dimensions $(\hat{x}, \hat{y}, \hat{z})$, \hat{z} représente l'axe le long du faisceau, \hat{x} pointe vers le centre de l'anneau du LHC et \hat{y} vers le haut. Initialement, les protons arrivent selon l'axe \hat{z} et donc, toute leur impulsion est dirigée dans cette direction. Si nous suivons la loi de conservation de l'impulsion, ceci signifie que l'impulsion finale totale après la collision devra également être selon l'axe \hat{z} .

Cependant, dans un milieu relativiste comme au détecteur ATLAS, il nous est plus utile d'utiliser d'autres quantités que $(\hat{x}, \hat{y}, \hat{z})$ pour décrire les positions des particules. L'angle ϕ représente l'angle dans le plan (\hat{x}, \hat{y}) . L'impulsion dans ce plan est nommée impulsion transverse (p_T). Nous utilisons aussi la rapidité y , définie à l'équation 11a [161]. De plus, comme l'énergie de la collision est beaucoup plus grande que les masses produites, nous pouvons utiliser la pseudo-rapidité, définie à l'équation 11b [161]. Ici, θ représente

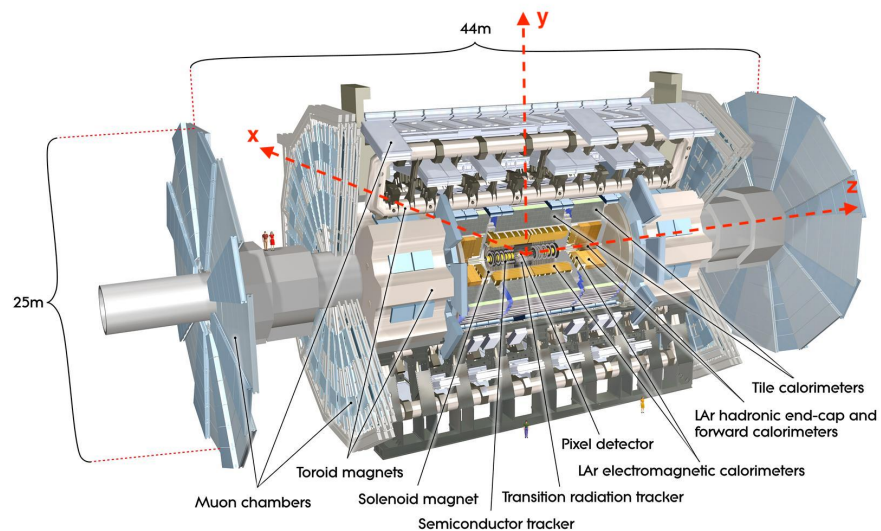


Figure 17. Coupe transversale du détecteur ATLAS. La Figure représente les trois systèmes différents de détection des particules : le détecteur interne, les calorimètres et le spectromètre à muons [130]. Les axes en rouge ont été ajoutés à l'image originale.

l'angle dans le plan (\hat{z}, \hat{y}) . Ces deux quantités nous sont utiles en milieu relativiste, car les différences en y et en η sont des quantités invariantes sous une transformation de Lorentz. Pour localiser les particules dans le détecteur, nous n'utiliserons donc plus $(\hat{x}, \hat{y}, \hat{z})$, mais plutôt (η, ϕ) , qui correspondent respectivement à un point dans le plan (\hat{z}, \hat{y}) et (\hat{x}, \hat{y}) . Une visualisation simplifiée de l'angle η est illustrée à la Figure 18.

$$y = \frac{1}{2} \log \left(\frac{E + p_z}{E - p_z} \right) \quad , \quad (11a)$$

$$\eta = -\log(\tan(0.5\theta)) \quad ; \quad (11b)$$

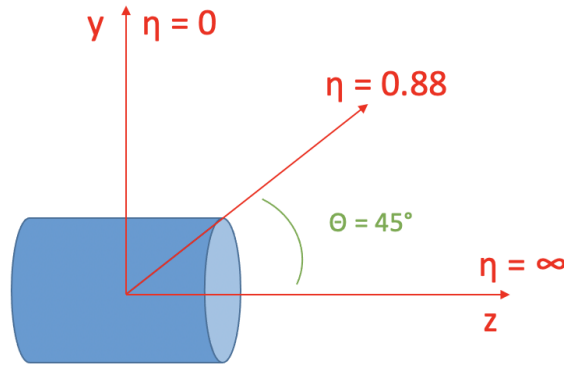


Figure 18. Représentation simplifiée de la variable η dans le détecteur ATLAS. Le détecteur ATLAS est imagé en bleu. L'axe \hat{z} correspond à la limite où $\eta = \infty$. L'axe \hat{y} correspond à une valeur de $\eta = 0$.

Maintenant que nous avons défini comment nous localisons les particules dans le détecteur, nous commencerons par regarder le système qui se trouve le plus près du point de collision, soit le détecteur interne.

2.2.2. Le détecteur interne

Le détecteur interne est le premier système du détecteur ATLAS à observer les produits des collisions. Pour identifier les particules à leur passage, nous utilisons la technique de trajectométrie. Cette technique vise à dévier la trajectoire des particules produites dans la collision pour mesurer leur impulsion et leur charge électrique. Les trajectoires sont courbées par des électro-aimants de type solénoïde qui produisent un champ de 2 T. En mesurant la courbure des trajectoires, nous pouvons obtenir l'impulsion de la particule, tel que l'indique l'équation 12 [126]. Ici, z représente la charge de la particule, B (Tesla) la valeur du champ, R (m) le rayon de courbure, p l'impulsion (GeV/ c) et λ l'angle d'inclinaison.

$$p \cos(\lambda) = 0.3zBR \quad ; \quad (12)$$

Le détecteur interne est formé de trois sous-détecteurs en couches d'oignon cylindriques qui permettent de couvrir la région $|\eta| < 2.5$. Ceux-ci sont affichés à la Figure 19. Le premier sous-détecteur se nomme le détecteur à pixels. Celui-ci est constitué de quatre couches de pixels semi-conducteurs, ainsi que de trois disques situés aux extrémités du détecteur, nommés *end-caps*. Ces pixels d'une dimension de $50 \mu\text{m}$ par $400 \mu\text{m}$ sont formés de silicium oxygéné de type-n avec des implants n^+ [17]. Lorsque des particules chargées produites dans les collisions passent dans le silicium, elles y déplacent des charges. Un champ électrique impose le mouvement de ces charges vers une électrode où elles sont recueillies et converties en un signal électrique. Ce signal indique alors le passage d'une particule dans le pixel. Les différentes couches de pixels permettent d'obtenir des points d'interaction à mesure que la particule progresse dans le détecteur et des techniques d'analyse sont utilisées pour reconstruire la trajectoire des particules, ou la trace, qui déposent leur énergie dans plus d'un pixel [24]. De plus, la première des quatre couches cylindriques est particulière : on la nomme *insertable B-Layer*. Elle fut installée en 2014, puisque la première couche du détecteur interne avait subi beaucoup de dommage de radiation et n'allait plus fonctionner proprement pour les Périodes (*Run*) 2 et 3 [50]. Le détecteur à pixels possède une grande résolution et nous permet d'obtenir des mesures des traces précises.

Après le détecteur à pixels, se trouve le trajectomètre semi-conducteur (SCT). Ce trajectomètre est aussi composé de quatre couches, formées cette fois-ci de senseurs micro-bandes en silicium. Ces senseurs d'une dimension de $80 \mu\text{m} \times 12 \text{cm}$ indiquent une mesure en une dimension (1D) [153]. Pour avoir des mesures en 2D, nous devons superposer des couches de deux senseurs. Cette disposition limite la résolution du détecteur, mais permet d'obtenir quatre points de positions supplémentaires pour les particules chargées dans le détecteur interne.

Le trajectomètre à radiation de transition (TRT) est le dernier des trois sous-détecteurs. Celui-ci est constitué de tubes à dérives de 4 mm de diamètre, placés en position parallèle au faisceau, dans un baril, et en position radiale dans deux disques *end-caps*. Ces tubes sont remplis d'un mélange de gaz de xénon (ou d'argon), de dioxyde de carbone et de dioxygène. Une anode est située en leur centre, et la paroi du tube joue le rôle de cathode. Une différence de potentiel de 1530 V est appliquée entre les deux [17]. Au passage d'une particule chargée, le gaz est ionisé et les ions et électrons produits, sous l'effet du potentiel,

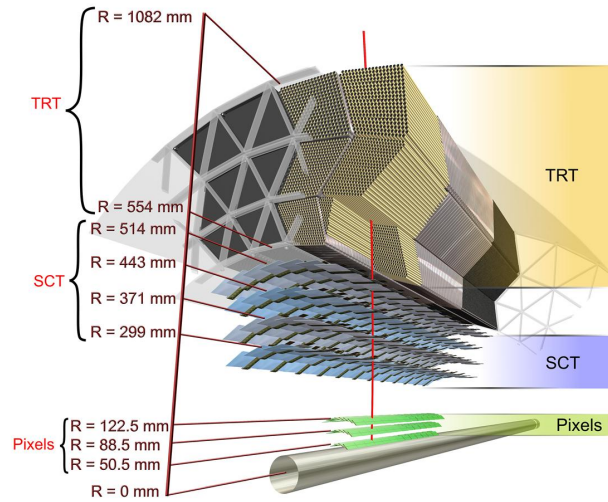


Figure 19. Schéma du détecteur interne. Celui-ci est composé du détecteur à pixels, du trajectomètre semi-conducteur (SCT) et du trajectomètre à radiation de transition (TRT) [128].

sont déplacés sur l’anode et convertis en un signal électrique. Des méthodes d’analyse sont utilisées pour reconstruire le point d’origine de ces cascades [37]. Aussi, lorsqu’une particule passe entre deux milieux (ex : l’espace entre deux tubes), il se produit un rayonnement de transition, soit la production d’un photon dans la gamme des rayons X. Ce rayonnement produit parfois des cascades de particules plus grandes que le signal dans le tube lui-même [37]. Nous exploitons ce rayonnement pour identifier les électrons e^- et les différencier des pions π (un type de mésons) grâce à leurs ratio énergie/masse. En effet, pour une même énergie, comme les électrons sont beaucoup plus légers, ils ont une plus grande probabilité de provoquer un rayonnement de transition [37].

Enfin, nous mentionnons que les muons vont laisser une trace dans le détecteur interne qui sera reconstruite. La combinaison du détecteur interne et du spectromètre à muons sera essentielle pour leur identification.

Le détecteur interne nous permet donc de bien mesurer les charges et les traces des particules chargées. Par contre, plusieurs autres particules passent à travers ce système sans être identifiées ou arrêtées. C’est le cas, entre autres, des particules neutres et des hadrons, qui nécessitent l’utilisation d’autres systèmes de détection.

2.2.3. Les calorimètres

Le système de calorimètres est réparti en couches autour du détecteur interne. Nous utilisons les calorimètres pour détecter les particules qui interagissent par la force forte et la force électromagnétique. Pour ce faire, nous exploitons la technique de calorimétrie, qui consiste à laisser les particules se propager dans le détecteur dans des couches successives de matériaux passif et actif. Une particule passant dans un matériau passif interagit avec celui-ci et crée des gerbes, ou cascades de particules. Ces produits rencontrent alors un matériau actif, qui absorbe leurs dépôts d'énergie en interagissant peu avec elles [74]. Le calorimètre mesure ces dépôts d'énergie, ce qui nous permet d'identifier les particules par la forme et la quantité des dépôts.

Deux types de cascades sont produites dans les calorimètres. Les cascades électromagnétiques sont constituées de leptons et de photons et sont produites par une production de paires et par du rayonnement Bremsstrahlung (voir section 1.1.3). Les cascades hadroniques sont plutôt formées de hadrons. Ces deux types de cascades nécessitent l'utilisation de systèmes de détection différents. Ces systèmes sont respectivement le calorimètre électromagnétique et le calorimètre hadronique.

La disposition des calorimètres est affichée à la Figure 20. D'abord, le calorimètre électromagnétique est composé d'un baril (*LAr electromagnetic barrel*) et deux *end-caps* (EMEC). Il couvre la région de $|\eta| < 3.2$ et utilise le plomb comme matériau actif et l'argon liquide comme matériau passif [17]. Ces matériaux sont disposés en couches accordéon pour bien couvrir l'ensemble des régions possible en ϕ . Lors de la création des cascades électromagnétiques, les particules seront converties en un signal électrique dans le matériau actif, ce qui nous permet d'inférer leur passage [17].

Si nous regardons la Figure 20, nous remarquons que le calorimètre hadronique est beaucoup plus volumineux que la calorimètre électromagnétique. Ceci s'explique par le fait que les cascades hadroniques sont beaucoup plus complexes et longues que les cascades électromagnétiques. En effet, les cascades hadroniques ont souvent une composante électromagnétique en plus d'une composante hadronique. Pour ces raisons, la résolution des calorimètres hadroniques est généralement moins bonne que celle des calorimètres électromagnétiques.

Le calorimètre hadronique est composé d'un baril (*Tile barrel*), de deux barils étendus (*Tile extended barrel*), de deux *end-caps* (HEC) et d'une partie avant (*FCal*). Il couvre la

région de $|\eta| < 4.9$. De l'acier fait office de matériau passif dans les barils. Des scintillateurs en tuile de plastique y sont utilisés comme matériau actif [17]. Ces scintillateurs mesurent l'énergie des cascades en recueillant la lumière produite dans le matériau au passage des particules. Nous recueillons cette lumière avec des tubes photo-multiplicateurs. Les disques HEC utilisent plutôt une combinaison de plaques de cuivre et d'argon liquide. Le calorimètre avant (FCAL) utilise quant à lui des tubes d'argon liquide et un mélange de cuivre-tungstène [17]. Le FCal est utilisé pour mesurer l'énergie des particules qui ont un angle de diffusion très grand, soit celles produites vers la partie avant du détecteur, ainsi que les jets hadroniques et électromagnétiques [17].

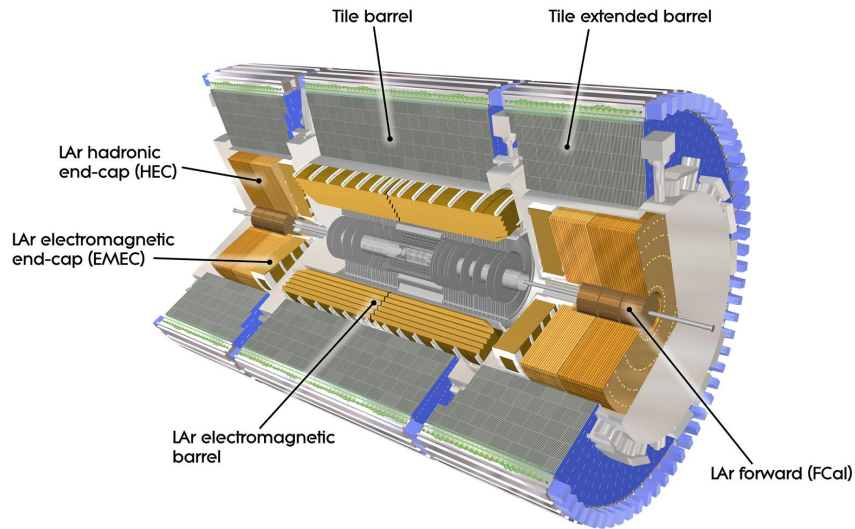


Figure 20. Le système de calorimètres du détecteur ATLAS. Le calorimètre électromagnétique est formé d'un baril (*LAr electromagnetic barrel*) et de deux disques (EMEC). Le calorimètre hadronique possède deux barils (*Tile barrel* et *Tile extended barrel*), ainsi que deux disques (HEC) et une partie avant (FCal) [127].

Par la combinaison du détecteur interne et des calorimètres, nous sommes capables d'identifier les électrons et les photons et d'arrêter les particules dans le détecteur. Cependant, d'autres particules sont capables de se faufiler à travers toutes ces couches sans être identifiées. Parmi celles-ci se trouvent les muons, qui nécessitent l'utilisation d'un spectromètre en plus de l'information laissée dans le détecteur interne pour leur identification.

2.2.4. Le spectromètre à muons

Le spectromètre à muons est situé autour des calorimètres, le plus à l'extérieur du détecteur ATLAS. Qu'est-ce qui rend les muons si difficiles à identifier ? La réponse se trouve dans le comportement du pouvoir d'arrêt dans le cuivre. Le pouvoir d'arrêt représente la perte moyenne d'énergie d'une particule par ionisation au passage dans un matériau. À des vitesses relativistes, ce pouvoir d'arrêt atteint une valeur minimale, nommée minimum d'ionisation [86]. Dans l'expérience ATLAS, les muons produits ont une énergie autour de cette valeur, qui est de l'ordre des GeV, ce qui leur permet de passer à travers le détecteur interne et les calorimètres sans être arrêtés [90]. Par contre, les muons vont tout de même laisser une trace dans le détecteur interne qui sera importante pour leur identification.

Le spectromètre à muons utilise la même technique de trajectométrie que le détecteur interne. Sa structure est affichée à la Figure 21. Un champ magnétique toroïdal (en forme de beigne) est créé par trois aimants, soit un baril (*Barrel toroid*) et 2 *end-caps* (*End-cap toroid*) pour courber la trajectoire des muons et mesurer leur impulsion indépendamment du détecteur interne. Deux sous-détecteurs sont utilisés pour détecter la présence des muons. D'abord, les tubes à dérive contrôlée (MDT) couvrent la région $|\eta| < 2.7$. Ceux-ci sont remplis d'un mélange de gaz (argon et dioxyde de carbone) qui sera ionisé par le passage des muons. Les ions et les électrons produits circulent sur une anode située au centre des tubes et faite d'un câble métallique en tungstène-rhénium plaqué or. Puis, ces ions et électrons sont convertis en un signal électrique [17]. Nous utilisons de trois à huit couches de ces tubes à dérive pour obtenir plusieurs points d'interaction et déterminer l'impulsion des muons. Des chambres à cathodes (CSC) sont également utilisées pour la détection des muons dans la région avant. Ces chambres nous permettent d'effectuer des mesures de précision supplémentaires aux extrémités du détecteur.

Finalement, le spectromètre à muons inclut deux systèmes qui contribuent au déclenchement (*trigger*) de l'enregistrement d'événements. Ceux-ci sont les chambres à plaques résistantes (RPC) et les chambres à intervalles minces (TGC), qui couvrent respectivement les régions $|\eta| < 1.05$ et $1.05 < |\eta| < 2.4$ [17].

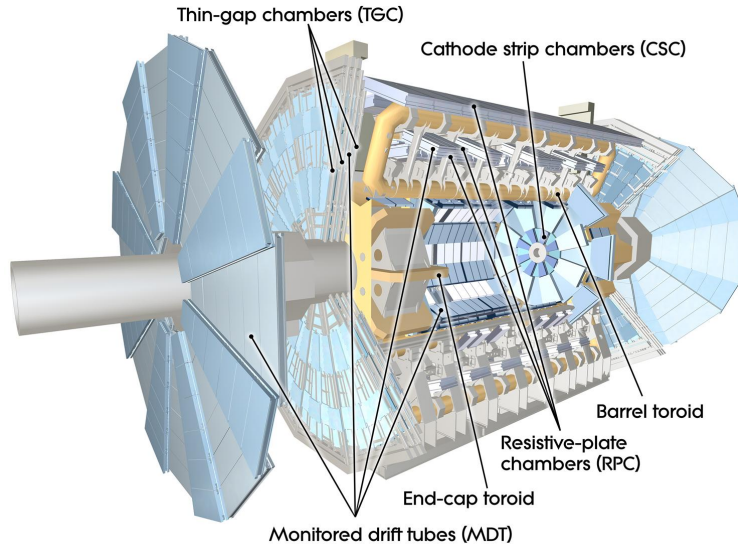


Figure 21. Le spectromètre à muons du détecteur ATLAS. Le spectromètre regroupe trois aimants (*Barrel* et *End-cap toroïd*) et quatre systèmes de détection et de déclenchement (MDT, TGC, CSC et RPC) [129].

2.2.5. Les détecteurs Timepix

Les détecteurs Timepix sont des types de détecteurs semi-conducteurs hybrides. Ils sont formés de capteurs de silicium de 256 x 256 pixels et liés à un système de lecture électronique [121]. Les technologies Timepix succèdent aux précédents Medipix utilisés dans l'expérience ATLAS [117]. Ces senseurs sont placés à divers endroits dans la caverne du détecteur ATLAS. Ils nous permettent de mesurer la luminosité de l'expérience, ainsi que d'étudier la radiation, incluant celles provenant des neutrons. Ces mesures sont importantes, car une bonne mesure de la luminosité nous permet d'améliorer la précision de diverses analyses, et parce qu'une mesure de la radiation nous permet de l'évaluer et de potentiellement d'en limiter les dommages sur l'électronique et les matériaux du détecteur.

Chaque pixel d'un Timepix, lorsqu'il détecte un signal, compare ce dernier à un discriminant et un seuil associé. Quand une particule dépose son énergie dans un pixel, celui-ci sera activé uniquement si l'énergie dépasse ce seuil prédéterminé. Les pixels activés forment alors des amas et, selon leur disposition et leur étalement, nous pouvons identifier le type de particules les ayant activés.

Les détecteurs Timepix peuvent opérer via trois modes différents. D'abord, le mode de comptage, ou *Hit counting* permet de compter le nombre d'événements durant un intervalle de temps donné. Le mode ToA, ou *Time of Arrival*, compte le temps écoulé entre

le déclenchement et la détection d'une particule de radiation. Enfin, le mode ToT, ou *Time over Threshold*, compte le temps durant lequel l'énergie d'un pixel demeure au-dessus de son seuil après un signal.

2.2.6. Le système d'acquisition de données

Les événements observés dans le détecteur ATLAS produisent une énorme quantité de données. Avec une collision à chaque 25 ns, ceci équivaut à un taux d'événements de 40 MHz et un taux de production de données de 60 TB/s. Il nous est impossible de conserver tous ces événements. Nous devons alors utiliser des systèmes pour réduire la quantité de données et choisir judicieusement celles qui seront enregistrées. Nous appelons ceux-ci le système de déclenchement, ou *trigger*, et le système d'acquisition de données. Nous décrirons brièvement le fonctionnement et l'utilité de chacun de ces systèmes.

Les événements détectés dans ATLAS passent d'abord dans le premier niveau de déclenchement (*Level 1 Trigger*). Ce système de déclenchement constitué d'électroniques sur mesure utilise l'information des calorimètres et du spectromètre à muons. C'est à ce niveau que nous déterminons les régions d'intérêt dans le détecteur et que nous cherchons pour des signatures de muons à haute impulsion, d'électron et de photons, de jets et de leptons τ . Ce niveau nous permet de réduire à 100 kHz (160 GB/s) le taux d'événements (de données) [17, 38]. Ces décisions sont prises extrêmement rapidement, soit moins de 2.5 μ s après la production des événements dans le détecteur [36].

Ensuite, le déclenchement de haut niveau (*High Level Trigger* ou HLT) combine la performance d'une ferme d'ordinateurs et de logiciels pour réduire à 1.5 kHz (1.5 GB/s) le taux d'événements (de données). En plus de l'information des calorimètres et du spectromètre, ce niveau utilise le détecteur interne pour améliorer les sélections du niveau précédent [17, 38]. Au total, le HLT sélectionne environ 1000 événements par seconde à analyser [36].

Finalement, les événements sont envoyés au centre de calcul du CERN pour être reconstruits et utilisés plus tard pour diverses analyses. Tout dépendant du processus étudié, divers critères au niveau du déclenchement sont demandés.

2.3. La reconstruction d'événements

Nous avons vu que le détecteur ATLAS utilise plusieurs systèmes pour détecter la présence de particules. Toutefois, pour analyser les processus enregistrés, nous devons également identifier ces processus et reconstruire les événements produits par les collisions proton-proton. Cette reconstruction est réalisée par divers algorithmes en fonction du type de particules. Nous verrons dans la section suivante comment certaines des particules sont reconstruites dans l'expérience ATLAS.

2.3.1. Les électrons et les photons

Pour reconstruire les électrons et les photons, nous utilisons l'information récoltée par le calorimètre électromagnétique et le détecteur interne. D'abord, nous définissons des fenêtres, ou cellules, dans le calorimètre électromagnétique et nous étudions la forme des cascades dans ces cellules. Pour les électrons, nous associons ensuite leur dépôt d'énergie dans ces fenêtres à une trajectoire (trace) dans le détecteur interne et à son origine, que l'on nomme « vertex primaire ». Toutefois, de multiples événements peuvent reproduire la signature expérimentale des électrons et ceux-ci peuvent provenir de la collision principale ou de collisions supplémentaires. L'identification des électrons est donc réalisée avec une certaine probabilité que l'électron corresponde à un signal ou à un bruit de fond. Pour définir cette probabilité, nous utilisons des variables qui calculent l'accord entre les dépôts d'énergie et les traces et qui évaluent la qualité de ces traces. Tout dépendant des critères, on dira que ces sélections sont lousse (*loose*), médium (*medium*) ou serrée (*tight*), avec des critères de plus en plus restreints [20]. Pour les photons, nous avons seulement des sélections lousse ou serrée et, contrairement aux électrons, les photons ne doivent pas posséder de trace devant leurs dépôts d'énergie dans le calorimètre.

De plus, pour nous permettre une meilleure identification d'électrons et de photons, nous demandons à ce que ceux-ci soient isolés d'autres particules. Ceci nous permet de réduire la contribution de bruit de fond. Pour ce faire, nous définissons des variables d'isolation [20]. Ces variables nous permettent de discriminer entre des électrons qui proviennent, par exemple, de désintégration à ceux issus de la conversion d'un photon ou d'un jet, et de différencier les photons qui proviennent, par exemple, de diffusion Compton à ceux issus de la désintégration de pions.

Enfin, les événements bien connus de type $Z \rightarrow e^+e^-$ et $J/\Psi \rightarrow e^+e^-$ (J/Ψ est

un méson) sont utilisés pour la calibration d'échelle d'énergie et pour assurer une bonne performance de reconstruction [162].

2.3.2. Les jets hadroniques

Pour reconstruire les jets hadroniques, nous récoltons l'information des cellules des calorimètres. À partir des cellules, nous définissons des amas topologiques en trois dimensions (*topological clusters* ou *topoclusters*). Chacun de ces amas devrait correspondre, en théorie, à une des particules qui constitue le jet. Pour construire un de ces amas, nous demandons d'abord à ce que celui-ci ait une énergie plus grande que le bruit de fond, qui est défini comme la somme quadratique du bruit lié à l'électronique et aux collisions qui ne proviennent pas de la collision proton-proton d'origine (*pile-up*). Ensuite, les cellules voisines de l'amas qui passent un certain seuil d'énergie sont ajoutées à l'amas par itérations. Finalement, un anneau de cellules voisines est ajouté aux amas pour former les amas topologiques finaux [162]. Ces trois étapes de formation sont affichées à la Figure 22 dans le cas du calorimètre hadronique avant (FCal).

Après la formation des amas, nous utilisons divers algorithmes pour recombinaison et former les jets, notamment les algorithmes de cône et d'agglomération (*clustering*). Le premier type identifie les jets en maximisant la quantité d'énergie qui peut être couverte par un cône d'un rayon R . Toutefois, l'utilisation de cet algorithme possède des lacunes lors de la recombinaison de certains types d'évènements, par exemple lorsque qu'une particule se désintègre en plus d'un jet [140]. C'est pourquoi le deuxième type d'algorithme est le plus utilisé, et nous discuterons de ce dernier plus en détails ci-bas.

Il existe une grande variété d'algorithmes d'agglomération, mais les plus populaires pour la reconstruction des jets hadroniques dans ATLAS sont les algorithmes de k_T , de Cambridge et Aachen et d'anti- k_T [140]. Dans l'algorithme k_T , nous commençons par définir une variable de distance d_{ij} entre deux particules i et j :

$$d_{ij} = \min(p_{Ti}^2, p_{Tj}^2) \Delta R_{ij}^2 \quad , \quad (13a)$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \quad . \quad (13b)$$

Dans ces équations, ΔR_{ij} représente la séparation entre les deux particules, p_T est leur impulsion transverse et y et ϕ leur rapidité et position azimutale. Si la valeur minimale

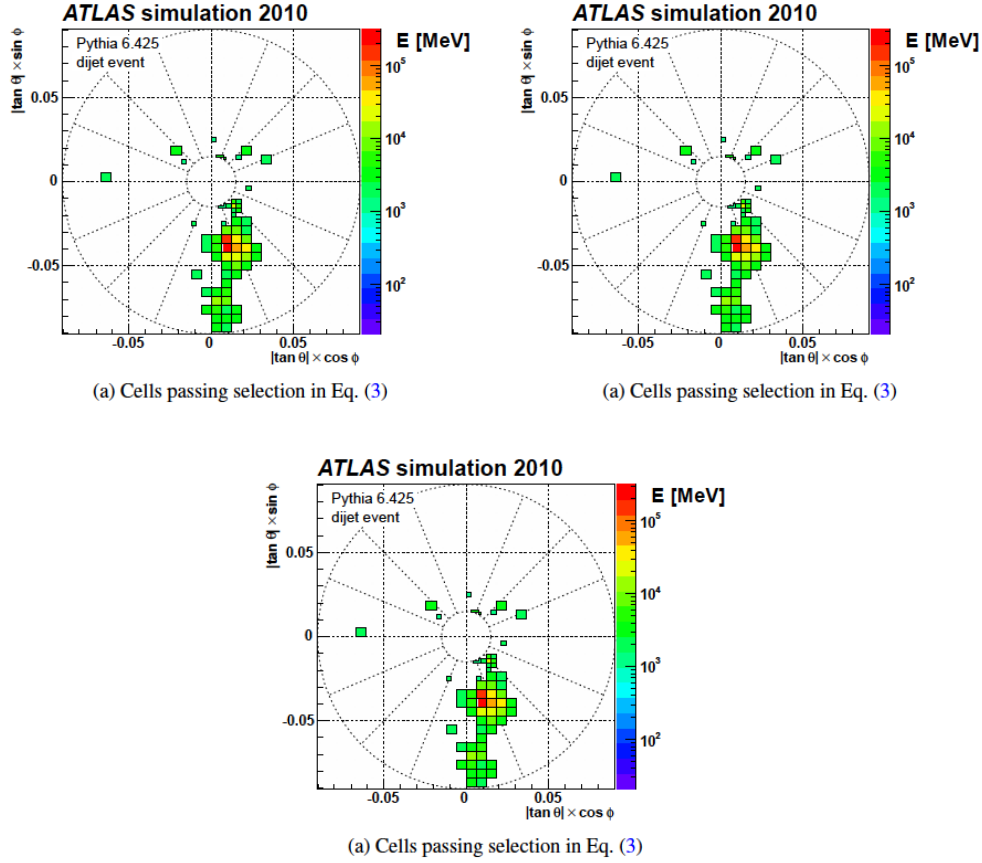


Figure 22. Formation des amas topologiques dans le calorimètre hadronique avant (FCAL). Les trois images représentent les étapes de formation des amas [25].

de d_{ij} est sous un certain seuil d_{cut} , alors les particules i et j sont recombinaées pour former un « pseudo-jet ». La procédure est répétée, jusqu'à ce que les particules restantes soient au-dessus du seuil d_{cut} et qu'elles soient déclarées comme des jets elles-mêmes [140]. Nous introduisons ensuite une deuxième distance d_{iB} entre la particule i et le faisceau, définie à l'équation 14, et qui correspond au carré de l'impulsion transverse de la particule i . Si cette distance s'avère plus petite que d_{ij} , nous recombinaons la particule i avec le faisceau.

$$d_{iB} = p_{Ti}^2 \quad ; \quad (14)$$

Une version inclusive de cet algorithme utilise les mêmes équations, à la différence près où la quantité ΔR_{ij}^2 est divisée par une quantité R^2 , qui représente un cône de rayon R . Dans cette version de l'algorithme, aucune valeur de d_{cut} n'est utilisée. On compare plutôt d_{ij} avec d_{iB} et si la première est plus petite que la deuxième, les particules i et j sont recombinaées. Sinon, la particule i est déclarée comme un jet. Ainsi, toutes les particules

sont incluses dans un jet final. Dans ce cas particulier, nous permettons aux particules « douces » (*soft*), soit les particules qui proviennent de résidus de désintégration et qui ont une quantité de mouvement plus basse, d'être des jets. Dans l'analyse des événements, il faut alors déterminer un p_T minimal pour que le jet soit d'intérêt [140]. Aussi, la valeur de $R = 0.4$ est normalement utilisée pour reconstruire les jets.

Il existe également un algorithme « anti- k_T ». Dans cette algorithme, nous introduisons une quantité p dans les équations :

$$d_{ij} = \min(p_{T_i}^{2p}, p_{T_j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad , \quad (15a)$$

$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \quad , \quad (15b)$$

$$d_{iB} = p_{T_i}^{2p} \quad . \quad (15c)$$

Pour l'algorithme anti- k_T , p prend la valeur de -1 . Dans ce cas précis, les jets reconstruits auront tendance à favoriser les particules « dures » (*hard*) qui ont une plus grande énergie, ce qui résulte en des jets plus réguliers et d'allures circulaires. Lorsque $p = 1$, nous retrouvons l'algorithme k_T inclusif. Le cas où $p = 0$ se nomme l'algorithme de Cambridge et Aachen. Celui-ci donnera des jets de formes plus irrégulières, mais sera plus utile pour étudier les jets possédant une sous-structure complexe à deux et trois sous-jets [68, 140]. Un exemple de reconstruction par ces algorithmes est affiché à la Figure 23.

Cependant, dans les cas des jets boostés, les produits des désintégrations des particules massives ne peuvent pas tous être inclus par l'algorithme d'anti- k_T en prenant l'usuel $R = 0.4$. Les produits sont donc reconstruits dans un seul grand jet, en prenant $R = 1.0$. Dans le cas où $R = 1.0$, nous utilisons une méthode de recombinaison supplémentaire qui se nomme l'élagage, ou *trimming*. Cette méthode consiste à recombinaison les jets déjà formés en utilisant un rayon plus petit et à conserver seulement les sous-jets au-dessus d'un certain seuil en énergie [140]. Cette méthode nous permet de réduire la contamination des événements par diverses sources [107].

Pour terminer, l'énergie des jets est étalonnée dans les vraies données en utilisant des processus tels que $Z + \text{jet}$, $\gamma + \text{jet}$ et dijet. D'abord, une correction liée au *pile-up* des simulations Monte Carlo est appliquée afin d'enlever l'énergie en trop ajoutée par les interactions proton-proton qui ne sont pas d'intérêt et pour éliminer la contribution de faux jets. Puis, la réponse en énergie à cette correction est simulée et une calibration résiduelle

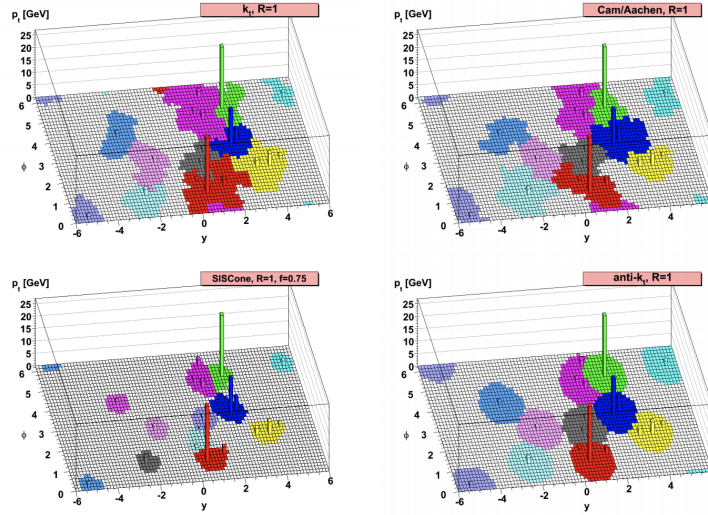


Figure 23. Le comportement de divers algorithmes de reconstruction de jets. Les algorithmes de k_T , de Cambridge et Aachen, de cône et d’anti- k_T reconstruisent différemment une collision de partons générée par un logiciel [48].

est effectuée aux jets dans les données pour tenir compte de la différence entre ceux-ci et ceux générés par les simulations Monte Carlo [21, 162].

2.3.3. Les muons

La reconstruction des muons se fait initialement d’une façon séparée dans le détecteur interne et le spectromètre. Ces deux systèmes nous permettent d’obtenir des mesures indépendantes de l’impulsion des muons. Dans le spectromètre, l’information de détection des muons est utilisée pour associer des segments de collision dans les diverses couches du spectromètre et pour reconstruire leurs traces [27, 62].

À partir de ces mesures indépendantes, nous utilisons quatre méthodes différentes de combinaison de traces et/ou de dépôts d’énergie pour reconstruire les muons. La première méthode consiste à combiner les traces du détecteur interne à celles du spectromètre. Nous appelons les muons reconstruits par cette méthode les muons combinés (CB). Une autre méthode consiste à identifier une trace dans le détecteur interne comme étant un muon si ce dernier est associé à au moins une trace dans le spectromètre. Ces muons seront dit identifiés par segment (*segment-tagged* ou ST). Nous pouvons aussi identifier une trace dans le détecteur interne comme un muon si elle est associée à un dépôt d’énergie dans le calorimètre : ce sont les muons identifiés par le calorimètre (*calorimeter-tagged* ou CT). Finalement, certains muons sont reconstruits seulement à partir des traces dans le

spectromètre, que nous appelons les muons extrapolés (ME) [27].

Comme pour les électrons, nous pouvons définir des sélections lousse, médium et serrée et définir des variables d'isolation pour différencier des muons provenant de l'interaction originale ou de bruit de fond. Nous utilisons également les événements de type $J/\Psi \rightarrow \mu\mu$ et $Z \rightarrow \mu\mu$ pour la calibration [27].

2.3.4. L'énergie manquante

Les algorithmes précédents sont efficaces pour reconstruire la majorité des particules qui sont produites à la suite des collisions de proton-proton. Cependant, certaines particules telles que les neutrinos arrivent à passer inaperçues, sans laisser de trace ni de dépôt d'énergie dans le détecteur. Lorsqu'une collision produit ces particules « invisibles » dans le détecteur, la somme de l'énergie finale dans le plan transverse (\hat{x}, \hat{y}) , qui est définie à partir de la quantité de mouvement dans le plan transverse, ne sera pas équilibrée avec l'énergie transverse initiale, qui, elle, est nulle. Nous appelons la quantité d'énergie nécessaire pour équilibrer l'énergie finale et initiale de la collision l'énergie manquante dans le plan transverse (E_T^{miss}). Nous pouvons alors inférer à partir de cette quantité la présence de neutrinos, ou peut-être même d'autres particules au-delà du modèle standard dans la collision [27].

Nous mesurons l'énergie manquante dans une collision à partir de deux quantités. Nous utilisons les particules et les jets reconstruit(e)s et calibré(e)s, ainsi que les traces des particules chargées associées à un vertex dans le détecteur interne [27, 162]. La somme de la quantité de mouvement de ces deux quantités dans les plan x et y nous permet de définir la quantité d'énergie manquante. Finalement, nous utilisons les événements bien connus et calibrés dans l'expérience ATLAS, tels que $Z \rightarrow \ell\ell$ (où ℓ est un lepton) et $W \rightarrow \ell\nu$ pour calibrer la performance de reconstruction de l'énergie manquante [162].

2.4. Enjeux démographiques et environnementaux au CERN

Comme toute organisation ou centre de recherche en physique, le CERN s'inscrit dans un contexte socioculturel. Concernant la démographie, le CERN regroupe des milliers de

scientifiques provenant d'horizons différents, soit de près de 90 nationalités différentes¹ [31]. Autant l'esprit de collaboration internationale permet de trouver des solutions créatives et innovatrices, il se crée parfois des conflits de valeurs entre les différentes cultures et personnes. En particulier, 80% des membres du CERN sont des hommes; une situation très similaire pour le domaine de la physique en général et dans de nombreuses disciplines scientifiques et technologiques [31, 152]. Selon la physicienne Pauline Gagnon, si la tendance actuelle se maintient, l'égalité homme-femme au CERN ne serait atteinte qu'en 2798 [80].

De plus en plus de membres de la communauté en physique adhèrent à l'idée que la diversité en science est essentielle et bénéfique pour la recherche. D'une part, il est crucial d'aborder les questions de diversité en sciences pour des questions de justice sociale et d'égalité des chances et des opportunités. Les sciences, y compris la physique, devraient être accessibles à tous et toutes, et les caractéristiques individuelles telles que le genre ou la nationalité ne devraient pas représenter des freins à l'accessibilité de ces domaines de recherche. De nombreuses études ont montré que des groupes hétérogènes proposent des solutions plus créatives et complètes aux questions et aux problèmes de recherche que les groupes homogènes [69, 125, 132].

Il est aussi important de noter que l'écart de représentativité de genre en sciences n'est pas causé par un écart de compétences. Les femmes sont aussi compétentes en sciences que les hommes [66, 97, 108, 155]. En fait, des facteurs comme les stéréotypes d'un milieu et la sous-représentation des femmes en sciences affectent leur confiance en leurs capacités personnelles [119, 150].

Pourquoi alors n'y a-t-il pas autant de femmes que d'hommes si leurs compétences sont égales? La question historique est impossible à ignorer. Auparavant, il était mal vu pour les femmes d'étudier en sciences. D'ailleurs, les femmes ont fréquemment vu leurs contributions effacées [8, 138]. Elles se sont aussi vu refuser l'obtention du prix Nobel en physique [80]. Même si, de nos jours, il est beaucoup plus normalisé pour les femmes d'oeuvrer dans les domaines scientifiques, les sciences sont toujours perçues comme un domaine masculin. Ceci perpétue les stéréotypes et provoque un syndrome de l'imposteur² chez les personnes qui ne correspondent pas à la description du scientifique typique [12, 97]. Les femmes et autres groupes historiquement marginalisés ont donc moins tendance à se diriger vers les sciences, puisqu'elles et ils ne considèrent pas correspondre aux caractéristiques du milieu.

¹À noter aussi que même si des personnes de 90 nationalités différentes y contribuent, les recherches du CERN sont largement dominées par les pays euro-américains [31].

²Le syndrome de l'imposteur se manifeste chez une personne comme un doute envers ses capacités et sa légitimité, malgré ses nombreuses réussites [57].

Comme la physique est un domaine majoritairement blanc et masculin, un climat organisationnel associé à une perspective homogène s’est installé et se perpétue toujours dans ses institutions. Le climat organisationnel peut être défini comme l’ensemble des pratiques, des valeurs et des politiques d’une organisation [114]. Un climat organisationnel est dit non favorable lorsque le comportement des personnes et les politiques en place nuisent aux personnes issues de groupes historiquement marginalisés. En particulier, au CERN, la communauté LGBTQ+ a vu certaines de ses affiches être vandalisées, en plus d’être elle-même couverte d’insultes [80, 115]. Le CERN n’est pas non plus à l’abri d’incidents sexistes, comme l’a démontré un incident impliquant un représentant du CERN lors d’une conférence [3, 4]. Alors que les statistiques démographiques offrent des indices quant au climat organisationnel, notamment à travers la rétention, la proportion de femmes diminue selon leur expérience et ce même en tenant compte des changements démographiques des 100 dernières années des domaines scientifiques. En effet, à la Figure 24, nous voyons que la fraction de femmes membres de l’expérience ATLAS diminue avec l’âge. Cette diminution se voit également dans la fraction de femmes chez les auteurs et autrices, illustrant un problème de rétention à long terme. De plus, nous voyons que la fraction des baccalauréats octroyés aux femmes stagne depuis le début des années 2000. Cette Figure confirme que le problème de rétention du CERN n’est pas seulement issu de la démographie initiale des groupes plus âgés, mais bien d’un problème de recrutement et de rétention des femmes en physique. Diverses études féministes, dont la référence [101], ont pointé vers certaines causes associées au climat de travail, incluant la faible reconnaissance des réussites, mais aussi à plusieurs enjeux associés aux microagressions³ ou encore à la conciliation travail-famille.

Heureusement, le CERN accorde de plus en plus d’importance aux enjeux de diversité et d’inclusion et fait des efforts conscients pour augmenter la diversité dans sa communauté [84]. Un comité LGBTQ+ existe depuis 2010 (<https://lgbtqcern.com/wordpress/>) au CERN, et un comité diversité a pris forme dans l’expérience ATLAS [35]. Ces comités, par leurs actions, visent à sensibiliser la communauté du CERN aux enjeux de diversité et d’inclusion. Malgré tout, le problème de climat demeure et ne pourra se résoudre qu’en opérant à un changement de culture. Par exemple, offrir des programmes de formation, combattre les microagressions et les stéréotypes par la sensibilisation et améliorer les conditions de travail pour les parents permettrait d’augmenter la rétention des femmes et des autres personnes issues de la diversité [61]. La professeure en histoires des sciences Londa Schiebinger recommande pour améliorer le sort de ces personnes dans les domaines

³« Une microagression est un terme utilisé pour décrire des indignités verbales, comportementales ou environnementales brèves et courantes, intentionnelles ou non, qui communiquent des affronts et des insultes hostiles, méprisantes et négatives envers tout groupe marginalisé » [159].

des STIM, d'à la fois « régler le recrutement », « régler les institutions et la culture organisationnelle » et « régler la connaissance en intégrant les concepts de genre à la recherche » [82].

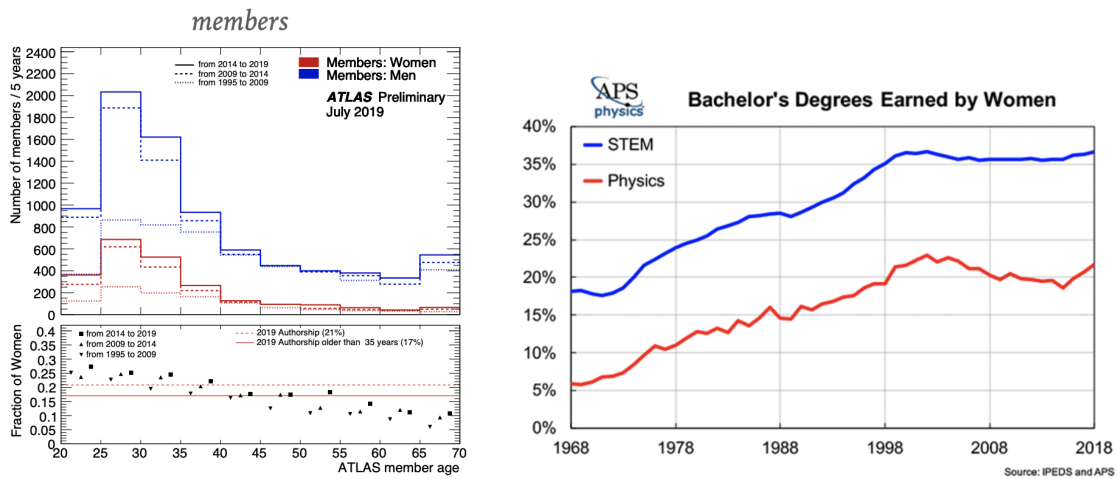


Figure 24. Figures discutant de l'évolution de la fraction des femmes en sciences, technologies, ingénierie et mathématiques (STIM) et au CERN. À gauche, la fraction des membres de l'expérience ATLAS qui sont des femmes en date de 2019 [31]. À droite, la fraction de baccalauréat obtenus par les femmes aux États-Unis dans les domaines des STEM [13].

En plus des défis liés à la diversité et à l'inclusion, le CERN rencontre plusieurs défis environnementaux. Le détecteur ATLAS utilise une combinaison de diverses ressources et matériaux comme le silicium, l'or et le tungstène pour ses détecteurs et son système d'électronique. Il ne faut pas oublier que ces matériaux sont extraits de mines qui ont un impact sur l'environnement et qui exploitent parfois les travailleurs et travailleuses dans des conditions de travail dangereuses [76, 98, 118, 135]. Aussi, à lui seul, le complexe utilise 2% de l'énergie de la Suisse [53]. Les systèmes de refroidissement et les détecteurs sont aussi responsables de la majorité des émissions de gaz à effet de serre du CERN, dont plus de 78% sont de gaz fluorés, ces gaz qui demeurent dans l'atmosphère très longtemps et qui contribuent à l'effet de serre.

Enfin, les « nuages » et serveurs de calcul, qui sont utilisés par les scientifiques du CERN, ne sont pas aussi aériens que l'expression ne l'entend. Ces installations consomment beaucoup de ressources minières et énergétiques. Le CERN et l'Université de Montréal ont tous deux accès à des centres de serveurs comme le centre de données du CERN et Calcul Québec. Ces centres d'ordinateurs et de serveurs ont un impact tangible sur le réchauffement climatique, comme le présente le rapport *Clicking Green* [87], et les chercheurs et chercheuses reçoivent rarement de l'information dans le cadre de leurs études pour réduire

leur empreinte écologique numérique.

Le CERN se responsabilise de plus en plus aux enjeux environnementaux et promet des efforts de réduction d'ici 2024. Cependant, en plus de ces promesses de réduction, le CERN pourrait réduire son impact sur l'environnement en réfléchissant à des alternatives pour la suite de ses opérations. En tant que membre de l'expérience ATLAS, je crois que d'autres solutions sont envisageables dans une perspective écologique. Par exemple, plutôt que d'augmenter sans arrêt la luminosité du LHC et d'effectuer rénovations après rénovations, consommant toujours plus de ressources, nous pourrions effectuer la recherche de façon plus écologique en utilisant de nouveaux outils pour analyser les données déjà disponibles, comme l'apprentissage machine. Les voyages pour conférence pourraient également être remplacés par des visio-conférences afin de réduire l'empreinte carbone des scientifiques. Des articles de sensibilisation commencent à être publiés dans des journaux scientifiques prestigieux [14, 70, 137] afin de sensibiliser la communauté scientifique aux effets environnementaux des voyages en avion et suggèrent à la communauté de réfléchir à l'adoption plus large des visio-conférences.

2.4.1. Conclusion

Enfin, nous avons vu dans le présent chapitre comment le détecteur ATLAS situé au CERN nous permet d'observer des collisions à très haute énergie grâce à la technologie de pointe. Nous comprenons également comment les jets et les diverses particules sont reconstruit(e)s dans le détecteur. Toutefois, nous ne savons toujours pas si l'ensemble des jets au LHC sont produits selon le modèle standard. La nouvelle physique pourrait-elle se manifester sous cette forme? Aussi, la quantité de données enregistrées à chaque jour au CERN est astronomique. Il nous est presque impossible de trouver la nouvelle physique dans cette mer de données manuellement. C'est pourquoi l'apprentissage machine pourrait être un outil essentiel dans cette recherche de nouvelle physique. Dans le prochain chapitre, nous nous intéresserons aux motivations et au contexte derrière le projet de ce mémoire et nous présenterons la méthodologie utilisée.

Chapitre 3

Motivations physiques et méthodologie

Le CERN, dans lequel s'inscrit l'expérience ATLAS, a prouvé être un moteur d'innovation depuis sa fondation en 1949. En plus d'effectuer de la recherche de pointe en physique des particules, le CERN a créé des protocoles pour gérer et diffuser d'énormes quantités de données. Nous nous trouvons dans une période de recherche et de développement, et cette phase d'innovation technique et numérique nous permet d'explorer d'autres outils, tel que l'intelligence artificielle, pour nous aider dans notre quête de nouvelle physique. Nous savons également que les jets boostés pourraient contenir cette nouvelle physique [1]. Bien que des algorithmes de reconstruction performants pour ces jets existent, ces événements demeurent complexes et partiellement incompris des physiciens et physiciennes. Heureusement, l'apprentissage machine est un outil performant dans les terrains de jeux où beaucoup de données sont disponibles [164, 165]. Nous arrivons au sujet de ce mémoire : la recherche de nouvelle physique dans les jets boostés au détecteur ATLAS par l'utilisation de l'apprentissage machine et plus précisément de la détection d'anomalies. Pour commencer le présent chapitre, nous aborderons un peu plus en détails les jets boostés, les résonances et notre choix de signal pour tester la performance de notre approche sans modèle, ou *model-independent*. Nous étudierons une méthode simplifiée de détection d'anomalies par la distance de Wasserstein. Puis, nous explorerons les bases de l'apprentissage machine nécessaires à la compréhension de ce mémoire. Enfin, nous étudierons la performance d'un modèle supervisé qui a réalisé une étude similaire à la nôtre à des fins de comparaison et nous terminerons par la discussion de certains défis liés à l'intelligence artificielle (IA).

3.1. Choix de signal et de bruit de fond

La première étape d'une recherche de nouvelle physique générale au LHC consiste à déterminer l'espace de recherche. Comme nous n'avons pas de modèle précis à observer et que la

quantité de données disponibles au LHC est immense, il est crucial de se restreindre à un espace précis pour limiter les recherches et optimiser le temps d’entraînement de nos algorithmes. Dans la présente section, nous expliquerons plus en détails le choix des jets boostés comme espace de recherche, et nous présenterons également le choix de signal test pour notre analyse.

3.1.1. Les jets boostés et les résonances

Les jets boostés, tels que présentés dans la section 1.1.5, sont souvent issus de la désintégration de particules massives avec un temps de vie très court. L’impulsion de ces particules doit aussi être beaucoup plus grande que leur masse pour que leurs produits de désintégration se retrouvent dans un seul et même jet. Les particules les plus communes à se désintégrer en ce type de jets sont le boson W , le boson Z , le boson de Higgs et le quark top. La plupart du temps, nous utilisons l’algorithme d’anti- k_T pour reconstruire ces jets dans le détecteur. Dans les cas boostés, pour que les produits des désintégrations des particules massives soient reconstruits dans un seul jet, nous utilisons souvent l’algorithme d’anti- k_T en prenant une valeur de $R = 1.0$ [28].

Il arrive que les jets boostés soient issus d’une résonance. Le concept de résonance fait référence à un pic dans la section efficace d’un processus à une certaine énergie. Ces pics représentent des particules intermédiaires s’étant désintégrées, leur temps de vie étant généralement très court. Le temps de vie de ces résonances est directement lié à la largeur du pic Γ par l’équation $\Gamma = \hbar/\tau$, avec \hbar la constante de Planck réduite et τ le temps de vie. Un exemple de résonance est celle du boson Z , observée la première fois au LEP par l’expérience ALEPH (*Apparatus for LEP PHysics at CERN*) par l’étude de l’interaction $e^+e^- \rightarrow q\bar{q}$. Le pic résultant, à la Figure 25, affiche l’observation de la résonance à la masse du boson Z à 91.1876 GeV [126].

Les résonances sont importantes dans l’optique d’une recherche de nouvelle physique dans les jets boostés. En effet, ces derniers abritent des possibilités de résonances inconnues, c’est-à-dire de nouvelles particules de masses inconnues s’étant désintégrées en des particules du modèle standard. Des résonances telles que le boson W' [89] et le boson Z' [141], ainsi que les quarks vectoriels (*Vector-Like-Quark*) [6] sont des exemples de nouvelles résonances qui pourraient se manifester dans les jets boostés. Bien que la présente étude ne se concentre pas sur un de ces modèles en particulier, ces possibilités justifient notre choix des jets boostés pour notre recherche.

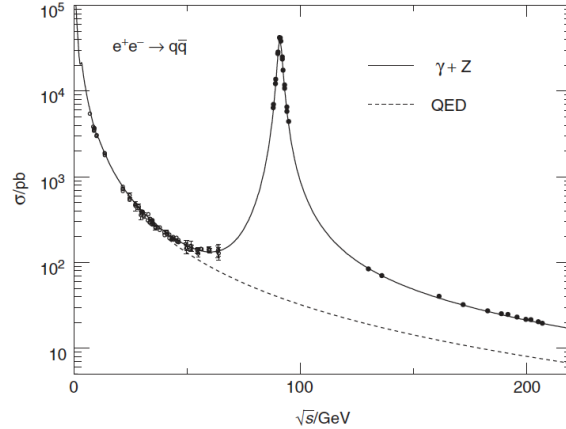


Figure 25. L’observation de la résonance du boson Z au collisionneur LEP par la collaboration ALEPH. La section efficace du processus $e^+e^- \rightarrow q\bar{q}$ est affichée pour les canaux avec un photon et un boson Z virtuel (ligne pleine), ainsi que pour le processus de la QED (ligne pointillée) [7].

Avant d’effectuer une recherche générale dans tous les jets boostés du détecteur ATLAS, nous devons toutefois développer une technique qui est capable de reconnaître les événements anormaux dans un jeu de données. Pour ce faire, nous déterminons d’abord un terrain d’entraînement pour tester notre capacité à découvrir de la nouvelle physique. Nous choisissons un signal déjà connu à retrouver parmi un bruit de fond dominant afin d’évaluer notre performance dans des données de simulation. Puis, nous développerons une technique afin de retrouver ce signal, qui sera expliquée un peu plus loin.

3.1.2. Signal et bruit de fond

Notre choix de signal parmi les jets boostés s’est arrêté sur le quark top. Le quark top a une grande masse ($m = 172.76$ GeV [126]) et se désintègre presque toujours en un quark bottom et un boson W . Au LHC, la création d’une paire de quark top et d’anti-top se produit tel qu’à la Figure 26. Tout dépendant de la désintégration du W , les jets auront diverses signatures. Dans tous les cas, le quark bottom produira une signature de jet dans le détecteur. La signature « lepton plus jet » implique que l’un des W se désintègre en un lepton et son neutrino associé (leptoniquement), et que l’autre produise une paire quark-antiquark (hadroniquement). Il arrive aussi que les deux bosons W se désintègrent leptoniquement et nous appelons cette signature « di-leptons ». Finalement, les deux bosons W peuvent se désintégrer hadroniquement en des quarks et anti-quarks. Les six quarks produits sont alors reconstruits en jets dans le détecteur.

Dans notre projet, nous utilisons l'information d'un seul des deux jets de top (ou d'anti-top). Ceci nous permet de simuler l'analyse que nous ferons dans le cadre d'une recherche générale de nouvelle physique dans les vraies données du détecteur où nous observerons seulement un seul jet, sans hypothèse initiale sur la nature de l'événement. Cette condition ne nous permettrait pas, par exemple, de distinguer entre les différents types de signatures pour un quark top, ou bien de tout autre type de signal.

Par contre, dans le cadre précis de notre tentative de redécouverte du quark top, notre jeu de données de signal est produit en imposant une signature hadronique aux événements. Ceci implique donc que chaque jet d'un quark top possède une structure à trois sous-jets, ou *3-prong* [42]. Aussi, comme nous nous intéressons ici aux jets boostés, les produits de désintégration du quark top se retrouveront à l'intérieur de $R = 1.0$. Ce type particulier de jet de quark top ne représente qu'un faible échantillon du nombre total d'événement de ce type produit au LHC, soit de l'ordre de 1%.

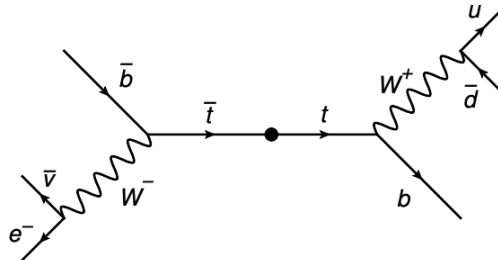


Figure 26. Production d'une paire de top anti-top au LHC. Le quark top se désintègre en un quark bottom et un boson W . Ce dernier peut se désintégrer leptoniquement (neutrino-lepton), comme à gauche, ou hadroniquement (quarks), comme à droite [47].

Les quarks top boostés sont dominés par le bruit de fond provenant des dijets QCD. Ces dijets proviennent d'un gluon ou d'un quark léger (u, d, c, s, b). La masse invariante de ce bruit de fond, reconstruite à partir des quadri-vecteurs des *topoclusters*, en opposition à celle du signal du top est affichée à la Figure 27, où nous utilisons les données de simulations ATLAS que nous décrirons en détails au chapitre 4. L'analyse incluse dans ce mémoire s'intéressera donc à la discrimination entre ce bruit de fond dominant, avec une signature typiquement formée d'un seul sous-jet 1(-*prong*), et les quarks top dans le régime boosté à une signature à trois sous-jets 3(-*prong*).

Nous notons que nous aurions également pu choisir de redécouvrir le boson de Higgs ou le boson W . Par contre, le boson de Higgs n'a été découvert que récemment (2012), alors que

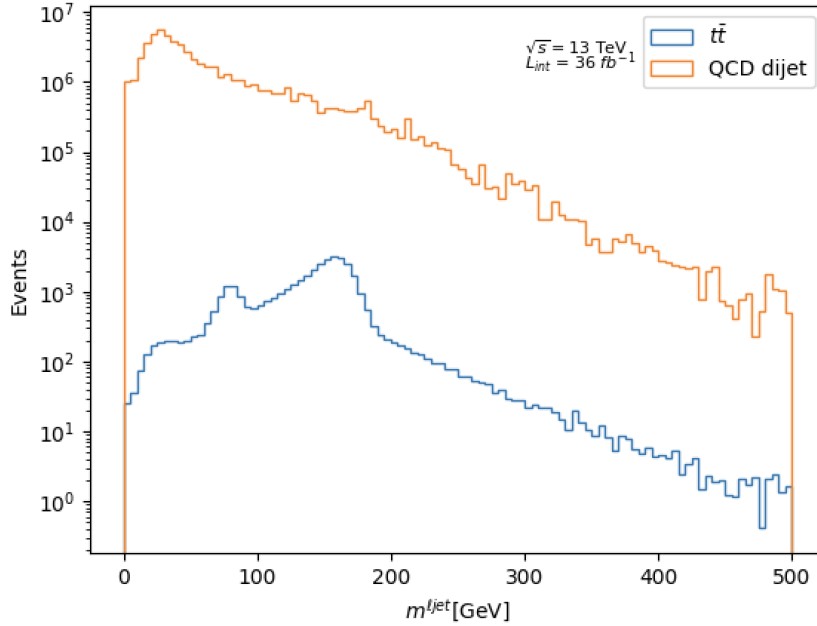


Figure 27. Masse invariante du premier jet, ou *leading jet* (m^{ljet}), pour le signal (bleu) et le bruit de fond (orange). Les masses sont obtenues par les données de simulation Monte Carlo de collision proton-proton à 13 TeV et ajustée à une luminosité intégrée de 36 fb^{-1} .

le quark top est de nos jours assez bien compris. De plus, sa masse culmine à un point un peu plus éloigné du bruit de fond des dijets QCD (autour de 60 GeV), comparativement au boson W ($m = 80.379 \text{ GeV}$) [126], ce qui augmente possiblement nos chances de discrimination entre les deux.

Plusieurs études ont démontré qu’il est possible de discriminer entre le quark top et le bruit de fond dominant issu de la QCD en utilisant des réseaux de neurones entraînés sur ce bruit de fond ou des techniques d’identification de sous-structure. Bien que les jets top boostés ne représentent qu’un faible échantillon des quarks top produits au LHC, il est possible de les identifier par-dessus ce bruit de fond pour tenter une recherche de nouvelles résonances, tel qu’exploré dans [19, 77].

Bref, notre terrain de jeu correspond à un bruit de fond dominant provenant des dijets QCD et d’un signal issu d’un « nouveau » quark anormal. Cependant, à cause du comportement de la constante de couplage de la QCD α_s à basse énergie, la théorie de la QCD devient non perturbative. Aux basses énergies, il nous est alors impossible de réaliser une approximation mathématique du processus de formation des jets à partir d’une solution connue et nous en n’avons pas de compréhension analytique, ce qui rend les calculs difficiles [116]. Ainsi, nous ne pouvons affirmer que nous comprenons complètement l’hadronisation

des jets au niveau théorique [148]. Nous nous retrouvons alors avec un ensemble de données complexes et un signal anormal à retrouver ; il s'agit du contexte idéal pour l'utilisation de l'IA. En effet, un algorithme d'apprentissage machine pourrait apprendre certaines caractéristiques de données que nous ne pouvons comprendre, et ainsi mieux performer dans la détection de signaux anormaux. Cependant, la première étape vers la détection d'anomalies par l'apprentissage machine est d'observer notre problème avec des méthodes simplifiées, sans et avec l'application de l'apprentissage machine. Dans la prochaine section, nous étudierons d'abord une méthode utilisant la distance de Wasserstein comme outil de détection d'anomalies.

3.2. Distance de Wasserstein

La distance de Wasserstein, ou EMD, est une métrique qui mesure la distance entre deux distributions de probabilité. Elle tire son nom de l'analogie des deux tas de terre, où elle correspond au travail minimal à accomplir pour transformer l'un des tas de terre en l'autre. Cette métrique peut être définie dans divers espaces, comme l'espace des collisions de particules.

Cette métrique a été adaptée à l'espace des jets, entre autres par les travaux des références [105] et [106], afin de calculer la quantité minimale de travail à accomplir pour transformer un événement en un autre. Dans cet espace, la métrique de EMD entre deux jets ε et ε' est définie selon les équations 16 [106].

$$\text{EMD}_{\beta,R}(\varepsilon,\varepsilon') = \min_{\{f_{ij}>0\}} \sum_{i=1}^M \sum_{j=1}^{M'} f_{ij} \left(\frac{\theta_{ij}}{R} \right)^\beta + \left| \sum_{i=1}^M E_i - \sum_{j=1}^{M'} E'_j \right|, \quad (16a)$$

$$\sum_{j=1}^{M'} f_{ij} \leq E_i, \quad \sum_{i=1}^M f_{ij} \leq E'_j, \quad \sum_{i=1}^M \sum_{j=1}^{M'} f_{ij} = E_{\min} \quad ; \quad (16b)$$

Ici, les sommes sont effectuées sur les particules contenues dans les deux jets contenant M et M' constituants. La quantité f_{ij} représente l'énergie déplacée entre les particules i et j , θ_{ij} est la distance angulaire entre les particules i et j , E est l'énergie d'une particule dans le jet et E_{\min} est la valeur minimale entre les deux énergies totales de chaque jet ($\sum_{i=1}^M E_i$ et $\sum_{j=1}^{M'} E'_j$). Le paramètre R est positif et permet de contrôler l'importance relative des deux termes de l'équation. Dans nos études, ce dernier est fixé à 1, tout comme l'exposant β . Cette définition nous indique que lorsque nous calculons le EMD entre deux

jets dans le détecteur ATLAS, par exemple, ceci revient à résoudre un problème de transport optimal entre les deux événements. Afin de comprendre l'utilité de cette métrique dans un cadre de détection d'anomalies, nous appliquerons celles-ci dans un échantillon de jets boostés QCD et top simulés.

L'échantillon que nous utilisons ici provient de simulations obtenues dans la référence [58]. Les jets QCD ont été simulés à partir des critères de jets avec un grand rayon R et reconstruits par l'algorithme d'anti- k_T avec un $R = 1.0$, sans élagage. Les jets produits par le logiciel `MadGraph` et des collisions de proton-proton à 13 TeV ont été reconstruits à partir des *particle flow objects*, soit les particules individuelles reconstruites par la combinaison des informations du détecteur interne et du calorimètre [22]. Les logiciels `Pythia8` et `Delphes` ont simulé les gerbes de partons et la réponse du détecteur. Les données ici n'incluent pas de *pile-up*. Un critère de sélection pour l'impulsion transverse des jets de $p_T > 450$ GeV/ c est imposé.

Afin de réduire les effets liés aux fluctuations statistiques et au positionnement des jets dans le détecteur, nous avons transformé les données. D'abord, nous avons effectué une rotation les centrant en $(0,0)$ dans l'espace (η, ϕ) , un *boost* pour renvoyer dans le référentiel au repos du jet et une rotation pour orienter l'axe central du jet verticalement. Nous restreignons également nos jets dans un intervalle d'impulsion de $(450,700)$ GeV/ c . Puis, nous avons échelonné l'impulsion transverse des particules i dans un jet par l'impulsion totale du jet selon l'équation $p_{Ti} = (p_{Ti}/p_T) * 600$, afin de centraliser l'impulsion des jets au centre de notre intervalle.

La Figure 28 permet de visualiser le calcul du EMD entre deux événements. Les couleurs rouge et bleu représentent chacune deux jets différents dans le plan de rapidité et d'angle azimutal (y, ϕ) . Les lignes noires représentent le chemin idéal pour transformer un jet en l'autre et la valeur correspondante de EMD est affichée en bas à droite de la figure. Ainsi, la métrique du EMD permet de visualiser les différences, jusqu'à un certain point, en termes de sous-structure des événements. Cette Figure illustre bien les structures à un sous-jet (*1-prong*) des QCD et à trois sous-jets des jets du top (*3-prong*).

Nous utilisons la métrique EMD comme score d'anomalie. Cette métrique pourrait s'avérer utile pour discriminer notre signal et notre bruit de fond, car elle risque de produire des valeurs plus élevées lors de la comparaison d'événements top-QCD que d'événements QCD-QCD. En effet, à cause des différences sous-jacentes dans la structure des événements top (*3-prong*) versus QCD (*1-prong*), nous misons sur le fait que la métrique devra accomplir un travail plus grand pour changer un événement QCD en un événement de quark top

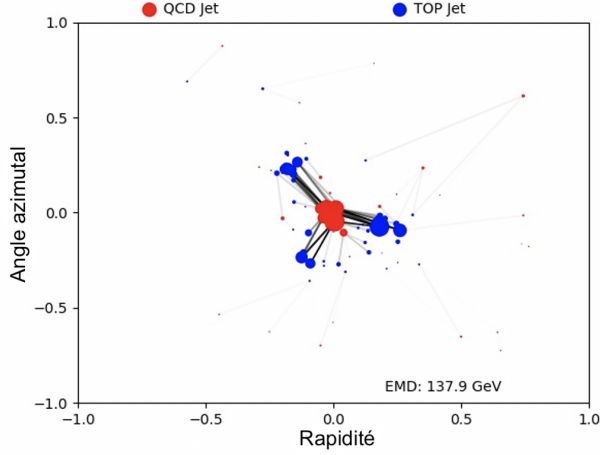


Figure 28. Transport optimal entre un jet de quark top de $p_T = 503.8 \text{ GeV}/c$ et un jet QCD de $p_T = 480.3 \text{ GeV}/c$. La valeur de EMD correspondant au transport optimal entre les deux événements est indiquée.

que pour changer un événement du bruit de fond en un autre. La métrique de EMD nous donnera ainsi des valeurs plus élevées lorsqu'un signal est impliqué.

Nous utilisons les informations de bas niveau des jets qui correspondent aux constituants sous la forme (p_T, η, ϕ) et nous calculons la distribution moyenne du EMD entre les jets d'un ensemble de données « d'entraînement » composé de bruit de fond QCD et un ensemble de données « test » composé d'un mélange de signal et de bruit de fond. Pour ce faire, nous calculons le EMD entre chaque jet de l'ensemble d'entraînement (17845 jets) et chaque jet QCD et de l'ensemble test (3553 jets). Nous prenons la valeur moyenne pour chaque jet de l'ensemble test pour obtenir notre distribution moyenne pour les QCD et nous répétons la procédure pour les jets top de l'ensemble test (3575 jets). Les deux distributions résultantes pour les jets QCD et top de l'ensemble test sont affichées à la Figure 29. Nous voyons clairement que la distribution moyenne du EMD est plus élevée pour les jets du quark top que QCD. Comme la métrique doit effectuer un travail plus grand pour transformer un jet de quark top en un jet du bruit de fond de QCD, nous nous retrouvons avec une distribution plus élevée pour nos signaux. Ceci nous indique que nos jets « anormaux » de notre ensemble test, soit nos jets de quarks top, diffèrent en termes de sous-structure de nos jets d'entraînement de QCD. La métrique EMD semble donc être un bon indicateur pour identifier les données qui semblent anormales ou hors de l'ordinaire.

La métrique EMD pourrait donc être un bon outil pour discriminer entre des jets normaux et anormaux. Pour analyser le pouvoir discriminant de la métrique, nous affichons la courbe ROC (Figure 30), de l'anglais *receiver operating characteristic* où la distribution moyenne du EMD est utilisée comme score. Les courbes ROC illustrent l'efficacité du signal (le taux

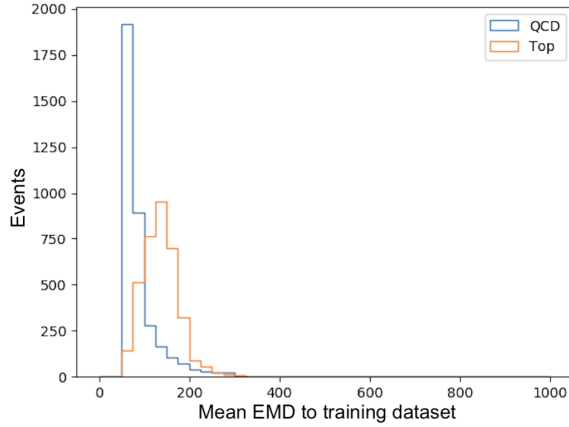


Figure 29. Distribution moyenne du EMD entre l'ensemble d'entraînement et les jets QCD et top de l'ensemble test.

de signal conservé) en fonction du rejet du bruit de fond (le taux de bruit de fond rejeté). Plus la courbe se rapproche du coin supérieur gauche, mieux est la performance. Le score AUC correspond simplement à l'aire sous la courbe. Ainsi, un score AUC se rapprochant de 1 indique une meilleure performance. Ici, nous obtenons un score AUC de 0.850, ce qui constitue une performance satisfaisante.

L'utilisation de la métrique EMD semble donc une méthode satisfaisante pour identifier des jets anormaux par rapport à un ensemble « d'entraînement ». Par contre, cette méthode ne nous permet pas nécessairement d'apprendre les caractéristiques intrinsèques de nos

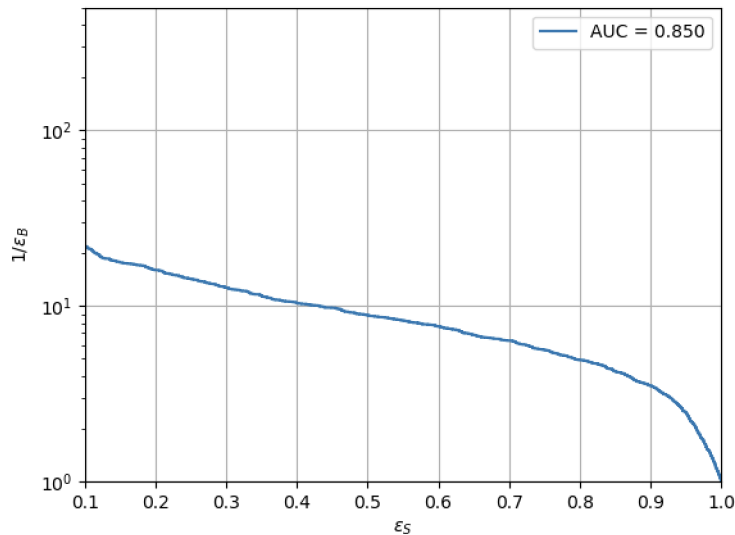


Figure 30. Courbe ROC obtenue en utilisant le EMD comme discriminant. La Figure affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l'efficacité du signal (ϵ_S). Plus la courbe se rapproche du coin supérieur droit, mieux est la performance.

échantillons. Son utilisation pourrait être pertinente pour vérifier la détection d'un nouveau signal, mais l'apprentissage machine pourrait possiblement contribuer à la détection d'anomalies en améliorant notre rejet de bruit de fond.

Pour réaliser notre analyse, nous utiliserons donc l'apprentissage machine, et plus spécifiquement la détection d'anomalies par les réseaux de neurones. Pour comprendre la méthode utilisée, il nous faut toutefois explorer quelques bases de l'apprentissage machine.

3.3. Les réseaux de neurones

Note : cette section se base sur les concepts théoriques présentés dans [85], sauf si une autre source est indiquée.

L'apprentissage machine est souvent utilisé pour réaliser des tâches trop complexes pour l'humain. L'objectif est d'apprendre à un algorithme à effectuer une tâche précise en entraînant celui-ci sur un ensemble de données ou de caractéristiques $\mathbf{x} \in \mathbb{R}^n$, où \mathbf{x} est un vecteur et \mathbb{R}^n l'ensemble des nombres réels de dimension n . En entraînant sur ces données, nous souhaitons que l'algorithme apprenne certaines caractéristiques importantes qui lui permettront de réaliser la tâche ciblée.

Les algorithmes d'apprentissage machine peuvent réaliser un grand nombre de tâches. Parmi celles-ci se trouvent la classification et la régression. L'objectif de la première est de déterminer une fonction $f : \mathbb{R}^n \rightarrow \{1, \dots, k\}$, tel que $y = f(\mathbf{x})$. Ici, la fonction f permet de classer les données \mathbf{x} dans une catégorie y allant de $\{1, \dots, k\}$. La tâche de régression, quant à elle, implique une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}$. L'algorithme de régression permet donc de prédire une valeur réelle à partir des données \mathbf{x} . Il existe aussi une multitude d'autres tâches comme l'estimation de densité, la traduction, la transcription et bien d'autres. Toutefois, la tâche qui nous intéresse ici est celle de la détection d'anomalies. Ici, l'algorithme qui s'entraîne dans un ensemble de données identifie celles qui semblent anormales. La détection d'anomalies est utilisée dans diverses applications, comme la détection de fraude ou de problèmes médicaux. Dans notre cas, nous utiliserons plutôt la détection d'anomalies pour trouver le quark top dans un bruit de fond dominant.

Pour déterminer la capacité d'un algorithme à réaliser une tâche, nous devons mesurer sa performance. Pour une tâche de classification, nous utilisons la précision, soit la quantité d'exemples bien classifiés, ou le taux d'erreur, soit la quantité d'exemples mal

classifiés. Pour des tâches qui impliquent des probabilités ou des valeurs continues, comme la régression et la détection d'anomalies, nous utilisons plutôt une métrique. Celle-ci retourne un score et mesure à quel point la réponse prédite par l'algorithme se trouve loin de la réalité.

Le vrai test de performance se trouve toutefois dans l'application de l'algorithme sur des données jamais vues à l'entraînement. C'est pourquoi nous séparons toujours nos ensembles de données en un ensemble d'entraînement, pour apprendre à l'algorithme à effectuer la tâche, et un ensemble de test qui nous indiquera la performance réelle de l'algorithme lors de son application sur de nouvelles données. Cette séparation nous permet d'éviter que l'algorithme apprenne par coeur l'ensemble d'entraînement.

Il existe différents types d'entraînement liés au concept de *supervision*. Les algorithmes *supervisés* utilisent les étiquettes des données durant l'entraînement. Ces algorithmes sont souvent utilisés dans les tâches de classification. Par exemple, un algorithme qui utilise le nom de fleurs à l'entraînement dans le but d'identifier différents types de fleurs (iris, tulipe, rose) à partir d'images est supervisé, puisque les étiquettes lui sont fournies. Les algorithmes *non supervisés* n'utilisent pas les étiquettes. Par exemple, un algorithme cherchant à regrouper les fleurs par caractéristiques (longueur de tige, type de pétales, couleur) sans avoir les étiquettes à l'entraînement est non supervisé. Cette approche est souvent utile lorsque l'ensemble de données possède plusieurs caractéristiques et que l'objectif est plutôt d'apprendre les propriétés et la structure des données. Toutefois, ce n'est pas la tâche qui définit le concept de supervision, mais bien si les étiquettes sont fournies ou non à l'algorithme. Il existe aussi plusieurs degrés de supervision. Nous pouvons avoir des algorithmes *semi-supervisés* où certaines données sont accompagnées d'une étiquette, alors que d'autres, non.

3.3.1. Les réseaux de neurones

Parmi tous les types d'algorithmes d'apprentissage machine qui existent, celui-ci qui sera utilisé ici se nomme « réseau de neurones ». Les algorithmes qui portent ce nom relient ensemble des couches de neurones artificielles, en analogie aux neurones du cerveau humain comme à la Figure 31. Ici, le cas le plus simple d'un réseau de neurones avec une seule couche centrale, ou cachée, est illustré à droite. Ces réseaux sont dits à propagation avant, car l'information s'y propage de gauche à droite, en partant des données d'entrée \mathbf{x} et en se propageant dans les couches cachées avant d'être envoyée dans la couche de sortie, où le réseau retourne sa prédiction $f(x)$.

Les fonctions f prédites par les modèles sont des approximations. Elles ne représentent jamais la réponse analytique exacte. D'ailleurs, le théorème d'approximation universelle nous garantit qu'il existe toujours une solution approximative au problème [63, 123]. La forme simplifiée de ce théorème nous indique qu'un réseau de neurones à une seule couche cachée peut trouver une approximation pour n'importe quelle fonction continue, mais ce résultat se généralise également aux réseaux qui possèdent plus d'une couche cachée et plus d'une neurone de sortie. Les réseaux de neurones peuvent donc être appliqués à un ensemble de problèmes.

Chaque couche cachée a une dimension, c'est-à-dire un certain nombre de neurones. L'information y produite par une neurone j peut être décrite par une équation de la forme [63, 123] :

$$y_j = \sigma \left(\sum_{k=0}^m w_{jk} x_k + b_j \right) \quad ; \quad (17)$$

Ici, les x_k représentent les entrées d'une dimension m , les w_{jk} sont les poids transmis à la neurone j et appris par le réseau et le b_j est le biais de la neurone j . La fonction σ se nomme une fonction d'activation et sert à transformer la sortie de la neurone. Généralement, une même fonction d'activation est appliquée à chaque neurone du réseau, mais la fonction d'activation à la couche de sortie diffère de celle utilisée dans les couches cachées.

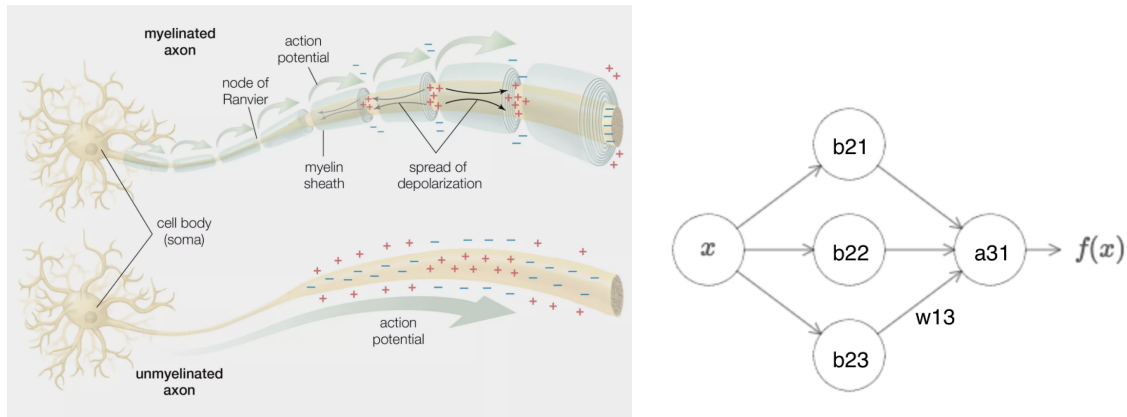


Figure 31. Comparaison entre les neurones du cerveau et un réseau de neurones artificielles. À gauche, des neurones traditionnelles qui propage un signal électrique vers l'avant [72]. À droite, un réseau de neurones avec une seule couche cachée qui propage l'information de gauche à droite [123]. Les b représentent les biais de chaque neurone, a l'activation de la neurone et w le poids transmis d'une neurone à l'autre (ces informations ont été ajoutées à l'image originale). La fonction $f(x)$ est celle prédite par le réseau.

La fonction d'activation σ appliquée à la couche de sortie du réseau de neurones permet généralement d'associer l'information du réseau à une probabilité en sortie, soit une valeur se trouvant entre l'intervalle $(0,1)$. Il existe une variété de fonction d'activation. Les plus communes sont affichées à la Figure 32. La fonction sigmoïde est une fonction logistique retournant une valeur entre $(0,1)$ qui correspond à une probabilité. La fonction *softmax* est une généralisation de la sigmoïde et souvent utilisée dans les données impliquant plusieurs classes (plusieurs étiquettes). La fonction ReLU (de l'anglais *rectifying linear unit*), envoie une valeur de 0 pour une valeur d'entrée négative, et devient la fonction linéaire pour un $x > 0$.

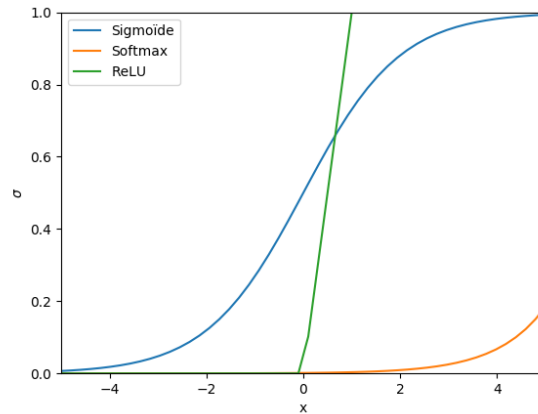


Figure 32. Fonctions d'activation les plus communes. Les fonctions sigmoïde et *softmax* permettent d'envoyer les données en sortie sur un intervalle entre $(0,1)$. La fonction ReLU prend la forme d'une fonction linéaire pour $x > 0$.

Enfin, nous différencions les réseaux de neurones dits profonds et peu profonds. Plus un réseau possède des couches cachées, plus le réseau sera profond, d'où l'appellation *Deep Neural Network* (DNN) en anglais.

3.3.2. Entraînement et descente du gradient

Nous avons vu que les réseaux de neurones prédisent une fonction $f(x)$ approximative. Plus précisément, cette fonction permet d'effectuer l'application $y = f(\mathbf{x}; \theta)$, où les y sont les valeurs prédites par le réseau. Les θ sont des paramètres appris par le réseau et qui permettent d'obtenir la meilleure approximation. Nous voulons donc que la fonction f se rapproche le plus possible de la vraie fonction analytique. Pour ce faire, nous entraînons le réseau pour trouver les meilleurs paramètres θ . L'entraînement des réseaux de neurones

devient un problème d'optimisation.

Soit un réseau avec les paramètres θ et soit les données $X, Y = (x_1, y_1), \dots, (x_n, y_n)$ représentant les paires valeur-étiquette d'entrées à l'entraînement. Nous pouvons obtenir une fonction dérivable de perte $L(X, Y, \theta)$. Cette fonction nous permet d'évaluer à quel point nos prédictions sont bonnes ou mauvaises. L'entraînement sert alors à minimiser la fonction de perte L , en altérant les paramètres.

Cette minimisation s'effectue par itérations avec la méthode de descente de gradient. La dérivée de la fonction L nous indique comment modifier les paramètres θ pour améliorer les prédictions y . Comme les données d'entrée \mathbf{x} ont souvent plus d'une dimension, nous utilisons le gradient. Le gradient pointe dans la direction où la fonction de perte augmente le plus rapidement. Ainsi, nous pouvons diminuer L en se déplaçant dans la direction opposée au gradient, d'où le terme « descente » de gradient. Ce dernier est calculé en fonction des poids du réseau par une méthode de rétropropagation (pour plus de détails, voir [85, 123]).

L'ajustement des paramètres au pas suivant θ' se fait par l'équation 18 en fonction des paramètres et de la perte au pas précédent. Les paramètres θ au temps initial sont généralement initialisés aléatoirement. La quantité ϵ s'appelle le taux d'apprentissage (*learning rate*) et nous permet de déterminer la grandeur des sauts de gradient pour ajuster la vitesse de convergence.

$$\theta' = \theta - \epsilon \nabla_{\theta} L(\theta) \quad ; \quad (18)$$

Il peut arriver que pour de gros ensembles de données, la convergence vers une solution finale soit longue. Nous utilisons alors une descente de gradient stochastique, où nous estimons le gradient sur des sous-ensembles des données (*batches*). Nous répétons cette procédure plusieurs fois (plusieurs époques) sur l'ensemble des données.

Pour éviter que le réseau stagne sur des points de selle dans le gradient, nous introduisons parfois un terme de moment dans l'équation de perte, comme le dernier terme à l'équation 19 et où le terme α posé manuellement permet de contrôler l'importance du moment.

$$\theta' = \theta - \epsilon \nabla_{\theta} L(\theta) + \alpha \nabla_{\theta-1} L(\theta) \quad ; \quad (19)$$

Un exemple de moment est le *Adaptive Moment Estimation* (ADAM). Ce dernier introduit divers moments correspondant à des fractions du gradient au pas précédent pour les additionner au gradient actuel [104].

Il existe une multitude de choix pour la fonction de perte L . L'erreur moyenne quadratique (ou *mean squared error*), est définie à l'équation 20. Ici, m représente le nombre d'événements dans les données d'entrée, \tilde{y} est la prédiction du réseau et y la valeur réelle.

$$\text{MSE} = \frac{1}{m} \sum_i \| y_i - \tilde{y}_i \|^2 \quad ; \quad (20)$$

Finalement, certains paramètres ne sont pas prédits par le réseau et doivent être ajustés à la main. Nous les appelons des hyperparamètres. Parmi ceux-ci se trouvent le taux d'apprentissage ϵ , le nombre de couches cachées et le nombre de neurones dans chaque couche. Il est essentiel d'optimiser ces hyperparamètres pour trouver la combinaison qui permet d'obtenir les meilleures prédictions. Généralement, nous les optimisons manuellement ou en effectuant un balayage sur un certain intervalle de valeurs.

Il est important de noter aussi qu'une grosse partie du travail se fait en amont. Le traitement initial des données avant l'entraînement est crucial afin de permettre à l'algorithme d'apprendre les caractéristiques essentielles et pour lui faciliter la tâche. Une pratique courante est d'échelonner les données entre (0,1) afin d'éliminer les effets liés aux ordres de grandeur. Ce traitement nous permet aussi de nettoyer et d'organiser les données et de concentrer l'attention du réseau sur les propriétés importantes des données et non sur des possibles fluctuations statistiques.

3.3.3. Généralisation

Après avoir entraîné le réseau, nous devons être capable de généraliser son application. La généralisation fait référence à la capacité d'un modèle à bien performer sur des données jamais vues. Nous voulons alors minimiser l'erreur sur l'ensemble d'entraînement, et l'erreur sur l'ensemble test. Pour ce faire, nous observons l'erreur de généralisation, qui se résume à l'erreur irréductible, le biais et la variance :

$$\text{Erreur} = \text{Irréductible} + \text{Variance} + \text{Biais}^2 \quad ; \quad (21)$$

L'erreur irréductible est liée au taux de Bayes et, comme son nom l'indique, ne peut être réduite [151]. Il s'agit alors de faire un compromis entre le biais et la variance. Lorsque nous favorisons le biais, le réseau a une faible capacité, c'est-à-dire qu'il peut trouver des approximations pour moins de fonctions. Par contre, le réseau performe bien sur des nouvelles données. Lorsque nous favorisons la variance, le réseau a une plus grande capacité, c'est-à-dire qu'il couvre plus de fonctions, mais sa performance varie beaucoup sur des nouvelles données.

La technique d'arrêt précoce ou *early stopping* permet d'arrêter l'entraînement quand l'erreur de généralisation sur un ensemble des données de validation ne s'améliore plus. Cette technique nous permet d'éviter le « *overfitting* », soit d'apprendre par coeur notre ensemble d'entraînement.

3.3.4. Les auto-encodeurs

Les auto-encodeurs (AEs) sont des algorithmes d'apprentissage non supervisés. Ce modèle combine deux réseaux de neurones, que nous appelons encodeur et décodeur (voir Figure 33). Ces deux réseaux possèdent généralement le même nombre de couches cachées, la même dimension d'entrée-sortie et le même nombre de neurones par couches (en ordre inversé). Le rôle de l'encodeur est de compresser les données d'entrées x pour en apprendre les caractéristiques saillantes. L'encodeur envoie ensuite les données compressées dans une couche centrale, ou couche latente, qui produit une représentation des données. Puis, le décodeur essaie de reproduire les données à partir de l'information fournie par la représentation latente. Comme l'encodeur et le décodeur sont eux-mêmes des réseaux de neurones, l'information des données est distribuée à travers ces réseaux, en plus de la couche latente. Pour mesurer la performance d'un auto-encodeur, il nous suffit alors de mesurer la différence entre les données en sortie et les entrées par une métrique quelconque, comme le MSE.

Durant l'entraînement, l'auto-encodeur cherche seulement à minimiser cette fonction de perte, ou erreur de reconstruction. Ceci peut parfois faire en sorte que l'espace latent apprenne les données d'entraînement par coeur. Pour contrer ce problème, il est possible d'imposer une structure à l'espace latent. Ceci correspond à un autre type d'algorithme, que nous appelons l'auto-encodeur variationnel (VAE). Cet algorithme permet de générer

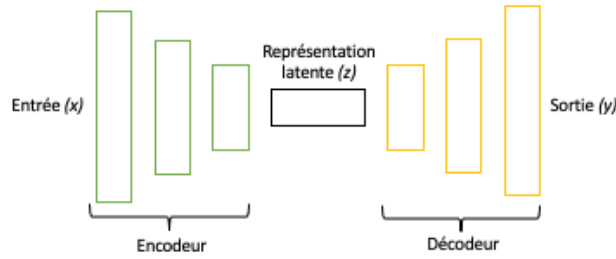


Figure 33. Schéma d'un auto-encodeur. À gauche, l'encodeur réduit la dimension des données d'entrées x et les envoie dans la couche latente. Puis, le décodeur génère des sorties y à partir de l'information de la couche latente.

de nouvelles données à partir d'un espace latent plus régulier.

Pour ce faire, nous demandons à ce que la représentation latente suive une distribution normale, comme visualisé à la Figure 34. Nous appliquons cette condition en minimisant une seconde fonction de perte à l'entraînement, soit la divergence Kullback-Leibler D_{KL} définie à l'équation 22 et qui est asymétrique [58]. Cette fonction mesure la « distance » entre deux distributions.

$$D_{KL}(Q(z|X) \parallel P(z)) = \sum_{x \in X} Q(z|X) \ln \left(\frac{Q(z|X)}{P(z)} \right) ; \quad (22)$$

Ici, la fonction $Q(z|X)$ correspond à l'application de l'entrée x à l'espace latent z , et la fonction $P(z)$ fait référence à la distribution normale. À l'entraînement, nous voulons minimiser le D_{KL} entre ces deux distributions, en plus de l'erreur de reconstruction. L'erreur à minimiser correspond donc à $L_{vae} = L_{reco} + \beta * L_{KL}$, où β est une valeur choisie manuellement et nous permet de déterminer la force de la régularisation de l'espace latent.

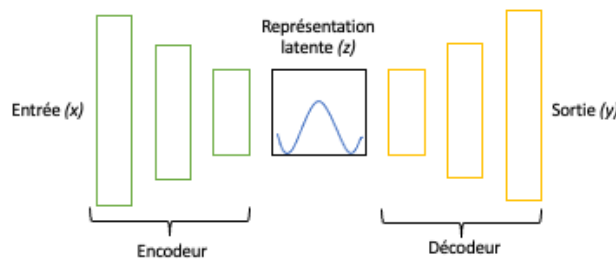


Figure 34. Schéma d'un auto-encodeur variationnel. À l'entraînement, nous demandons à l'espace latent de suivre une distribution normale. L'information encodée pourra être échantillonnée à partir de cette distribution.

En résumé, les données d’entrées sont d’abord encodées comme une distribution dans l’espace latent. Un point de l’espace est alors échantillonné, puis décodé. Finalement, l’erreur de reconstruction entre la sortie et l’entrée et la divergence de Kullback-Leibler sont calculées et la rétropropagation et la descente de gradient s’effectuent.

Lorsque nous utilisons les auto-encodeurs et les auto-encodeurs variationnels pour la détection d’anomalies, ces algorithmes performant bien sur les données d’entraînement et moins bien sur des données anormales ou aberrantes. Ainsi, lorsque nous appliquons un AE ou un VAE à de nouvelles données, nous pouvons utiliser divers scores ou métriques entre les données d’entrée et de sortie comme score d’anomalie. Par exemple, si nous choisissons le MSE comme score, sa valeur sera plus grande pour une donnée aberrante ou anormale.

Avant de tester la performance d’un auto-encodeur variationnel pour la redécouverte du quark top dans le bruit de fond de QCD dominant, nous devons savoir quelle est la performance idéale à laquelle nous pouvons aspirer. Pour ce faire, nous commencerons par présenter la performance d’un modèle supervisé d’apprentissage machine pour la détection du quark top dans les jets boostés.

3.4. Identification des jets de quark top boostés avec modèle supervisé

Dans la présente section, nous explorerons un cas particulier de réseau de neurone supervisé dédié à l’identification des quarks top boostés. L’objectif ici sera d’évaluer quelle est la performance « optimale » à laquelle un réseau de neurones peut aspirer. En effet, le cas supervisé où le réseau connaît la nature des événements et est entraîné à la fois sur le bruit de fond de QCD que sur le signal du top est le cas le plus encadré possible. Plus tard, nous entraînerons notre réseau uniquement sur notre bruit de fond, et ceci nous permettra de comparer notre performance à celle du cas optimal.

Plusieurs études ont tenté d’identifier les quarks top boostés dans les données d’ATLAS par divers moyens. Ici, nous décrirons plus en détails le réseau de neurones supervisé d’identification des quarks top en opposition aux jets QCD et intégré aux logiciels d’ATLAS, soit le TopoDNN [29]. Nous notons toute de même que plusieurs autres études ont également tenté cette identification à l’aide de l’apprentissage machine, entre autres [103].

Le cadre d’analyse du TopoDNN comporte des similarités au nôtre. D’abord, les quarks

top boostés sont utilisés comme signaux et les événements utilisés comme bruit de fond correspondent aux simulations de dijets issus de la QCD, donc de gluons et de quarks légers. De plus, le réseau est entraîné sur les quadri-vecteurs des *topoclusters*. Nous utilisons ainsi sensiblement les mêmes informations des jets à l’entraînement. Par contre, nous rappelons que le TopoDNN est entraîné à la fois sur le signal et le bruit de fond, ce qui n’est pas notre cas.

Le TopoDNN utilise seulement les dix premiers constituants des jets à l’entraînement et à l’évaluation du modèle, comme ceux-ci transportent la majorité de l’impulsion du jet. Les constituants sont utilisés dans la forme (p_T, η, ϕ) et sont transformés en (η, ϕ) par une translation, une rotation et un *flip* selon leur topologie à trois sous-jets (*3-prong*). Une spécification d’impulsion transverse des jets de $p_T > 450 \text{ GeV}/c$ est aussi imposée.

Le bruit de fond de la QCD est produit en tranches d’impulsion transverse allant de 200 à 2500 GeV/c . Les poids des événements individuels, ou « événement par événement », sont appliqués lors de l’évaluation du modèle afin de retrouver la distribution lisse et tombante de l’impulsion transverse des vraies données. Ces poids correspondent à un facteur multiplicatif appliqué au nombre d’événements observés dans la simulation Monte Carlo, afin que ceux-ci correspondent au nombre d’événements attendus dans les vraies données pour une certaine luminosité intégrée. Il est à noter que les poids sont en même temps utilisés pour corriger la simulation pour que celle-ci corresponde le plus possible aux vraies données. Par exemple, le poids associé au *pile-up* (voir section 2.3.2) ajuste la distribution du nombre de *pile-up* dans la simulation à celui attendu dans les vraies données.

Les signaux correspondent aux quarks top à grande impulsion transverse et issus du processus de désintégration de nouvelle physique de $Z' \rightarrow t\bar{t}$. Le boson Z' est simulé avec des masses allant de 400 à 5000 GeV/c^2 . La signature des jets du quark top utilisée dans cette analyse est celle du « lepton plus jet ». Les dijets et les signaux sont simulés par le déclenchement de jets à grand rayon R et reconstruits avec $R = 1.0$ par l’algorithme d’anti- k_T , avec élagage. Enfin, le *pile-up* est considéré dans les simulations.

Le modèle consiste en un réseau de neurones profond avec quatre couches cachées de 300, 102, 12 et 6 neurones. Le modèle est supervisé, puisque les étiquettes sont fournies à l’entraînement. Le modèle est également entraîné sur un mélange de signal et de bruit de fond. La fonction d’activation ReLU relie les couches cachées et la fonction sigmoïde est utilisée à la sortie. L’algorithme ADAM opère durant 40 époques et la méthode de *early stopping* est implémentée. Finalement, la fonction de perte du *binary cross-entropy*, qui est utilisée dans les cas de classification binaire, est appliquée ici. Les hyperparamètres sont

optimisés manuellement, en explorant des intervalles de valeurs pour le nombre de couches (entre 4 et 6) et le nombre de neurones par couche (entre 40 et 10000).

Pour comparer la performance du TopoDNN utilisant les informations de bas niveau (*topoclusters*) avec divers autres algorithmes présentés dans cette étude qui utilisent les informations de haut niveau, nous observons la courbe ROC (Figure 35). Ici, plus la courbe se dirige vers le coin supérieur droit, plus le modèle performe pour discriminer le bruit de fond du signal.

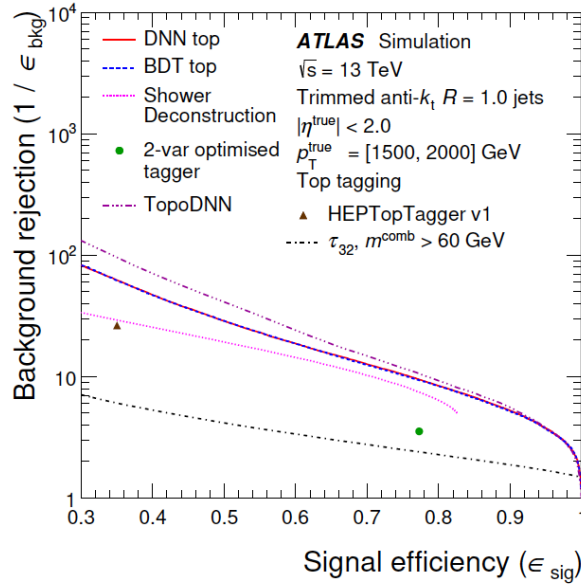


Figure 35. Courbes ROC pour le TopoDNN intégré dans les logiciels d’ATLAS. La Figure affiche le rejet du bruit de fond ($1/\epsilon_{\text{bkg}}$) en fonction de l’efficacité du signal (ϵ_{sig}). Les jets considérés sont à grande impulsion transverse [29].

Le TopoDNN performe mieux que d’autres algorithmes utilisant des variables de haut niveau. En effet, les autres modèles exploitent des variables calculées à partir des quadri-vecteurs des constituants des jets, telles que τ_1 , τ_{32} [147] et D_2 [110] qui contiennent de l’information sur la sous-structure des jets. Le TopoDNN, bien qu’il n’utilise que l’information des constituants eux-mêmes, semble donc un outil très efficace et compétitif avec les autres modèles pour identifier les jets de quarks tops à haut p_T .

Enfin, l’objectif de notre propre analyse sera d’observer les jets boostés qui pourraient inclure des traces de nouvelle physique. Ainsi, nous ne ferons aucune supposition sur la nature de ces jets ou sur la nature de cette nouvelle physique. Cependant, comme nous souhaitons tester la performance de notre approche avant d’effectuer notre recherche

générale, nous utiliserons le quark top comme exemple de signal de « nouvelle physique » à retrouver dans nos données de simulation. Ces jets de quark top boostés, qui ont une grande masse et une signature particulière à trois sous-jets, contrasteront avec notre bruit de fond dominant issu de la QCD. Le TopoDNN nous servira alors de modèle de référence pour comparer notre propre performance dans la tentative d'identification du quark top. Nous utiliserons un réseau de neurones non supervisé pour notre détection d'anomalies, correspondant à un auto-encodeur variationnel entraîné uniquement sur notre bruit de fond, dans le but de retrouver le quark top par-dessus le bruit de fond provenant des dijets QCD.

3.5. Diversité et inclusion en intelligence artificielle

Ce projet de recherche est en partie financé par l'Institut de valorisation des données (IVADO), dont l'une des priorités est de valoriser l'équité, la diversité et l'inclusion en intelligence numérique : « à IVADO, nous sommes convaincus qu'une culture universitaire inclusive stimule la réflexion et la créativité et permet d'attirer, de motiver et de retenir des talents diversifiés. Pour toutes ces raisons, les valeurs d'équité, de diversité et d'inclusion (ÉDI) sont indissociables de notre mission » [99]. Dans ce mémoire, nous visons à mettre en pratique ce cadre de référence, notamment l'axe 3, qui consiste à « renforcer la qualité, les retombées et la pertinence sociale de la formation et des travaux de recherche en science des données par la prise en compte des enjeux de diversité et de biais dans la production du savoir ». Nous développons dans ce qui suit une réflexion de cet objectif.

Contrairement aux collisions de particules dans l'anneau de LHC, la recherche dans le milieu académique ne s'effectue pas dans le vide. Même si certains projets en physique, en sciences des données et en intelligence artificielle n'adressent pas des enjeux sociaux, les techniques développées sont souvent transférées à d'autres domaines. Les physiciens et physiciennes qui terminent leurs études de 2^e et 3^e cycles et qui entrent sur le marché de l'emploi en tant que spécialistes de l'IA et des données vont probablement manipuler des données touchant les humains un jour ou l'autre. Si ces personnes ne sont pas informées sur les biais en IA, elles perpétueront les nombreuses inégalités.

Tout comme en physique, l'IA est dominée par les hommes. Dans le milieu de l'emploi, la proportion de femmes tourne autour de 25% [11, 99, 134, 154]. Dans le milieu universitaire, les femmes représentent 20,6% du corps professoral dans les domaines de mathématiques, d'informatique et des sciences de l'information [99]. Alors que le recrutement et la rétention des femmes dans ces domaines est difficile, le problème est amplifié

pour les groupes minoritaires en milieu universitaire, tels que les étudiant(e)s de première génération, les étudiant(e)s autochtones et minorités visibles et les étudiant(e)s en situation de handicap. Comme le mentionne IVADO dans son cadre de référence sur l'équité, la diversité et l'inclusion, « la promotion de la diversité et de l'inclusion répond d'abord à un enjeu de justice sociale et d'équité » [99]. Augmenter la représentation de ces groupes minoritaires en IA est d'abord un enjeu d'équité et d'égalité des chances.

La sous-représentation de certains groupes dans un milieu est souvent le reflet d'un climat organisationnel peu inclusif [16, 41, 46, 160], où le climat est fortement influencé par les comportements et les perceptions des individus. Ce climat est souvent généré par un manque de sensibilisation aux enjeux d'équité, de diversité et d'inclusion. Ceci laisse alors place à l'apparition de biais qui se répercutent dans la production du savoir en science et technologie. Dans un domaine tel que l'IA, il a été observé que des biais peuvent se transmettre dans les algorithmes [67]. Nous nous retrouvons alors avec des algorithmes supposément objectifs dont les résultats sont influencés par de nombreux biais et préjugés intégrés dans le tissu de la société. Selon le contexte d'application de ces algorithmes, ces biais peuvent créer ou même accentuer des inégalités déjà présentes. Ne pas en tenir compte limite les bénéfices potentiels pour la société d'un bassin diversifié de spécialiste en données.

Par ailleurs, la manière dont l'IA exploite les bases de données la rend particulièrement sensible à reproduire et renforcer certaines inégalités. Dans le livre *Data Feminism*, les autrices présentent plusieurs exemples de transmission des biais dans les algorithmes [67] : l'incapacité d'un logiciel de détection faciale à reconnaître les personnes noires et la préférence des CV de candidats masculin et l'évaluation de cas par les services de protection de la famille de la Pennsylvanie. Dans le contexte québécois, nous pensons à l'évaluation de cas par la Direction de la protection de la jeunesse (DPJ) à l'aide de l'IA [149]. Ces cas mettent en relief, d'une part, le danger d'automatiser des inégalités déjà présentes, et, d'autre part, que l'importance d'avoir des bases de données diversifiées pour contrer les biais. Adopter le « féminisme des données », c'est ainsi adopter « une façon de réfléchir aux données, leur utilisation et leurs limites, par une façon qui est informée par les expériences, l'engagement et une perspective intersectionnelle » [67] (traduit directement de l'anglais).

Également, derrière les chiffres se cachent des expériences diverses et complexes. Ainsi, lorsque nous transformons les expériences individuelles en données, nous perdons en complexité et en subtilité. Aussi, les données en physique, comme celles de toutes sciences, ne tombent pas du ciel. De nombreuses personnes ont travaillé à la transformation et l'analyse de ces données. Ces tâches nécessaires aux contributions scientifiques ont été historiquement (et sont toujours) attribuées aux femmes et sont souvent occultées [109, 113].

Enfin, la production et l'entreposage des données représente un travail de grande échelle qui nécessite beaucoup de ressources humaines et financières. Alors, la propriété des données revient souvent à des organismes de grande envergure : les universités, les gouvernements et les corporations, et ce sont majoritairement ces entités qui en bénéficient. Souvent ce sont les personnes marginalisées qui en souffrent, les institutions n'ayant pas pris en compte leurs intérêts. Heureusement, le CERN a pris conscience dans les dernières années de l'importance de la démocratisation des données de recherche et a rendu disponible certaines de ces données par les projets de Open Data [34].

Pour conclure, la sensibilisation aux enjeux d'équité, de diversité et d'inclusion est essentielle pour prendre conscience des problèmes de représentation en IA, de l'utilisation des données et des biais inconscients de société transmis aux algorithmes. En physique, tout comme en informatique, nous avons souvent tendance à évacuer ces questionnements de notre recherche. Ces aspects sont cependant importants et indissociables de notre recherche, tout comme le mentionne le troisième axe d'IVADO, d'autant plus que les universités ont la responsabilité, le pouvoir et les ressources pour former la relève d'une façon responsable. Une formation sur les biais en intelligence artificielle est cruciale pour toutes les personnes qui effectuent un projet de recherche avec l'IA, même si celui-ci ne touche pas directement les humains.

3.5.1. Conclusion

Nous avons déterminé dans le présent Chapitre quel bruit de fond et quel signal nous utiliserons dans notre recherche de nouvelle physique préliminaire. Nous avons exploré une méthode simplifiée de détection d'anomalies par la métrique de EMD. Puis, nous avons présenté quelques bases de l'apprentissage machine et compris le fonctionnement des auto-encodeurs variationnels que nous utiliserons dans notre redécouverte du quark top. Finalement, nous avons étudié la performance idéale atteinte par le TopoDNN d'ATLAS à laquelle nous pouvons espérer. Dans le prochain chapitre, nous décrirons plus en détails nos échantillons de simulation et l'architecture de notre auto-encodeur variationnel. Enfin, nous analyserons nos premiers résultats et discuterons de la performance de notre modèle dans l'identification du quark top boosté dans nos données.

Chapitre 4

Résultats et discussion

Les trois derniers chapitres ont mis la table pour notre recherche de nouvelle physique indépendante des modèles théoriques. Maintenant que nous avons les connaissances nécessaires sur le modèle standard, l'expérience ATLAS, les jets boostés, ainsi que l'apprentissage machine, nous entraînerons un réseau de neurones dans un ensemble de jets boostés du MS. De cette façon, le réseau, qui apprendra la structure des jets du MS, aura de la difficulté à reconstruire les jets issus de la nouvelle physique. Pour analyser la capacité de notre réseau à identifier ces jets « anormaux », nous tenterons d'abord de retrouver une particule bien connue dans un ensemble de jets boostés simulés : le quark top. Dans la première section de ce chapitre, nous décrirons nos échantillons de simulation de jets boostés qui jouent les rôles de signaux et de bruit de fond. Dans la deuxième section, nous présenterons l'architecture du réseau, soit notre auto-encodeur variationnel, et les détails techniques liés à son entraînement et au traitement de nos données. Nous présenterons aussi notre méthode d'analyse pour tenter de retrouver le quark top dans le spectre de masse des échantillons à partir de l'information apprise par notre réseau de neurones. Dans la troisième section, nous pourrions voir nos premiers résultats, et nous explorerons un problème de corrélation entre la masse de nos échantillons et les métriques d'anomalies de notre réseau, qui pourrait nuire à notre recherche. Enfin, nous proposerons à la quatrième section une méthode de semi-supervision pour résoudre ce problème, et nous explorerons nos résultats finaux en lien avec cette nouvelle méthode.

4.1. Description des échantillons de simulation

Pour notre étude, nous utilisons des échantillons générés à partir de simulations Monte Carlo (MC) pour le signal et le bruit de fond. Ces événements de dijets furent simulés par la collaboration ATLAS à partir de collisions proton-proton à une énergie de centre

de masse de 13 TeV. Celles-ci furent d’abord générés avec le logiciel `Pythia8`, qui permet de générer des événements du modèle standard par des collisions à hautes énergies [143]. À cette étape de génération, le logiciel `Pythia8` obtient, pour le bruit de fond, les éléments de matrice du processus du modèle standard $pp \rightarrow jj$ (dijet). Dans la chaîne de production, les particules stables ayant un temps de vie supérieur à 10^{-10} s sont ensuite propagées dans le logiciel `GEANT4`, qui simule l’interaction des particules avec le matériel du détecteur, ainsi que la géométrie de ce dernier [81]. Ensuite, une étape de numérisation permet de formater la sortie de simulation précédente, pour que les données de simulation ressemblent aux signaux électriques retournés par le détecteur. Enfin, l’étape de reconstruction, qui est exactement la même pour les vraies données, nous permet de reconstruire les traces, cellules, vertex, énergie manquante et jets, tel que décrit à la section 2.3.

La collaboration ATLAS a généré les événements de dijets QCD par tranches d’impulsion transverse pour obtenir assez d’événements sur l’ensemble de la distribution d’impulsion attendue au LHC. Pour que nos données simulées reproduisent le spectre d’impulsion lisse et tombant attendu dans les vraies données, nous produisons et appliquons par la suite les poids « événement par événement », de la même façon qu’à la section 3.4. Le Tableau 1 répertorie les tranches d’impulsion et la section efficace effective des différents échantillons JZW de dijets QCD.

Nous retenons uniquement les événements qui possèdent deux jets avec une impulsion $p_T > 200$ GeV/ c , puis les événements déclenchés par le *single-jet trigger* qui possèdent au moins un grand jet, soit un jet reconstruit avec un rayon $R = 1.0$ dans le plan (η, ϕ) , ayant une impulsion $p_T > 450$ GeV/ c et un $|\eta| < 2.0$. Pour notre analyse, nous utilisons seulement l’information du premier grand jet dans les événements. Après ces sélections, chaque échantillon JZW inclut entre 35 000 et 16 millions de jets simulés.

Les signaux de quark top boostés furent également générés par la collaboration ATLAS à partir des logiciels `Pythia8` et `Powheg`. Ce dernier permet de générer les partons, ainsi que les diagrammes d’ordre supérieur, ou *Next to Leading Order* (NLO) [10, 78]. De la même façon que pour les dijets, les logiciels `Pythia8` et `Powheg` obtiennent pour le signal les éléments de matrice des processus $pp \rightarrow t\bar{t}$ et $t \rightarrow Wb$, ainsi que toutes les désintégrations du modèle standard pour le W . Il fut demandé à ce qu’aucun lepton ne soit présent dans l’état final. Les mêmes étapes de production que pour le bruit de fond sont ensuite réalisées. La collaboration ATLAS a généré les événements par tranches de masse invariante des deux quarks top initiaux $m_{t\bar{t}}$, répertoriées au Tableau 2 avec la section efficace effective de chaque échantillon. Pour notre analyse, nous utilisons seulement l’information du premier grand jet. Nous appliquons les mêmes sélections que pour les événements

Échantillon	Impulsion transverse (<i>truth</i> p_T) (GeV/ c)	Section efficace effective (fb)
JZ3W	160-400	8469512.64
JZ4W	400-800	135305.29
JZ5W	800-1300	4207.84
JZ6W	1300-1800	242.70
JZ7W	1800-2500	6.37
JZ8W	2500-3200	6.36
JZ9W	3200-3900	0.24
JZ10W	3900-4600	$7.07 * 10^{-3}$
JZ11W	4600-5300	$1.13 * 10^{-4}$
JZ12W	5300- ∞	$4.42 * 10^{-7}$

Tableau 1. Description des échantillons de dijets QCD. Les événements sont générés par tranches d'impulsion correspondant aux divers échantillons JZW.

dijets QCD. Après ces sélections, chaque échantillon contient entre 15 000 et 100 000 jets simulés. Nous appliquons également les poids « événement par événement » pour notre signal.

Échantillon	Masse invariante (m_{tt}) (GeV/ c^2)	Section-efficace effective (fb)
410284	1100-1300	3231.64
410285	1300-1500	1315.68
410286	1500-1700	571.23
410287	1700-2000	350.13
410288	2000-14 000	199.15

Tableau 2. Description des échantillons de jets de quark top boostés. Les divers échantillons incluent les événements générés par tranches de masse invariante des deux quarks top initiaux (m_{tt}).

Le spectre de l'impulsion transverse à la Figure 36 nous permet de voir l'allure tombante du bruit de fond des dijets QCD. Ici, les simulations ont été ajustées à une luminosité intégrée de 36 fb^{-1} . Ceci signifie que le nombre de jets indiqué sur l'axe vertical correspond au nombre réel attendu dans 36 fb^{-1} de données, soit les données enregistrées de 2015 à 2016 par l'expérience ATLAS. Ainsi, le spectre d'impulsion du quark top se trouve environ deux ordres de grandeurs sous le bruit de fond de QCD, tel que visible sur la Figure 36 de droite. Il en va de même pour le pic à la masse du quark top ($m = 172.76 \text{ GeV}$) à la Figure 27.

Enfin, nous devons trouver l'architecture idéale permettant au réseau de différencier le bruit de fond issu de la QCD et de signature typique 1-*prong* de notre signal provenant du quark top ayant une signature typique à 3-*prong*. De plus, pour retrouver le pic de masse du quark top

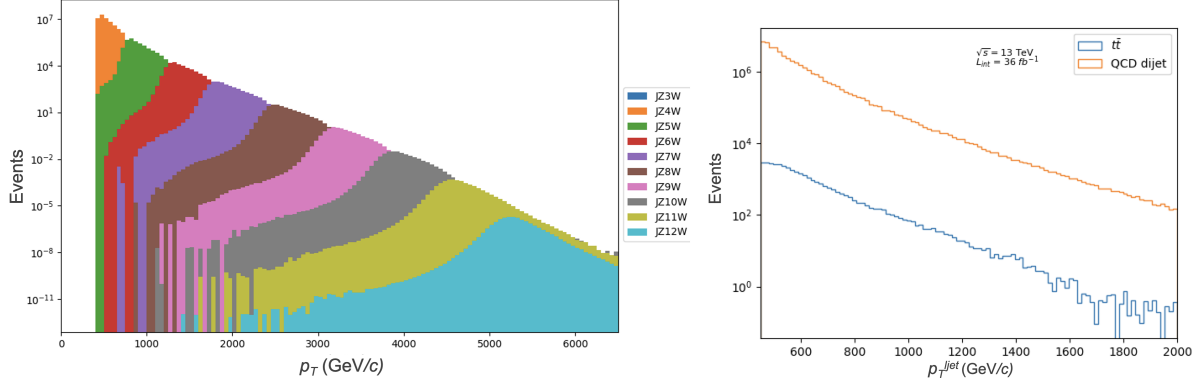


Figure 36. Impulsion transverse du premier jet (p_T^{1jet}) pour le signal et le bruit de fond. À gauche, nous voyons le spectre d'impulsion tombant du bruit de fond des divers échantillons JZW de dijets QCD. À droite, nous voyons le spectre d'impulsion du signal en comparaison au bruit de fond. Les jets sont obtenus par des simulations Monte Carlo de collisions proton-proton à $\sqrt{s} = 13$ TeV et ajustées à une luminosité intégrée de 36 fb^{-1} .

dans le spectre de masse invariante de notre échantillon d'évaluation, nous devons déterminer une façon de réduire le bruit de fond de QCD, qui est initialement deux ordres de grandeur plus important que notre signal. Dans la prochaine section, nous décrirons l'architecture de notre réseau, ainsi que notre méthode d'analyse pour retrouver le quark top dans le spectre de masse invariante de notre échantillon de simulation.

4.2. Architecture du réseau et méthode d'analyse

Pour notre tentative de « redécouverte » du quark top, nous exploitons un auto-encodeur variationnel. D'abord, nous devons déterminer l'architecture de notre réseau et faire le choix de nos hyperparamètres. Par la suite, nous pourrions entraîner notre réseau sur le bruit de fond et établir une méthode d'analyse pour retrouver notre signal. La prochaine sous-section présentera notre choix d'architecture pour notre modèle.

4.2.1. Description de l'architecture

Nous entraînons notre auto-encodeur variationnel (beta-VAE) sur les informations de bas niveau des jets correspondant aux amas topologiques (*topoclusters*). Nous choisissons d'utiliser les informations de bas niveau plutôt que les variables de haut niveau, car ceci pourrait permettre au réseau d'apprendre de l'information pertinente qui pourrait ne pas être contenue dans les variables calculées à partir des quadri-vecteurs des constituants des

jets. Ainsi, nous reprenons une approche similaire à celles de [40, 43, 75, 95].

Les données d'entrée des *topoclusters* prennent la forme de quadri-vecteurs de type (E, p_x, p_y, p_z) . Pour l'entraînement et l'évaluation du réseau, nous utilisons seulement les 20 premiers constituants des jets, en ordre décroissant de p_T , car ceux-ci transportent la majorité de l'impulsion totale des jets. La variation de ce nombre a peu d'impact sur la performance du réseau. Nous transformons également les jets au préalable en les centrant en $(0,0)$ en (η, ϕ) , en leur imposant un *boost* vers leur référentiel au repos et en effectuant une rotation pour aligner leur axe central verticalement. De plus, avant d'envoyer les données dans le réseau, nous échelonnons les quadri-vecteurs des jets à l'aide du *Quantile Transformer* [142] en prenant $n_{\text{quantiles}} = 100000$, ce qui nous permet de rediriger les valeurs des données entre 0 et 1 et ainsi de faciliter la convergence du réseau.

Pour éviter que notre réseau apprenne notre ensemble d'entraînement par coeur, nous séparons nos données en un ensemble d'entraînement de dijets QCD et un ensemble de validation, qui représentent chacun 80% et 20% de l'échantillon fourni à l'entraînement. Notre ensemble d'entraînement possède ainsi 480 000 jets QCD simulés et notre ensemble de validation 120 000 jets simulés. Ces ensembles contiennent des jets provenant des échantillons de JZ3W à JZ12W, mélangés aléatoirement.

Notre ensemble d'évaluation (ou test) contient 20 millions de dijets QCD et 203 942 jets top boostés simulés. Les poids « événements par événements » sont appliqués lors de l'évaluation pour reproduire les résultats attendus dans les vraies données à une luminosité intégrée de 36 fb^{-1} .

Nous avons optimisé les hyperparamètres de notre réseau à la main, en effectuant quelques essais dans des intervalles de valeurs. Le Tableau 3 résume les valeurs que nous avons choisies. De plus, nous avons appliqué l'algorithme ADAM pour l'optimisation à l'entraînement avec les paramètres par défaut. Nous effectuons l'entraînement par descente de gradient stochastique sur des sous-ensembles (*batches*), pendant 100 époques.

Enfin, nous utilisons l'équation 23 comme fonction de perte à l'entraînement, avec L_{reco} l'erreur quadratique moyenne (MSE). Ici, β est une valeur déterminée à la main et permet de contrôler la force de la régularisation de l'espace latent. Pour un $\beta = 0$, nous retrouvons un modèle d'auto-encodeur.

$$L = L_{\text{reco}} + \beta L_{\text{KL}} \quad ; \quad (23)$$

Hyperparamètre	Valeur
Nombre de couches cachées	3 (encodeur) 3 (décodeur)
Nombre de neurones cachées	256,128,64 (encodeur) 64,128,256 (décodeur)
Taux d'apprentissage	0.001
Activation (couches cachées)	ReLU
Activation (sortie)	Linéaire
Dimension de l'espace latent	10
β	0.1
Grandeur des <i>batches</i>	1000

Tableau 3. Hyperparamètres de l'auto-encodeur variationnel.

Le terme L_{KL} correspond à la divergence de Kullback-Leibler entre la distribution postérieure, qui encode l'information sur la structure latente d'un événement, et la distribution normale. Dans notre cas, ceci équivaut à l'équation 24, où i est sommé sur la dimension de l'espace latent et avec σ et μ représentant la variance et la moyenne de la distribution.

$$L_{\text{KL}} = -\frac{1}{2} \sum_i^N (1 - \sigma_i^2 - \mu_i^2 + \log(\sigma_i^2)) \quad . \quad (24)$$

4.2.2. Méthode d'analyse

Maintenant que nous avons choisi l'architecture de notre auto-encodeur variationnel, que nous nommerons beta-VAE, nous entraînons celui-ci sur notre échantillon d'entraînement composé de dijets QCD, en espérant que le réseau apprenne des informations sur la sous-structure de ces événements. Nous évaluons ensuite notre beta-VAE sur notre mélange de dijets QCD et de jets de quark top et nous calculons un score d'anomalie. Cependant, comme nous souhaitons retrouver le quark top dans le spectre de masse invariante de notre échantillon d'évaluation, nous effectuerons une « recherche de dos d'âne » (*bump hunt*). Cette procédure consiste à chercher un pic (ou dos d'âne) dans le spectre de masse lisse de l'échantillon d'évaluation. Ici, nous tenterons de retrouver le pic de masse du quark top, mais dans une recherche générale, ce pic pourrait correspondre à n'importe quel signal de nouvelle physique.

Nous illustrons la procédure correspondante à la Figure 37. Pour démontrer la capacité de notre réseau à identifier un nouveau signal, ici le quark top, nous effectuons des sélections supplémentaires à partir des prédictions du beta-VAE. Dans les « recherches de dos d’âne » traditionnelles réalisées au LHC, nous utilisons l’information du spectre de masse invariante pour estimer et réduire le bruit de fond dans une certaine région. Un algorithme balaie ensuite le spectre pour trouver un excès, ou dos d’âne, correspondant à une nouvelle particule ou résonance.⁴ Nous tenterons une stratégie similaire. Cependant, le signal est initialement noyé par le bruit de fond dominant du modèle standard. Pour pouvoir retrouver notre signal, nous devons réduire au maximum notre bruit de fond dominant pour essayer de faire ressortir le signal du quark top dans notre spectre de masse.

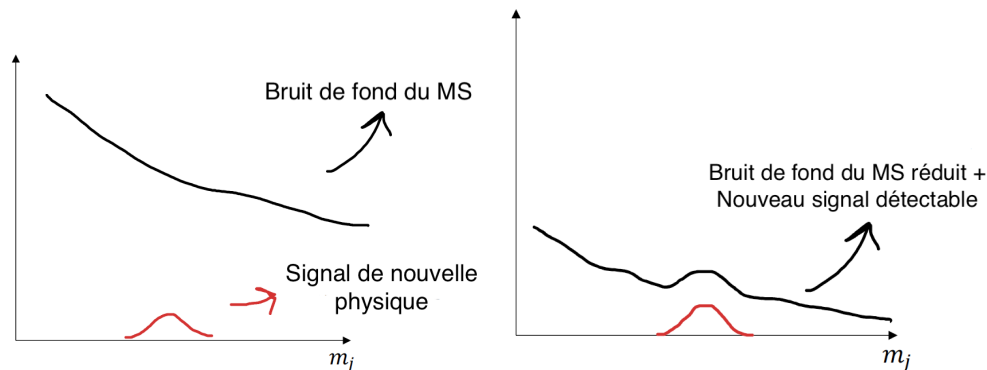


Figure 37. Effet espéré de la sélection à partir de la métrique d’anomalie sur le spectre de masse invariant. À gauche, le spectre de masse invariante du premier jet (m_j), ou *leading jet*, avant la sélection. À droite, le spectre de masse invariante après la sélection effectuée à partir de la métrique d’anomalie.

Pour réduire notre bruit de fond, nous effectuons une sélection à partir de notre score d’anomalie. Cette sélection nous permettra de conserver les jets les plus anormaux, tels que prédits par le beta-VAE. Nous visons à conserver la plus grande quantité de signal, tout en rejetant la plus grande quantité de bruit de fond possible. Nous définissons le rejet du bruit de fond (ou *background rejection*) en fonction des valeurs de la métrique d’anomalie sur le bruit de fond de QCD. Par exemple, nous déterminons la valeur de la métrique à laquelle effectuer notre sélection et qui nous permet de conserver seulement 1% de notre bruit de fond. Ensuite, nous appliquons cette sélection sur notre ensemble test complet incluant notre signal et nous observons l’effet de cette sélection sur notre spectre de masse invariante. Si le réseau a appris à bien distinguer les événements du modèle standard des événements anormaux, le bruit de fond sera diminué et le signal

⁴Un exemple de ce type de recherche peut être trouvé dans [64] ou, pour une description vulgarisée, dans le chapitre 4 traitant de la découverte du boson de Higgs du livre de Pauline Gagnon [80].

sera détectable. Nous espérons en effet diminuer au maximum notre bruit de fond et faire apparaître notre signal qui, auparavant, était trop faible pour être détecté. Pour obtenir un résultat concluant, la courbe devrait suivre un comportement similaire à celui décrit à la Figure 37, où le signal de nouvelle physique devient apparent après la sélection.

Nous devons également choisir la meilleure métrique (ou meilleur score) d'anomalie pour l'évaluation sur l'ensemble test. Nous explorons les différents choix suivant :

- L'erreur quadratique moyenne (MSE) entre la prédiction et l'entrée (équation 20).
- La divergence Kullback-Leibler (KLD) dans l'espace latent (équation 22) et dans l'espace de reconstruction (prédiction - entrée).
- La distance de Wasserstein (EMD) entre la prédiction et l'entrée (équations 16).
- La distance de Jensen-Shannon (JSD) entre la prédiction et l'entrée, qui mesure la différence entre deux distributions de probabilités. Cette métrique est symétrique et définie à partir de la divergence de Kullback-Leibler D tel que :

$$\text{JSD}(P|Q) = \sqrt{\frac{D(P|M) + D(Q|M)}{2}}, \text{ où } M = (P + Q)/2 \quad . \quad (25)$$

Nous calculons les scores d'anomalie précédents lors de l'évaluation de notre modèle sur l'ensemble test de jets QCD et de quark top boostés. Nous affichons également les courbes ROC et l'erreur de reconstruction pour étudier la performance générale de l'auto-encodeur variationnel.

Enfin, comme il est fort probable que le dos d'âne dans le spectre de masse invariante soit trop petit pour être visible à l'oeil nu, nous évaluerons la performance de nos auto-encodeurs variationnels en utilisant l'efficacité du signal et la signification statistique (*significance*) en fonction de différentes valeurs de rejet du bruit de fond. L'efficacité sera définie comme le taux de signal restant après la sélection, et la signification statistique comme le ratio entre le nombre de signal et la racine carré du nombre de bruit de fond dans la région de masse du signal du quark top, que nous définissons comme (140,200) GeV/ c^2 . Nous choisissons cette fenêtre pour englober à la fois le pic du quark top et ses bandes latérales.

4.3. Résultats pour l’auto-encodeur variationnel

Nous commençons par afficher les courbes ROC et de gain de signal sur l’ensemble test pour les différentes métriques présentées. Les courbes ROC de la Figure 38 de gauche affichent la performance attendue du beta-VAE dans 36 fb^{-1} pour les divers scores d’anomalie. Les quatre métriques présentent un score AUC similaire, mais nous remarquons que le score AUC maximal atteint par le beta-VAE est de 0.760 avec la métrique du MSE pour l’architecture choisie. La Figure de droite montre le gain en signal selon son efficacité. À de basses efficacités de signal, le MSE est encore la métrique qui permet d’obtenir la meilleure performance et le meilleur gain. Nous remarquons toutefois que nos courbes ROC se trouvent beaucoup plus loin que celles obtenues par le TopoDNN à la Figure 35. En comparaison, nous obtenons un score AUC de 0.921 en utilisant les scores retournés sur notre ensemble test par le TopoDNN, soit le cas optimal obtenu par un réseau de neurones supervisé dont la tâche est d’identifier les quarks top.

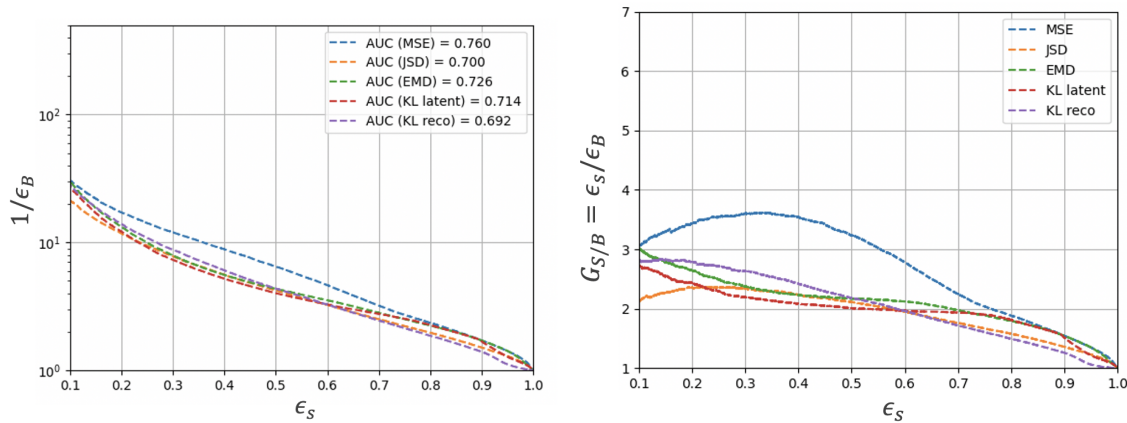


Figure 38. Courbes ROC (gauche) et de gain de signal (droite) pour le beta-VAE sur l’ensemble test. La Figure de gauche affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l’efficacité du signal des quark top boostés (ϵ_S). La Figure de droite affiche le gain en signal (ϵ_S/ϵ_B) en fonction de l’efficacité du signal (ϵ_S).

Nous observons ensuite l’allure des différents scores d’anomalie pour le bruit de fond et le signal à la Figure 39. Dans les quatre cas présentés, les jets provenant du quark top résultent en une erreur de reconstruction plus élevée. Ceci nous indique que le beta-VAE a, tel qu’espéré, plus de difficulté à reproduire les jets de signaux, que nous ne lui avons pas fourni à l’entraînement. À première vue, ces résultats sont prometteurs, puisque nous pouvons espérer conserver plus de signaux que de bruit de fond en sélectionnant les jets avec des valeurs plus élevées de score d’anomalie. Cependant, nous devons mentionner que le MSE et la KLD dans l’espace latent ont des allures particulières. En effet, nous

nous attendons à voir des distributions lisses comme celles du JSD et de la KLD dans l'espace de reconstruction, mais nous faisons plutôt face à un phénomène de double bosse. Ce phénomène est surtout présent pour le bruit de fond provenant des dijets QCD. Après une investigation des propriétés des jets qui peuplent la deuxième bosse (masse, impulsion, nombre de constituants), nous ne pouvons conclure clairement sur les causes de ce phénomène. Nous soupçonnons que, pour les valeurs élevées de MSE (KLD), nous retenons les jets QCD à plus basse impulsion transverse provenant des échantillons de JZ3W à JZ5W. Comme ces événements possèdent des poids d'événement par événement plus élevés, la distribution migre vers la droite.

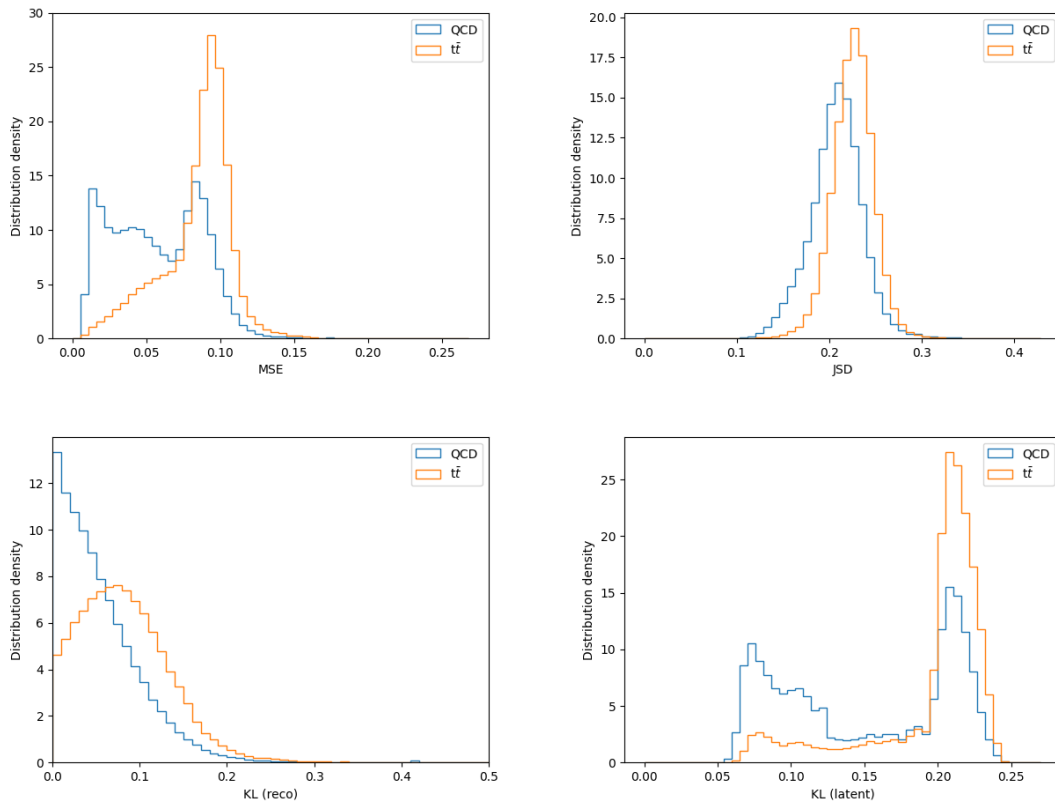


Figure 39. Allure des différentes erreur de reconstruction pour le bruit de fond (bleu) et le signal (orange). Les histogrammes correspondent aux distributions attendues dans 36 fb^{-1} de données.

À partir de la métrique d'anomalie, nous effectuons la procédure présentée à la section 4.2.2. Nous choisissons d'observer l'effet sur le spectre de masse invariante pour des rejets de bruit de fond de 10, 100, 10^3 et 10^4 , correspondant à des efficacités de bruit de fond de 10%, 1%, 0.1% et 0.01%. Pour l'instant, nous effectuons ces sélections en utilisant les métriques d'anomalie de MSE, de JSD et de KLD dans l'espace latent, mais les autres métriques

donnent des résultats similaires. Nous commençons par observer l'effet des sélections sur la masse des jets de bruit de fond de QCD à la Figure 40. Un problème important est déjà apparent. Ici, les métriques semblent fortement corrélées avec la masse des jets. Comme nous pouvons le voir, les sélections plus sévères, c'est-à-dire les plus grands rejets de bruit de fond de fond, sélectionnent les événements de QCD avec une haute masse, ce qui déplace la distribution de masse vers la droite. Ainsi, notre signal, qui se trouve à haute masse, risque d'être toujours noyé par notre bruit de fond. Cette corrélation est aussi beaucoup plus prononcée pour le MSE et le KLD.

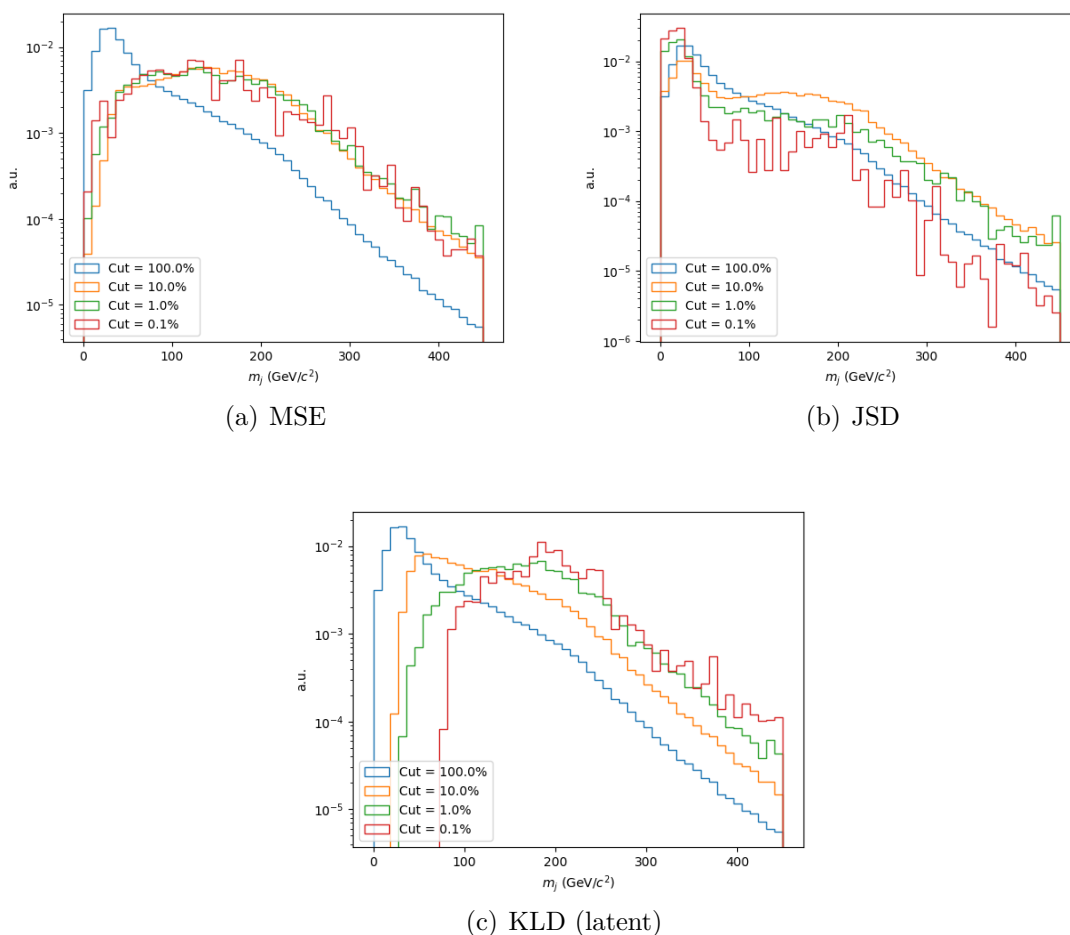


Figure 40. Effet des sélections sur le spectre de masse invariante du premier jet du bruit de fond (m_j) dans 36 fb^{-1} pour le beta-VAE. Les distributions sont normalisées.

Cette corrélation avec la masse des jets est aussi visible à la Figure 41. Dans les trois Figures, nous voyons qu'à mesure que nous retenons les jets avec un score d'anomalie plus élevé, nous retenons les jets de bruit de fond avec une plus grande masse en moyenne.

Cette corrélation est d'autant plus importante pour le KLD (figure du bas), qui atteint une valeur de masse moyenne de près de 1700 GeV/c², ce qui correspond à l'effet observé sur le spectre de masse à la Figure 40 c. Cette corrélation pourrait être problématique. En effet, dans les « recherches de dos d'âne », nous utilisons l'information en lien avec la masse pour estimer et réduire le bruit de fond. Si notre bruit de fond ne diminue pas de façon constante en fonction de la masse à mesure que nous effectuons des sélections plus sévères, tel qu'à la Figure 37, nous nous retrouverons possiblement avec un problème pour notre recherche de résonances et l'estimation du bruit de fond. Si la corrélation entre la masse des jets et le score d'anomalie est très élevée, le réseau pourrait éliminer uniquement les jets à basse masse. Notre réseau serait alors inutile pour chercher des résonances à haute masse, puisqu'il ne couperait aucun des jets du bruit de fond de QCD à haute masse. Aussi, si la corrélation est élevée, le spectre de masse des jets QCD, qui était à l'origine lisse, pourrait être sculpté et rendre plus difficile l'observation d'un dos d'âne au-dessus du bruit de fond.

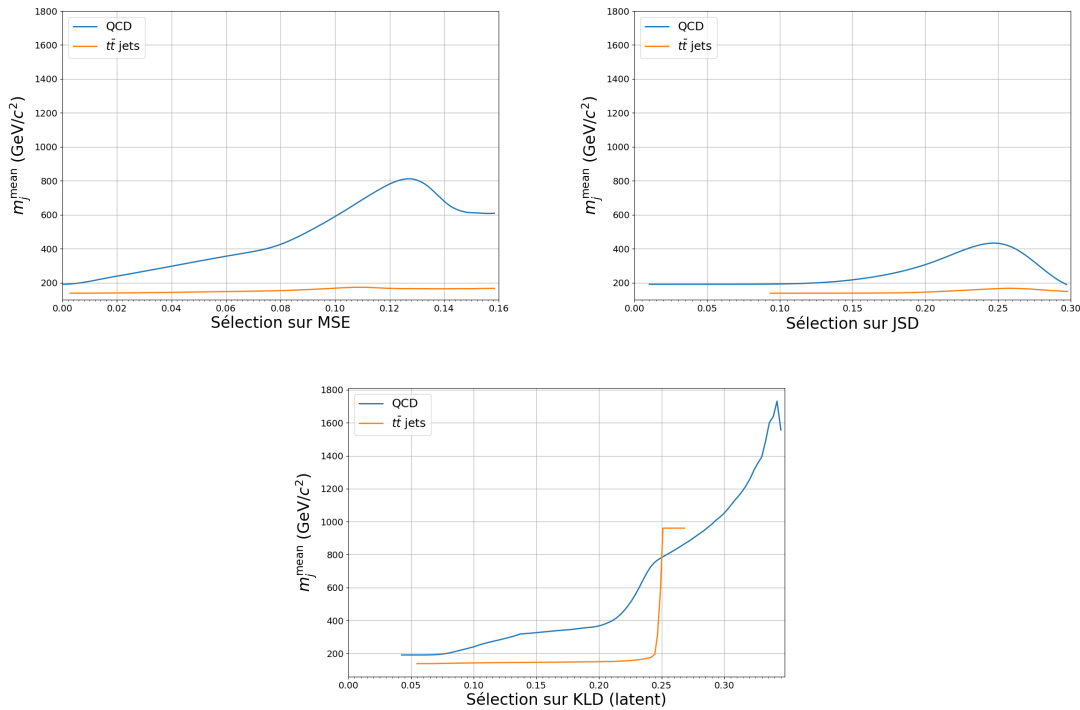


Figure 41. Corrélation entre les métriques d'anomalie de MSE, JSD et KLD (latent) et la masse des jets. La Figure affiche la masse moyenne des jets (m_j^{mean}) en fonction de la sélection effectuée sur le score d'anomalie.

Toutefois, comme nous l'avons mentionné, il est fort probable que le « dos d'âne » ne soit pas visible à l'oeil nu. Voyons alors ce que les chiffres nous disent. Les sélections sur la métrique d'anomalie à partir du beta-VAE nous permettent-elles d'obtenir une signification statistique

de 3 ou 5 σ dans notre ensemble de test incluant les signaux? Le Tableau 4 présente les résultats en fonction de l’efficacité du bruit de fond pour la métrique de MSE. Malheureusement, à mesure que nous effectuons des sélections plus sévères, nous éliminons trop de signal par rapport au bruit de fond. Effectivement, même si la signification statistique était déjà élevée sans sélections, nous espérons augmenter celle-ci avec notre réseau. Pour nous permettre de purifier notre signal et d’obtenir 3 ou 5 σ , nous devons éliminer plus de dijets QCD.

Modèle	beta-VAE			
	ϵ_B (%)	ϵ_S (%)	S	B
100	100	20 109	4 245 466	9.756
10	36.4	9818	1 753 795	7.414
1	2.98	749	148 852	1.941
0.1	0.31	81	15 165	0.658
0.01	0.031	8	1559	0.203

Tableau 4. Signification statistique pour le beta-VAE en fonction de l’efficacité du bruit de fond (MSE). L’efficacité du signal/bruit de fond (ϵ_S/ϵ_B) correspond au taux de signal/bruit de fond restant après la sélection. Le nombre de signal (S) ou de bruit de fond (B) correspond au nombre de jets restant dans la région de masse d’intérêt de (140,200) GeV/ c^2 . Les chiffres correspondent au nombre de jets attendus dans 36 fb $^{-1}$ dans les vraies données.

Les résultats obtenus aux Tableaux 5 et 6 pour les métriques de JSD et du KLD de l’espace latent ne sont pas plus optimistes. Dans le cas du JSD, même si la corrélation de masse est moins prononcée, les sélections effectuées éliminent plus de signaux que pour le MSE. Dans le cas du KLD, nous conservons plutôt des événements de bruit de fond supplémentaires par rapport au MSE.

Modèle	beta-VAE			
	ϵ_B (%)	ϵ_S (%)	S	B
100	100	20 109	4 245 466	9.756
10	25.1	6462	1 195 604	5.910
1	1.13	254	47 304	1.167
0.1	0.059	12	2061	0.264
0.01	0.0052	2	12	0.577

Tableau 5. Signification statistique pour le beta-VAE en fonction de l’efficacité du bruit de fond (JSD). Les chiffres correspondent au nombre de jets attendus dans 36 fb $^{-1}$ dans les vraies données.

Modèle	beta-VAE			
	ϵ_B (%)	ϵ_S (%)	S	B
100	100	20 109	4 245 466	9.756
10	25.6	6258	1 271 318	5.550
1	3.01	764	217 689	1.637
0.1	2.53	62	30 311	0.356
0.01	0.02	5	2592	0.098

Tableau 6. Signification statistique pour le beta-VAE en fonction de l’efficacité du bruit de fond (KLD latent). Les chiffres correspondent au nombre de jets attendus dans 36 fb^{-1} dans les vraies données.

Ainsi, il devient important de trouver une façon d’éliminer ou de réduire la corrélation entre la masse des jets et nos métriques d’anomalie, et d’améliorer notre rejet de bruit de fond sans en faire souffrir notre signal.

4.4. Auto-encodeur variationnel avec exposition aux valeurs aberrantes

Dans la section précédente, nous avons rencontré un obstacle important à notre tentative de « redécouverte » du quark top. Cet obstacle nous oblige à nous arrêter et à réfléchir : comment pouvons-nous forcer notre réseau à discriminer les jets selon leurs caractéristiques intrinsèques et non leur masse ? Ceci nous amène à notre prochaine méthode, soit l’exposition aux valeurs aberrantes à l’entraînement. Nous décrirons cette méthode dans ce qui suit et nous découvrirons si nos résultats en bénéficient.

4.4.1. Méthode d’exposition aux valeurs aberrantes

Pour tenter de dé-corréler nos métriques d’anomalie de la masse des jets, nous utilisons une méthode d’exposition aux valeurs aberrantes, de l’anglais *Outlier Exposure*. L’objectif de cette méthode de semi-supervision est d’injecter des événements aberrants, ou hors distribution, lors de l’entraînement pour acclimater le réseau. Nous nommons ces événements des valeurs aberrantes, ou *outliers*. Ainsi, à l’entraînement, nous espérons que ce type d’auto-encodeur variationnel, que nous appellerons maintenant OE-VAE, s’entraîne à reconnaître des événements qui sortent de la distribution normale des dijets QCD.⁵

⁵Pour plus de détails sur les origines et les motivations de cette méthode d’exposition, voir les travaux de [96].

Nous choisissons des jets boostés issus de bosons W comme valeurs aberrantes. Le boson W est issu d’une résonance massive et se désintègre généralement en deux jets. Ainsi, à cause de la masse élevée de la résonance, le boson W sera boosté et les deux jets produits sont reconstruits dans un seul même grand jet. Ce type de jet boosté a donc une signature typique à deux sous-jets 2(-*prong*), et diffère de la signature typique des dijets QCD 1(-*prong*), mais aussi de celle de notre signal de quark top boosté 3(-*prong*).

Puisque nous souhaitons que les résultats prédits par le réseau soient indépendants de la masse des échantillons, nous ajustons la masse de ces jets aberrants afin qu’elle corresponde, en moyenne, à la distribution de masse des dijets QCD. De ce fait, nous souhaitons que notre OE-VAE se concentre sur la sous-structure des jets et non sur leur masse pour définir ce en quoi constitue un jet anormal.

Nous notons que les échantillons de W furent générés dans la référence [58] avec **Delphes** selon les mêmes échantillons qu’à la section 3.2. Ces jets de W furent simulés à partir d’interactions de type $pp \rightarrow W' \rightarrow W(jj)Z(\nu\nu)$. Les valeurs aberrantes ne sont donc pas exactement produites comme nos simulations d’échantillons ATLAS. Par contre, un entraînement du OE-VAE effectué sur des échantillons **Delphes** nous a permis d’obtenir une bonne performance lors de l’évaluation sur des échantillons simulés d’ATLAS. Ceci nous donne espoir qu’il est possible d’utiliser des échantillons de valeurs aberrantes produites par **Delphes** pour ce test préliminaire.

Concrètement, l’injection des valeurs aberrantes à l’entraînement requiert l’ajout d’un deuxième terme dans la fonction de perte. Celle-ci devient :

$$L = L_{\text{vae}} - \lambda L_{\text{OE}}, \text{ où } L_{\text{vae}} = L_{\text{reco}} + \beta L_{\text{KL}} \quad . \quad (26)$$

Ici, nous contrôlons le paramètre λ manuellement. Comme la perte L_{OE} est soustraite à la perte totale, elle se comporte comme une pénalité. Une valeur de λ plus grande signifie donc que le OE-VAE se concentre plus sur la tâche de supervision et de dé-corrélation de masse que sur la tâche de reconstruction des jets.

Nous distinguons deux types de perte L_{OE} , présentées aux équations 27. Selon celle que nous utilisons, nous distinguons deux types de OE-VAE : le KL-OE-VAE et le MSE-OE-VAE. Dans ces équations, σ correspond à la fonction d’activation sigmoïde et *margin* est

une valeur posée à la main qui nous permet de forcer les valeurs aberrantes à avoir une perte plus élevée qu’un certain seuil. Également, z' correspond à la représentation latente des valeurs aberrantes et z à celle des événements QCD. Ainsi, à l’évaluation du modèle, nous utilisons la divergence Kullback-Leibler dans l’espace latent comme score d’anomalie pour le KL-OE-VAE, et le MSE pour le MSE-OE-VAE, puisque la dé-corrélation à l’entraînement s’effectue à partir de ces métriques.

$$L_{\text{OE}} = \sigma(\text{MSE}(W) - \text{MSE}(\text{QCD})) \quad , \quad (27\text{a})$$

$$L_{\text{OE}} = \min \{0, L_{\text{KL},W}(z') - L_{\text{KL},\text{QCD}}(z) - \text{margin}\} \quad ; \quad (27\text{b})$$

Nous transformons les événements de la même façon que pour le beta-VAE et nous utilisons le même nombre d’événements QCD et top pour l’entraînement et l’évaluation du réseau que précédemment. Nous utilisons un total de 120 000 jets simulés de W lors de l’entraînement pour l’exposition aux valeurs aberrantes.

Ensuite, nous fixons le paramètre λ à 500 pour le MSE-OE-VAE et à 2 pour le KL-OE-VAE. Nous utilisons également une valeur de *margin* de 1. Ces paramètres ont été optimisés à la main dans l’étude [58]. De plus, nous faisons évoluer le paramètre λ de façon cyclique à l’entraînement pour essayer de trouver un juste milieu entre les pertes du beta-VAE et du OE. Cette procédure se nomme *annealing* et consiste à faire osciller λ entre 0 et sa valeur définie manuellement, selon l’époque.

Dans ce qui suit, nous présenterons enfin les résultats pour le KL-OE-VAE et pour le MSE-OE-VAE, et nous comparerons ces derniers aux résultats obtenus pour le beta-VAE.

4.4.2. Résultats pour l’auto-encodeur variationnel avec exposition aux valeurs aberrantes

Nous commençons à nouveau par afficher les courbes ROC sur l’ensemble test pour les différents scores d’anomalie avec l’application de la méthode d’exposition. Les courbes ROC de la Figure 42 affichent la performance attendue du KL-OE-VAE dans 36 fb^{-1} pour les divers scores d’anomalie. Les scores AUC obtenus par le nouveau réseau n’ont que légèrement varié par rapport aux performances du beta-VAE (lignes pointillées). Nous voyons que le KL-OE-VAE (ligne brune) permet d’obtenir un score de 0.706, en comparaison

au 0.714 obtenu par le beta-VAE avec le KLD dans l'espace latent (ligne rouge). Pour le MSE-OE-VAE (ligne rose), nous obtenons un score de 0.721, en comparaison au 0.760 du beta-VAE (ligne bleue). Ainsi, les scores sont bel et bien similaires, bien que les OE-VAE présentent une légère baisse de performance. Comme l'ajout du paramètre λ à l'entraînement nous permet de favoriser la régularisation de l'espace latent à la reconstruction des jets, cette baisse de performance était à prévoir. Aussi, la métrique de MSE demeure la métrique qui nous permet d'obtenir le meilleur gain en signal lorsque nous comparons les courbes de gain du KL-OE-VAE et du MSE-OE-VAE. Par contre, ces gains ne sont toujours pas comparables à ceux obtenus par le beta-VAE avec le MSE comme score d'anomalie.

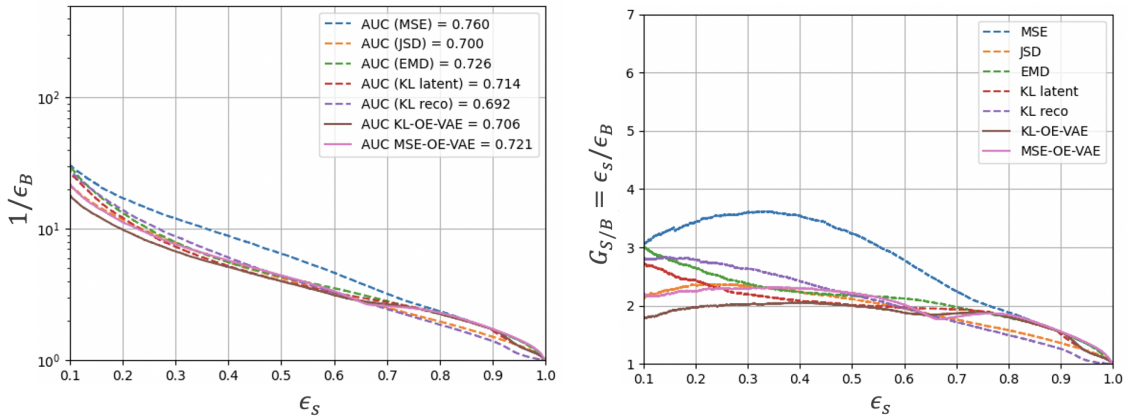


Figure 42. Courbes ROC (gauche) et de gain de signal (droite) pour le KL-OE-VAE sur l'ensemble test. La Figure de gauche affiche le rejet du bruit de fond ($1/\epsilon_B$) en fonction de l'efficacité du signal des quark top boostés (ϵ_S). La Figure de droite le gain en signal (ϵ_S/ϵ_B) en fonction de l'efficacité du signal des quark top boostés (ϵ_S).

Nous comparons ensuite l'allure des scores de MSE et de KLD (latent) pour le bruit de fond et le signal à la Figure 43. Les distributions pour le signal présentent toujours une erreur de reconstruction généralement plus élevée, signe que les OE-VAE ont, tel qu'attendu, de la difficulté à reconstruire ces jets considérés anormaux. Cependant, pour les deux métriques, le phénomène de double bosse est très prononcé. Nous voyons une séparation claire entre les jets qui ont une faible erreur de reconstruction, qui sont majoritairement les jets QCD, et ceux qui ont une erreur plus élevée. Ces derniers voient des contributions à la fois des jets de QCD et des jets de quark top, bien que ceux-ci sont présent en plus grand nombre. Le comportement du OE-VAE et des métriques est cependant fortement influencé par le choix des paramètres β et λ . Par exemple, un entraînement avec le KL-OE-VAE en fixant $\beta = 0.01$ nous permettait d'obtenir un KLD avec une double bosse légèrement moins prononcée. Le comportement du OE-VAE et des métriques semble très sensible aux choix des hyperparamètres, et une petite variation peut faire varier grandement la capacité de reconstruction du réseau. La tâche d'ajustement des hyperparamètres avec notre jeu

de données devient importante si nous souhaitons effectuer un bon compromis entre la reconstruction et la régularisation de l'espace latent, et une optimisation par balayage plus rigoureuse sera de mise dans le futur.

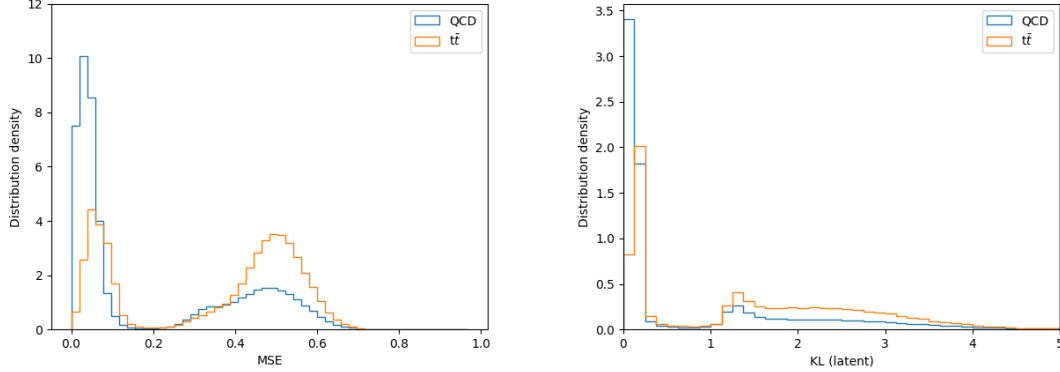


Figure 43. Allure des erreurs de reconstruction du bruit de fond (bleu) et le signal (orange) pour le MSE-OE-VAE (gauche) et le KL-OE-VAE (droite).

Ensuite, nous effectuons notre sélection des jets les plus anormaux et nous observons l'effet sur le spectre de masse invariante pour des rejets de bruit de fond de 10, 100, 10^3 et 10^4 . À la Figure 45, nous pouvons voir l'effet des sélections sur le bruit de fond de QCD pour le MSE-OE-VAE, et à la Figure 44 pour le KL-OE-VAE. Les résultats pour le MSE-OE-VAE semblent prometteurs. En effet, nous voyons un léger aplatissement du spectre de masse du bruit de fond à la Figure 44 b), signe qu'une certaine dé-corrélation est atteinte. Nous voyons également une grande amélioration pour la KL-OE-VAE à la Figure 45 b) en comparaison aux résultats du beta-VAE. À partir d'une efficacité de bruit de fond de 10%, nous voyons que la distribution de masse reste relativement stable. Le KL-OE-VAE semble donc dé-corréler la masse jusqu'à un certain niveau. Nous pouvons ainsi espérer qu'une recherche de dos d'âne réalisée sur ce spectre de masse puisse réaliser une meilleure estimation de bruit de fond, et possiblement trouver notre signal. À la Figure 46, nous voyons qu'une corrélation de masse en fonction des sélections est toujours présente pour les deux OE-VAE. Par contre, cette corrélation a effectivement diminué. Pour le MSE-OE-VAE, nous atteignons une masse moyenne maximale d'environ $400 \text{ GeV}/c^2$, en comparaison à $800 \text{ GeV}/c^2$ pour le beta-VAE avec le MSE à la Figure 41. Pour le KL-OE-VAE, nous obtenons une valeur maximale de $350 \text{ GeV}/c^2$, en comparaison à $1700 \text{ GeV}/c^2$ pour le beta-VAE avec le KLD (latent). La corrélation a ainsi grandement diminuée pour le KL-OE-VAE. Il semblerait donc à première vue que les modèles d'auto-encodeur variationnel avec exposition sont efficaces pour dé-corréler la masse des jets des scores d'anomalies et qu'ils pourraient

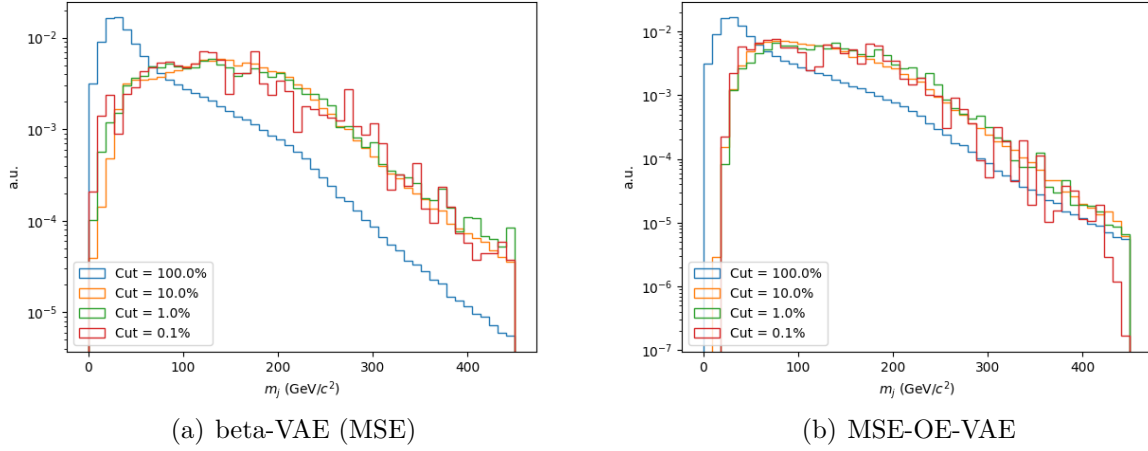


Figure 44. Effet des sélections sur le spectre de masse invariante du premier jet (m_j) du bruit de fond dans 36 fb^{-1} pour le beta-VAE (gauche) et le MSE-OE-VAE (droite). Les distributions sont normalisées.

être utilisés dans une perspective de recherche de bos d'âne.

Les modèles de OE-VAE nous permettent d'atteindre une dé-corrélation de la masse satisfaisante. Mais ces résultats nous permettent-ils d'améliorer notre signification statistique? Nous pouvons enfin comparer les significations statistiques obtenues avec le KL-OE-VAE et le MSE-OE-VAE à celles obtenues par le beta-VAE. Les Tableaux 7 et 8 présentent les résultats en fonction de l'efficacité du bruit de fond pour les deux réseaux.

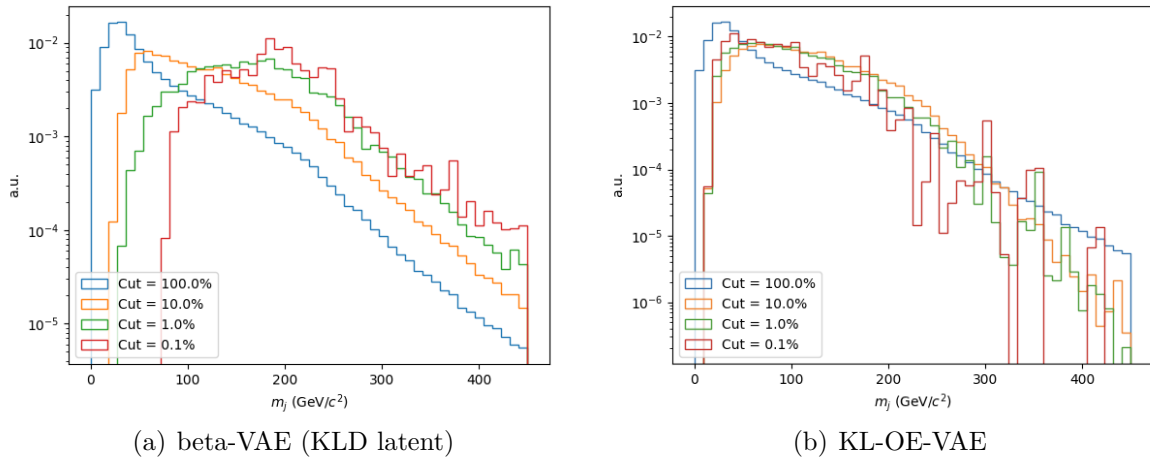


Figure 45. Effet des sélections sur le spectre de masse invariante du premier jet (m_j) du bruit de fond dans 36 fb^{-1} pour le beta-VAE (gauche) et le KL-OE-VAE (droite). Les distributions sont normalisées.

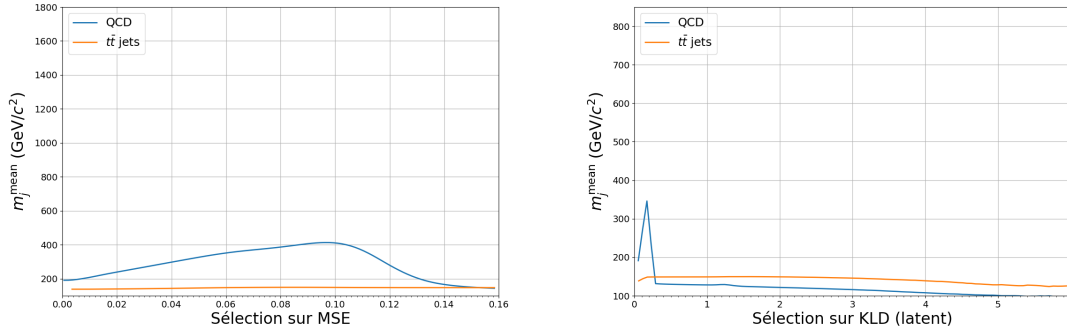


Figure 46. Corrélation entre les métriques d’anomalie de MSE et KLD (latent) et la masses des jets pour le MSE-OE-VAE (gauche) et le KL-OE-VAE (droite). La Figure affiche la masse moyenne des jets (m_j^{mean}) en fonction de la sélection effectuée sur la métrique d’anomalie.

Malheureusement, même si nous parvenons à dé-corréler la masse des jets de nos scores d’anomalie, nos sélections les plus sévères éliminent trop de signal par rapport au bruit de fond. En effet, nous obtenons une signification statistique de 0.173 (KL-OE-VAE) et 0.178 (MSE-OE-VAE) pour une efficacité de bruit de fond de 0.1%, en comparaison à 0.356 (KLD) et 0.658 (MSE) pour le beta-VAE. Il semble que l’allure à double bosse des métriques de KLD et de MSE pour les modèles de OE-VAE nuise à la performance en terme de signification statistique. Comme une certaine population de jets QCD se trouvent dans la deuxième bosse, les sélections des jets anormaux conservent une plus grande quantité de jets de bruit de fond. La performance pour les OE-VAE ne nous permet donc pas d’obtenir de meilleure signification statistique avec nos modèles actuels. Bref, même si une certaine dé-corrélation de la masse est atteinte, nos résultats souffrent d’une baisse de performance.

Modèle	KL-OE-VAE			
ϵ_B (%)	ϵ_S (%)	S	B	$\sigma = S/\sqrt{B}$
100	100	20 109	4 245 466	9.756
10	19.45	4457	1 181 436	4.100
1	1.24	222	101 931	0.695
0.1	0.10	16	8535	0.173
0.01	0.01	2	403	0.099

Tableau 7. Signification statistique pour le KL-OE-VAE en fonction de l’efficacité du bruit de fond. Les chiffres correspondent au nombre de jets attendus dans 36 fb^{-1} dans les vraies données.

Finalement, nous pouvons comparer nos performances avec le TopoDNN de l’expérience ATLAS, soit le cas optimal supervisé. Nous effectuons les mêmes sélections sur nos jets, cette

Modèle	MSE-OE-VAE			
	ϵ_B (%)	ϵ_S (%)	S	B
100	100	20 109	4 245 466	9.756
10	23.04	5576	1 351 997	4.796
1	1.80	397	163 656	0.981
0.1	0.13	22	15 315	0.178
0.01	0.01	2	1714	0.048

Tableau 8. Signification statistique pour le MSE-OE-VAE en fonction de l'efficacité du bruit de fond. Les chiffres correspondent au nombre de jets attendus dans 36 fb^{-1} dans les vraies données.

fois-ci à partir du score retourné par le TopoDNN. Ce score, situé entre (0,1), correspond à la probabilité qu'un jet soit issu d'un quark top. En définissant les mêmes sélections pour des rejets de bruit de fond de 10, 100, 10^3 et 10^4 , nous comparons les significations statistiques obtenues par le TopoDNN avec nos divers réseaux de neurones au Tableau 9. Ainsi, nous pouvons voir que nos algorithmes non-supervisé (beta-VAE) et semi-supervisé (ici MSE-OE-VAE) performant beaucoup moins bien que l'algorithme supervisé du TopoDNN, et que les sélections effectuées à partir du TopoDNN nous permettent de conserver beaucoup plus de signaux. Bien sûr, ce dernier est conçu pour retrouver les quarks top, mais cette comparaison nous donne une idée de comment nous nous approchons, ou dans notre cas nous éloignons, du cas optimal. Nous devons effectuer plusieurs améliorations à nos réseaux dans le but de s'approcher autant que possible de la performance du TopoDNN.

Modèle	beta-VAE (MSE)		MSE-OE-VAE		TopoDNN		
	ϵ_B (%)	ϵ_S (%)	$\sigma = S/\sqrt{B}$	ϵ_S (%)	$\sigma = S/\sqrt{B}$	ϵ_S (%)	$\sigma = S/\sqrt{B}$
100	100	100	9.756	100	9.756	100	9.756
10	36.4	36.4	7.414	23.04	4.796	78.2	15.292
1	2.98	2.98	1.941	1.80	0.981	36.9	21.161
0.1	0.31	0.31	0.658	0.13	0.178	10.4	19.430
0.01	0.031	0.031	0.203	0.01	0.048	1.2	7.999

Tableau 9. Comparaison des significations statistiques entre le beta-VAE, le KL-OE-VAE et le TopoDNN en fonction de l'efficacité du bruit de fond. Les chiffres correspondent au nombre de jets attendus dans 36 fb^{-1} dans les vraies données.

4.4.3. Améliorations futures

Il est indéniable que nous devons effectuer du travail au niveau de l'architecture, du choix d'hyperparamètres et des données fournies au réseau pour améliorer nos résultats dans le futur. Pour ce faire, nous envisageons de nombreuses avenues. La première serait d'intégrer des variables de haut niveau à l'entraînement, de façon comparable au réseau de neurones profond (DNN) et à l'arbre de décision (BDT) présentés dans [29]. Par exemple, les variables liées à la sous-structure, telles que τ_{21} et τ_{32} , pourraient améliorer la performance et fournir de l'information supplémentaire au réseau. Nous espérons que ces variables puissent améliorer la discrimination du réseau entre les jets de QCD et les signaux.

Aussi, nous considérons utiliser l'information des traces et des *particle flow objects* plutôt que des *topoclusters*. Ces objets représentent généralement mieux les particules individuelles qui forment les jets. Même si les *topoclusters* nous fournissent de l'information sur la sous-structure des jets, ces derniers ne sont pas parfaits, et les traces et les *particle flow objects* pourraient augmenter notre discrimination entre le bruit de fond et le signal. En effet, ces derniers procurent généralement une meilleure résolution spatiale et résolution en énergie des constituants des jets et pourraient permettre au réseau de mieux sonder la sous-structure des jets.

Une autre possibilité envisagée est de combiner les informations des quadri-vecteurs avec les images de dépôts d'énergie des jets dans différentes couches du détecteur. Pour ce faire, nous considérons transformer les couches d'entrées de notre VAE en un réseau neuronal convolutif (CNN). Les constituants des jets seraient alors traités comme des images plutôt qu'en quadri-vecteurs.

Enfin, des modifications sur le plan du nombre de couches, du nombre de neurones et des autres hyperparamètres sont à envisager. Nos premiers tests nous laissent croire que la tâche de « redécouverte » du quark top dans les jets boostés sera difficile, et que toute optimisation possible sera nécessaire pour mettre les chances de notre côté et prouver que notre réseau pourrait être un outil utile dans la recherche de nouvelle physique dans l'expérience ATLAS.

4.4.4. Conclusion

Nos résultats présentés dans le Chapitre 4 représentent un premier pas vers l'élaboration d'une nouvelle méthode de recherche de physique au-delà du modèle standard indépendante des modèles théoriques. À notre connaissance, nous sommes les premiers et premières à tenter une recherche aussi ambitieuse, qui ne s'intéresse pas à un modèle de nouvelle physique en particulier et qui s'effectue dans l'ensemble des jets boostés du détecteur ATLAS. Nous avons pu établir une base à notre projet et nous avons maintenant une idée concrète d'à quel point nous sommes près (ou loin) de notre objectif. Nous avons également proposé des pistes d'amélioration possibles. Même si nos résultats ne sont pas concluants, toute recherche en physique se fonde sur des essais, des erreurs et de multiples itérations. Nos travaux pourront certainement être une fondation sur laquelle les prochains et prochaines contributeurs et contributrices ajouterons de multiples briques.

Conclusion

Dans ce mémoire, nous avons exploré l'apprentissage machine comme avenue possible pour la détection de nouvelle physique dans les jets boostés de l'expérience ATLAS. Nous avons amené une nouvelle approche indépendante des modèles théoriques en développant un réseau de neurones entraîné sur des jets boostés du modèle standard, dans le but d'identifier des événements anormaux dans les données du LHC. Au premier chapitre, nous avons présenté le modèle standard de la physique des particules et nous avons illustré ses nombreuses limites. Nous avons vu que les recherches réalisées au LHC et à l'expérience ATLAS se concentrant sur des modèles précis de nouvelle physique peinent à porter fruit, bien qu'il y ait eu des développements intéressants dans les expériences LHCb et Muon $g-2$ au moment de soumettre ce mémoire [2, 112]. Ceci a motivé notre projet de recherche indépendante des modèles théoriques. Ensuite, au deuxième chapitre, nous avons décrit l'appareillage expérimental de l'expérience ATLAS et nous avons expliqué comment les jets se produisent et sont enregistrés dans le détecteur. Puis, nous avons poursuivi au troisième chapitre avec les bases des réseaux de neurones et la description des jets boostés, ce type de jets qui pourrait dissimuler des particules et résonances candidates à la nouvelle physique et qui sont l'objet de notre recherche. Nous avons déterminé notre terrain d'entraînement, soit les jets issus de la QCD (gluons et quarks u, d, c, s, b), et ciblé un signal à retrouver dans nos données, les jets issus du quark top, dans le but de tester la faisabilité de notre approche. Finalement, au dernier chapitre, nous avons testé notre réseau de neurones dans des données de simulation ATLAS. Nous avons tenté de retrouver notre signal en effectuant des sélections sur les scores d'anomalie retournés par le réseau. Un problème de corrélation entre les scores et la masse des jets nous a amené à développer une technique de semi-supervision pour optimiser nos chances de réussite dans une optique de chasse aux dos d'ânes.

Notre analyse ne nous a pas permis d'atteindre une signification statistique de 3σ . Les quark top boostés représentant seulement $O(1\%)$ des quarks top produits au LHC, notre analyse s'avérait dès le départ très difficile. Comme notre réseau ne pouvait purifier notre signal à mesure que nous effectuions des sélections plus sévères, des améliorations sont nécessaires pour obtenir une meilleure discrimination entre les événements du modèle

standard et de futurs événements anormaux. Cette nouvelle technique de détection d'anomalies, si prouvée performante, pourrait toutefois être un outil précieux pour la recherche de nouvelle physique dans les jets boostés du détecteur ATLAS.

Enfin, mes connaissances en physique acquises au cours de ma maîtrise et de mon baccalauréat sont inséparables de mon expérience au sein du département et de mes connaissances acquises en ÉDI. Ainsi, à travers ce mémoire, nous avons posé des réflexions sur des enjeux environnementaux, d'équité, de diversité et d'inclusion en physique. D'abord, nous avons constaté que les domaines de la physique et de l'intelligence artificielle sont de nos jours encore dominés par les hommes et sont très peu diversifiés. Nous avons vu que l'exclusion des femmes et des personnes issues de la diversité était souvent synonyme d'un manque de sensibilisation aux enjeux d'ÉDI, ce qui instaurait un climat organisationnel non inclusif au sein des diverses organisations de recherche. En IA, ce climat résulte en l'enracinement de nos biais dans les algorithmes, qui causent ou accentuent les inégalités vécues par ces groupes sous-représentés. Ce manque de sensibilisation aux enjeux d'ÉDI dans les domaines des STIM fait en sorte que nous perdons de nombreuses personnes qui pourraient contribuer à la recherche. Par des efforts de sensibilisation et de discussion, nous visons ainsi à créer des milieux de recherche ouverts à tous et toutes. Heureusement, les communautés en physique et en IA changent. De plus en plus de personnes reconnaissent les liens entre notre recherche, l'environnement dans lequel elle s'effectue et nos pratiques en tant que scientifiques, et les conférences de sensibilisation sur le sujet sont de plus en plus communes dans les conférences internationales de STIM et les organisations, tel qu'au symposium sur la matière sombre *A Rainbow of Dark Sectors* (<https://indico.cern.ch/event/1019224/>) ou dans le cadre des activités d'IVADO (<https://ivado.ca/evenements/bias-and-discrimination-in-ai/>). Même si ces conversations sont parfois difficiles à avoir, les institutions démontrent de plus en plus d'écoute et de volonté, et c'est en continuant notre travail de discussion et de remise en question que nous atteindrons nos objectifs de recherche inclusive.

Références bibliographiques

- [1] ABDESSELAM, A., BELYAEV, A., KUUTMANN, E. B. *et al.* (2011). Boosted objects : a probe of beyond the standard model physics. *European Physical Journal C*, 71(1661). 10.1140/epjc/s10052-011-1661-y.
- [2] ABI, B., ALBAHRI, T., AL-KILANI, S. *et al.* (2021). Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. *Physical Review Letters*, 126(14). 10.1103/physrevlett.126.141801.
- [3] AGENCE FRANCE-PRESSE. (2018). Un scientifique émet des propos sexistes lors d'un atelier du CERN. *La Presse*. <https://www.lapresse.ca/international/europe/201810/01/01-5198654-un-scientifique-emet-des-propos-sexistes-lors-dun-atelier-du-cern.php>.
- [4] AGENCE FRANCE-PRESSE. (2019). Cern : un scientifique exclu pour sexisme. *Le Figaro*. <https://www.lefigaro.fr/flash-actu/2019/03/07/97001-20190307FILWWW00116-cern-un-scientifique-exclu-pour-sexisme.php?web=1&wdLOR=c4D7F76B4-1E2B-114C-A4F3-610338CE2BE1>.
- [5] AGOSTINI, M., BENATO, G., DELL'ORO, S. *et al.* (2021). Discovery probabilities of Majorana neutrinos based on cosmological data. *Physical Review D*, 103(3). 10.1103/PhysRevD.103.033008.
- [6] AGUILAR-SAAVEDRA, J. A., BENBRIK, R., HEINEMEYER, S. *et al.* (2013). Handbook of vector-like quarks : Mixing and single production. *Physical Review D*, 88(9). 10.1103/PhysRevD.88.094010.
- [7] ALEPH COLLABORATION. (2006). Precision electroweak measurements on the Z resonance. *Physics Reports*, 427(5). 10.1016/j.physrep.2005.12.006.
- [8] ALIC, M. (1982). The history of women in science : A women's studies course. *Women's Studies International Forum*, 5(1). 10.1016/0277-5395(82)90066-8.
- [9] ALICE COLLABORATION. (2008). The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08). <https://doi.org/10.1088/1748-0221/3/08/S08002>.
- [10] ALIOLI, S., NASON, P., OLEARI, C. *et al.* (2010). A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX. *Journal of High Energy Physics*, 2010(43). 10.1007/JHEP06(2010)043.

- [11] AMERICAN ASSOCIATION OF UNIVERSITY WOMEN. (s.d.a). The STEM Gap. <https://www.aauw.org/issues/education/stem/>.
- [12] AMERICAN ASSOCIATION OF UNIVERSITY WOMEN. (s.d.b). The STEM Gap : Women and Girls in Science, Technology, Engineering and Math. <https://www.aauw.org/resources/research/the-stem-gap/>.
- [13] APS PHYSICS. (s.d.). Bachelor’s Degrees in Physics and STEM Earned by Women. <https://www.aps.org/programs/education/statistics/womenstem.cfm>.
- [14] ARSENAULT, J., TALBOT, J., BOUSTANI, L. *et al.* (2019). The environmental footprint of academic and student mobility in a large research-oriented university. *Environmental Research Letters*, 14(9). 10.1088/1748-9326/ab33e6.
- [15] ABMANN, R., LAMONT, M. *et* MYERS, S. (2002). A Brief History of the LEP Collider. *Nuclear Physics B - Proceedings Supplements*, 109(2). 10.1016/S0920-5632(02)90005-8.
- [16] ATHERTON, T. J., BARTHELEMY, R. S., DECONINCK, W. *et al.* (2016). LGBT Climate in Physics : Building an Inclusive Community. Rapport technique, American Physical Society. <https://www.aps.org/programs/lgbt/upload/LGBTClimateinPhysicsReport.pdf>.
- [17] ATLAS COLLABORATION. (2008). The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08). <https://doi.org/10.1088/1748-0221/3/08/S08003>.
- [18] ATLAS COLLABORATION. (2012). Observation of a new particle in the search for the Standard Model Higgs Boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1). <https://doi.org/10.1016/j.physletb.2012.08.020>.
- [19] ATLAS COLLABORATION. (2013). Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2013(116). 10.1007/jhep01(2013)116.
- [20] ATLAS COLLABORATION. (2016). Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. Rapport technique, CERN, Genève. <https://cds.cern.ch/record/2157687>.
- [21] ATLAS COLLABORATION. (2017a). Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Physical Review D*, 96(7). 10.1103/PhysRevD.96.072002.
- [22] ATLAS COLLABORATION. (2017b). Jet reconstruction and performance using particle flow with the ATLAS Detector. *European Physical Journal C*, 77(466). 10.1140/epjc/s10052-017-5031-2.
- [23] ATLAS COLLABORATION. (2017c). Measurement of the Higgs boson mass in the H to ZZ^* to 4ℓ and H to $\gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector. Rapport technique, CERN, Genève. <http://cdsweb.cern.ch/record/2273853/>.
- [24] ATLAS COLLABORATION. (2017d). Performance of the ATLAS Track Reconstruction

- Algorithms in Dense Environments in LHC Run 2. *European Physical Journal C*, 77(673). 10.1140/epjc/s10052-017-5225-7.
- [25] ATLAS COLLABORATION. (2017e). Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *The European Physical Journal C*, 77(490). 10.1140/epjc/s10052-017-5004-5.
- [26] ATLAS COLLABORATION. (2018a). Luminosity Public Results Run 2. <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>.
- [27] ATLAS COLLABORATION. (2018b). Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV. *European Physical Journal C*, 78(903). 10.1140/epjc/s10052-018-6288-9.
- [28] ATLAS COLLABORATION. (2019a). In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector. *European Physical Journal C*, 79(135). 10.1140/epjc/s10052-019-6632-8.
- [29] ATLAS COLLABORATION. (2019b). Performance of top-quark and W -boson tagging with ATLAS in Run 2 of the LHC. *European Physical Journal C*, 79(375). 10.1140/epjc/s10052-019-6847-8.
- [30] ATLAS COLLABORATION. (2019c). Summary Plots for Exotics Heavy Particle Searches and Exotics/SUSY Long-lived Particle Searches. Rapport technique, CERN, Genève. <http://cds.cern.ch/record/2682064>.
- [31] ATLAS COLLABORATION. (2020a). Evolution of Regional, Age and Gender Demographics in the ATLAS Collaboration. ATL-OREACH-SLIDE-2020-057. <http://cds.cern.ch/record/2711704?ln=en>.
- [32] ATLAS COLLABORATION. (2020b). Standard Model Summary Plots Spring 2020. Rapport technique, CERN, Genève. <https://cdsweb.cern.ch/record/2718937>.
- [33] ATLAS COLLABORATION. (2020c). SUSY July 2020 Summary Plot Update. Rapport technique, CERN, Genève. <http://cds.cern.ch/record/2725258>.
- [34] ATLAS COLLABORATION. (s.d.). ATLAS Open Data. <http://opendata.atlas.cern/>.
- [35] ATLAS EXPERIMENT. (s.d.a). ATLAS Diversity and Inclusion. <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/DiversityInclusion>.
- [36] ATLAS EXPERIMENT. (s.d.b). Trigger and Data Acquisition. <https://atlas.cern/discover/detector/trigger-daq>.
- [37] ATLAS TRT COLLABORATION. (2012). The ATLAS TRT and its performance at LHC. *Journal of Physics : Conference Series*, 347. 10.1088/1742-6596/347/1/012025.
- [38] AVOLIO, G., BALLESTRERO, S. et VANDELLI, W. (2012). ATLAS TDAQ system : current status and performance. *Physics Procedia*, 37. 10.1016/j.phpro.2012.03.755.
- [39] BAER, H. et TATA, X. (2006). *Weak Scale Supersymmetry*. Cambridge University Press.
- [40] BALDI, P., SADOWSKI, P. et WHITESON, D. (2014). Searching for exotic particles

- in high-energy physics with deep learning. *Nature Communications*, 5(4308). 10.1038/ncomms5308.
- [41] BERRAH, N. (2007). Gender Equity : Strengthening the Physics Enterprise in Universities and National Laboratories. Rapport technique, American Physical Society. <https://www.aps.org/programs/women/workshops/gender-equity/upload/genderequity.pdf>.
- [42] BERTELLA, C. (2011). Fully hadronic ttbar cross section measurement with ATLAS detector. *arXiv :[hep-ex]*. 1111.3745v1.
- [43] BLANCE, A., SPANNOVSKY, M. et WAITE, P. (2019). Adversarially-trained autoencoders for robust unsupervised new physics searches. *Journal of High Energy Physics*, 2019(47). 10.1007/JHEP10(2019)047.
- [44] BRÜNING, O. et ROSSI, L. (2020). Chapter 1 : High-Luminosity Large Hadron Collider. *CERN Yellow Reports : Monographs*, 10(2019). 10.23731/CYRM-2020-0010.1.
- [45] BUREAU DE VALORISATION DE LA LANGUE FRANÇAISE ET DE LA FRANCO-PHONIE. (2019). *Inclusivement : guide d'écriture pour tous et toutes*. https://francais.umontreal.ca/fileadmin/francophonie/documents/Guide_de_redaction_inclusive/UdeM_Guide-ecriture-inclusive.pdf.
- [46] BUTTNER, E. H., LOWE, K. B. et BILLINGS-HARRIS, L. (2010). Diversity climate impact on employee of color outcomes : Does justice matter? *Career Development International*, 15(3). 10.1108/13620431011053721.
- [47] BÄRNREUTHER, P. (2012). *Top Quark Pair Production at the LHC*. Thèse de doctorat, RWTH Aachen University, Aix-la-Chapelle. https://inis.iaea.org/collection/NCLCollectionStore/_Public/45/012/45012790.pdf?r=1&r=1.
- [48] CACCIARI, M., SALAM, G. P. et SOYEZ, G. (2008). The anti-kT jet clustering algorithm. *Journal of High Energy Physics*, 2008(4). 10.1088/1126-6708/2008/04/063.
- [49] CAKIR, A. (2015). Searches for Beyond the Standard Model Physics at the LHC : Run1 Summary and Run2 Prospects. *Proceedings of Science*, FPCP2015:024. 10.22323/1.248.0024.
- [50] CAPEANS, M., DARBO, B., EINSWEILLER, K. *et al.* (2010). ATLAS Insertable B-Layer Technical Design Report. Rapport technique, CERN, Genève. <https://cds.cern.ch/record/1291633>.
- [51] CARON, J.-L. (1986). Vue aérienne du CERN. CERN. <http://cds.cern.ch/record/841506>.
- [52] CAST COLLABORATION. (2020). CAST - Status Report to the SPSC for the 139th Meeting. Rapport technique, CERN, Genève. <https://cds.cern.ch/record/2738387>.
- [53] CERN. (2020). Rapport sur l'environnement 2017–2018. Rapport technique, CERN, Genève. <https://doi.org/10.25325/CERN-Environment-2020-001>.
- [54] CERN. (s.d.a). Chiffres clés du LHC. <https://home.cern/fr/resources/faqs/facts-and-figures-about-lhc>.

- [55] CERN. (s.d.b). La naissance du Web. <https://home.cern/fr/science/computing/birth-web>.
- [56] CERN. (s.d.c). Nos États membres. <https://home.cern/fr/about/who-we-are/our-governance/member-states>.
- [57] CHASSANGRE, K. et CALLAHAN, S. (2017). « J’ai réussi, j’ai de la chance... je serai démasqué » : revue de littérature du syndrome de l’imposteur. *Pratiques Psychologiques*, 23(2). 10.1016/j.prps.2017.01.001.
- [58] CHENG, T., ARGUIN, J.-F., LEISSNER-MARTIN, J., PILETTE, J. et GOLLING, T. (2020). Variational Autoencoders for Anomalous Jet Tagging. *arXiv :[hep-ph]*. 2007.01850.
- [59] CMS COLLABORATION. (2008). The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08). <https://doi.org/10.1088/1748-0221/3/08/S08004>.
- [60] CMS COLLABORATION. (2012). Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Physics Letters B*, 716(1). 10.1016/j.physletb.2012.08.021.
- [61] CORBETT, C. et HILL, C. (2015). Solving the Equation : The Variables for Women’s Success in Engineering and Computing. Rapport technique, American Association of University Women. <https://www.aaup.org/app/uploads/2020/03/Solving-the-Equation-report-nsa.pdf>.
- [62] CORRADI, M. (2015). The ATLAS Muon System. [Présentation]. Istituto Nazionale di Fisica Nucleare, Rome. <https://www.roma1.infn.it/~bini/AtlasMuons.pdf>.
- [63] CYBENKO, G. (1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals, and Systems*, 2. 10.1007/BF02551274.
- [64] CZAKON, M., HEYMES, D. et MITOV, A. (2016). Bump hunting in LHC $t\bar{t}$ events. *Physical Review D*, 94(11). 10.1103/PhysRevD.94.114033.
- [65] DE SWART, J. G., BERTONE, G. et VAN DONGEN, J. (2017). How Dark Matter Came to Matter. *Nature Astronomy*, 1(0059). 10.1038/s41550-017-0059.
- [66] DEW, M., PERRY, J., FORD, L. *et al.* (2021). Gendered performance differences in introductory physics : A study from a large land-grant university. *Physical Review Physics Education Research*, 17(1). 10.1103/PhysRevPhysEducRes.17.010106.
- [67] D’IGNAZIO, C. et KLEIN, L. F. (2020). *Data Feminism*. The MIT Press.
- [68] DOKSHITZER, Y. L., LEDER, G. D., MORETTI, S. *et al.* (1997). Better Jet Clustering Algorithms. *Journal of High Energy Physics*, 1997(08). <https://doi.org/10.1088/1126-6708/1997/08/001>.
- [69] EDITORIAL BOARD. (2018). Science benefits from diversity. *Nature*, 558(5). 10.1038/d41586-018-05326-3.
- [70] EDITORIAL BOARD. (2020). Rethinking conferences. *Nature Reviews Physics*, 2(67). 10.1038/s42254-020-0151-2.
- [71] ELLIS, J. R. (2002). Limits of the Standard Model. Rapport technique, CERN, Genève.

<https://cds.cern.ch/record/591472>.

- [72] ERULKAR, S. D. (2020). Nervous system. <https://www.britannica.com/science/nervous-system>.
- [73] EVANS, L. et BRYANT, P. (2008). LHC Machine. *Journal of Instrumentation*, 3(08). <https://doi.org/10.1088/1748-0221/3/08/S08001>.
- [74] FABJAN, C. W. et GIANOTTI, F. (2003). Calorimetry for particle physics. *Review of Modern Physics*, 75(4). 10.1103/RevModPhys.75.1243.
- [75] FARINA, M., NAKAI, Y. et SHIH, D. (2020). Searching for new physics with deep autoencoders. *Physical Review D*, 101(7). 10.1103/PhysRevD.101.075021.
- [76] FITZPATRICK, C., OLIVETTI, E., MILLER, T. R. *et al.* (2015). Conflict Minerals in the Compute Sector : Estimating Extent of Tin, Tantalum, Tungsten, and Gold Use in ICT Products. *Environmental Science and Technology*, 49(2). 10.1021/es501193k.
- [77] FLEISCHMANN, S. (2013). Boosted top quark techniques and searches for $t\bar{t}$ resonances at the LHC. *Journal of Physics : Conference Series*, 452(1). 10.1088/1742-6596/452/1/012034.
- [78] FRIXIONE, S., NASON, P. et OLEARI, C. (2007). Matching NLO QCD computations with parton shower simulations : the POWHEG method. *Journal of High Energy Physics*, 11. 10.1088/1126-6708/2007/11/070.
- [79] FUNDAMENTAL PHYSICS BREAKTHROUGH PRIZE. (2013). Laureates - Fabiola Gianotti. <https://breakthroughprize.org/Laureates/1/L48>.
- [80] GAGNON, P. (2015). *Qu'est-ce que le boson de Higgs mange en hiver et autres détails essentiels*. Multimondes, Montréal.
- [81] GEANT4 COLLABORATION. (2020). GEANT4 – a simulation toolkit. *Nuclear Instruments and Methods in Physics A*, 506(3). 10.1016/S0168-9002(03)01368-8.
- [82] GENDER SUMMIT 6. (2015). GS6 Speakers - Londa Schiebinger. <https://gender-summit.com/gs6-speakers/1012-schiebinger-gs6a>.
- [83] GIAGU, S. (2019). WIMP Dark Matter Searches With the ATLAS Detector at the LHC. *Frontiers in Physics*, 7(75). 10.3389/fphy.2019.00075.
- [84] GIANOTTI, F. (2017). L'intolérance n'a plus sa place parmi nous. CERN. <https://home.cern/fr/news/opinion/general-cern/consigning-intolerance-history>.
- [85] GOODFELLOW, I., BENGIO, Y. et COURVILLE, A. (2016). *Deep Learning*. MIT Press.
- [86] GREEN, D. (2000). *The Physics of Particle Detectors*. Cambridge University Press.
- [87] GREENPEACE (2017). Clicking Clean : Who is Winning The Race to Build a Green Internet? Rapport technique. <https://www.greenpeace.org/usa/ending-the-climate-crisis/click-clean/>.
- [88] GRIFFITHS, D. (2008). *Introduction to Elementary Particles*. Wiley-VCH, 2e édition.
- [89] GROJEAN, C., SALVIONI, E. et TORRE, R. (2011). A weakly constrained W' at the early LHC. *Journal of High Energy Physics*, 07(2). 10.1007/JHEP07(2011)002.

- [90] GROOM, D. E., MOKHOV, N. V. et STRIGANOV, S. I. (2002). Muon Stopping Power and Range Tables 10 MeV - 100 TeV. *Atomic Data and Nuclear Data Tables*, 78(2). <https://doi.org/10.1006/adnd.2001.0861>.
- [91] GU, P.-H. (2021). Neutrinoless double beta decay without Majorana neutrinos. *arXiv :[hep-ph]*. 2101.00212.
- [92] HAFFNER, J. (2013). Complexe des accélérateurs du CERN. CERN. <https://cds.cern.ch/record/1621894>.
- [93] HANSER, D. (2006). *Architecture of France*. Greenwood Publishing Group.
- [94] HAWKING, S. (2002). *The Theory of Everything : The Origin and Fate of the Universe*. New Millennium Entertainment.
- [95] HEIMEL, T., KASIECZKA, G., PLEHN, T. *et al.* (2019). QCD or what ? *SciPost Physics*, 6(3). 10.21468/SciPostPhys.6.3.030.
- [96] HENDRYCKS, D., MAZEIKA, M. et DIETTERICH, T. (2019). Deep Anomaly Detection with Outlier Exposure. *arXiv :[cs.LG]*. 1812.04606.
- [97] HILL, C., CORBETT, C. et ST. ROSE, A. (2010). Why So Few ? Rapport technique, American Association of University Women. <https://www.aauw.org/app/uploads/2020/03/why-so-few-research.pdf>.
- [98] HOFMANN, H., SCHLEPER, M. C. et BLOME, C. (2018). Conflict Minerals and Supply Chain Due Diligence : An Exploratory Study of Multi-tier Supply Chains. *Journal of Business Ethics*, 147. 10.1007/s10551-015-2963-z.
- [99] INSTITUT DE VALORISATION DES DONNÉES. (2018). *Cadre de référence d'IVADO en matière d'équité, de diversité et d'inclusion*. <https://ivado.ca/PDF/Cadre-reference-EDI/>.
- [100] INSTITUT DE VALORISATION DES DONNÉES. (s.d.). Équité, Diversité, Inclusion. <https://ivado.ca/equite-diversite-inclusion/>.
- [101] IVIE, R., TEFAYE, C. L., CZUJKO, R. *et al.* (2013). The global survey of physicists : A collaborative effort illuminates the situation of women in physics. *AIP Conference Proceedings*, 1517(1). 10.1063/1.4794221.
- [102] JEGERLEHNER, F. (2019). The Hierarchy Problem and the Cosmological Constant Problem Revisited. *Foundations of Physics*, 49(9). 10.1007/s10701-019-00262-2.
- [103] KASIECZKA, G., PLEHN, T., BUTTER, A. *et al.* (2019). The Machine Learning Landscape of Top Taggers. *SciPost Physics*, 7(1). 10.21468/scipostphys.7.1.014.
- [104] KINGMA, D. P. et BA, J. (2014). ADAM : A Method for Stochastic Approximation. *arXiv :[cs.LG]*. 1412.6980.
- [105] KOMISKE, P. T., METODIEV, E. M. et THALER, J. (2019). Metric Space of Collider Events. *Physical Review Letters*, 123(4). 10.1103/PhysRevLett.123.041801.
- [106] KOMISKE, P. T., METODIEV, E. M. et THALER, J. (2020). The Hidden Geometry of Particle Collisions. *Journal of High Energy Physics*, 2020(6). 10.1007/JHEP07(2020)006.

- [107] KROHN, D., THALER, J. et WANG, L.-T. (2010). Jet Trimming. *Journal of High Energy Physics*, 2010(84). 10.1007/JHEP02(2010)084.
- [108] LABELLE, A. (2019). Filles et garçons égaux en mathématique. *Radio-Canada*. <https://ici.radio-canada.ca/nouvelle/1388422/genre-filles-garcons-mathematiques-science>.
- [109] LARIVIÈRE, V., PONTILLE, D. et SUGIMOTO, C. R. (2021). Investigating the division of scientific labor using the Contributor Roles Taxonomy (CRediT). *Quantitative Science Studies*, 2(1). 10.1162/qss_a_00097.
- [110] LARKOSKI, A. J., MOULT, I. et NEILL, D. (2014). Power Counting to Better Jet Observables. *Journal of High Energy Physics*, 2014(12). 10.1007/jhep12(2014)009.
- [111] LHCb COLLABORATION. (2008). The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08). <https://doi.org/10.1088/1748-0221/3/08/S08005>.
- [112] LHCb COLLABORATION. (2021). Test of lepton universality in beauty-quark decays. *arXiv :[hep-ex]*. 2103.11769.
- [113] MACALUSO, B., LARIVIÈRE, V., SUGIMOTO, T. *et al.* (2016). Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship. *Academic Medicine*, 91(8). 10.1097/ACM.0000000000001261.
- [114] MADHUKAR, V. et SHARMA, S. (2017). Organisational climate : A conceptual perspective. *International Journal of Management and Business*, 7(8). https://www.ijmra.us/project%20doc/2017/IJMIE_AUGUST2017/IJMRA-12034.pdf.
- [115] MALM, S. (2016). Homophobia row at CERN : LGBT society subjected to abuse at home of Large Hadron Collider with group posters defaced with bible quotes and 'pigs' graffiti. *Daily Mail*. <https://www.dailymail.co.uk/news/article-3502362/Homophobia-row-CERN-LGBT-society-subjected-abuse-home-Large-Hadron-Collider-group-posters-defaced-bible-quotes-pigs-graffiti.html>.
- [116] MARZANI, S., SOYEZ, G. et SPANNOVSKY, M. (2019). *Looking Inside Jets : An Introduction to Jet Substructure and Boosted-object Phenomenology*, volume 958 de *Lecture Notes in Physics*. Springer, Suisse.
- [117] MEDIPIX COLLABORATION. (s.d.). Medipix. <https://medipix.web.cern.ch/home>.
- [118] MISHRA, A. (2015). Impact of silica mining on environment. *Journal of Geography and Regional Planning*, 8(6). 5897/JGRP2015.0495.
- [119] MIYAKE, A., KOST-SMITH, L. E., FINKELSTEIN, N. D. *et al.* (2010). Reducing the gender achievement gap in college science : A classroom study of values affirmation. *Science*, 330(6008). 10.1126/science.1195996.
- [120] MOHAMADI, I. (2017). A mathematical proof for the existence of a possible source for dark energy. *arXiv :[math.FA]*. 1704.04430.
- [121] N. ASBAH, C. Leroy, S. P. *et al.* (2014). Measurement of the efficiency of the pattern recognition of tracks generated by ionizing radiation in a TIMEPIX detector. *Journal of*

- Instrumentation*, 9(5). 10.1088/1748-0221/9/05/c05021.
- [122] NATIONAL WOMEN’S HALL OF FAME. (s.d.). Chien-Shiung Wu. <https://www.womenofthehall.org/inductee/chienchiung-wu/8>.
- [123] NIELSEN, M. A. (2015). *Neural Networks and Deep Learning*. Determination Press.
- [124] OSQAR COLLABORATION. (2015). New Exclusion Limits for the Search of Scalar and Pseudoscalar Axion-Like Particles from Light Shining Through a Wall. *Phys. Rev. D*, 92(9). 10.1103/PhysRevD.92.092002.
- [125] PAGE, S. (2007). *The Difference : How the Power of Diversity Creates Better Groups*. Princeton University Press.
- [126] PARTICLE DATA GROUP (2020). Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2020(8). 10.1093/ptep/ptaa104.
- [127] PEQUENAO, J. (2008a). Computer Generated image of the ATLAS calorimeter. CERN. <https://cds.cern.ch/record/1095927>.
- [128] PEQUENAO, J. (2008b). Computer generated image of the ATLAS inner detector. CERN. <https://cds.cern.ch/record/1095926>.
- [129] PEQUENAO, J. (2008c). Computer generated image of the ATLAS Muons subsystem. CERN. <http://cds.cern.ch/record/1095929?ln=en>.
- [130] PEQUENAO, J. (2008d). Computer generated image of the whole ATLAS detector. CERN. <https://cds.cern.ch/record/1095924>.
- [131] PERKINS, D. (2000). *Introduction to High Energy Physics*. Cambridge University Press, 4e édition.
- [132] PHILLIPS, K. W., LILJENQUIST, K. A. et NEALE, M. A. (2009). Is the Pain Worth the Gain ? The Advantages and Liabilities of Agreeing With Socially Distinct Newcomers. *Personality and Social Psychology Bulletin*, 35(3). 10.1177/0146167208328062.
- [133] PICO COLLABORATION. (2016). Dark matter search results from the PICO-60 CF₃I bubble chamber. *Phys. Rev. D*, 93(5). 10.1103/PhysRevD.93.052014.
- [134] POSNER, T. et FEI-FEI, L. (2020). AI will change the world, so it’s time to change AI. *Nature*, 588(S118). 10.1038/d41586-020-03412-z.
- [135] PRÉVOST, H. (2018). Il y a 20 fois plus de déchets miniers que domestiques produits au Québec, avance une coalition. *Radio-Canada*. <https://ici.radio-canada.ca/nouvelle/1136952/politique-quebec-environnement-mine-impact>.
- [136] QUIGG, C. (1983). *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*. Princeton University Press, 2e édition.
- [137] QUINTON, J. N. (2020). Cutting the carbon cost of academic travel. *Nature Reviews Earth and Environment*, 1(13). 10.1038/s43017-019-0008-3.
- [138] RAYNER-CANHAM, M. F. et RAYNER-CANHAM, G. W. (1997). *A Devotion to Their Science, Pioneer : Women of Radioactivity*. McGill-Queen’s University Press.

- [139] ROSSI, L. (2003). The LHC Superconducting Magnets. (CERN-LHC-Project-Report-660). <https://cds.cern.ch/record/630341>.
- [140] SALAM, G. (2010). Towards Jetography. *The European Physical Journal C*, 67. <https://doi.org/10.1140/epjc/s10052-010-1314-6>.
- [141] SALVIONI, E., VILLADORO, G. et ZWIRNER, F. (2009). Minimal Z-prime models : Present bounds and early LHC reach. *Journal of High Energy Physics*, 11(068). 10.1088/1126-6708/2009/11/068.
- [142] SCIKIT-LEARN : MACHINE LEARNING IN PYTHON (s.d.). Quantile transformer. <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.QuantileTransformer.html>.
- [143] SJÖSTRAND, T., MRENNNA, S. et SKANDS, P. (2008). A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178(11). 10.1016/j.cpc.2008.01.036.
- [144] SNO COLLABORATION. (2002). Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Physical Review Letters*, 89(1). <https://doi.org/10.1103/physrevlett.89.011301>.
- [145] SOLAECHE, G. (2015). *Top tagging at the LHC experiments with proton-proton collisions at $\sqrt{s} = 13\text{TeV}$* . Thèse de doctorat, Université de València. 10.13140/RG.2.1.4521.4560.
- [146] SUPERCDMS COLLABORATION. (2020). Constraints on low-mass, relic dark matter candidates from a surface operated SuperCDMS single-charge sensitive detector. *Phys. Rev. D*, 102(9). 10.1103/physrevd.102.091101.
- [147] THALER, J. et VAN TILBURG, K. (2011). Identifying boosted objects with N-subjettiness. *Journal of High Energy Physics*, 2011(3). 10.1007/JHEP03(2011)015.
- [148] THOMSON, M. (2013). *Modern Particle Physics*. Cambridge University Press.
- [149] TISSEYRE, C. (ANIMATEUR) ET MONTPETIT, I. (JOURNALISTE). (2001). L'intelligence artificielle au service de la DPJ. Dans F. Charron (réalisatrice), *Découverte*. Société Radio-Canada. <https://curio.ca/fr/video/lintelligence-artificielle-au-service-de-la-dpj-3712/>.
- [150] TULLIO, I. D. (2019). Gender Equality in STEM : Exploring Self-Efficacy Through Gender Awareness. *Sociology of Education*, 11(3):226–245. 10.14658/pupj-ijse-2019-3-13.
- [151] TUMER, K. et GHOSH, J. (1996). Estimating the Bayes error rate through classifier combining. *Proceedings - International Conference on Pattern Recognition*, 2. 10.1109/ICPR.1996.546912.
- [152] UNESCO. (s.d.). Women in Science. <http://uis.unesco.org/en/topic/women-science>.
- [153] UNNO, Y. (2000). ATLAS silicon microstrip Semiconductor Tracker (SCT). *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 453(1). 10.1016/S0168-9002(00)00615-X.

- [154] U.S. BUREAU OF LABOR STATISTICS (2021). Labor Force Statistics from the Current Population Survey. BLS Beta Labs. <https://beta.bls.gov/dataViewer/view/timeseries/LNU02070002Q>.
- [155] VIDAL, C. (2011). Le cerveau a-t-il un sexe? *L'école des parents*, 593(6). 10.3917/e-par.593.0026.
- [156] WARDROPE, D. R. (2013). *Boosted Heavy Particles and Jet Substructure*. 43rd International Symposium on Multiparticle Dynamics, Chicago, IL, USA. <https://cds.cern.ch/record/1604479>.
- [157] WIKIMEDIA COMMONS. (2006). File : Standard Model of Elementary Particles.svg. https://commons.wikimedia.org/w/index.php?title=File:Standard_Model_of_Elementary_Particles.svg&oldid=341423479.
- [158] WIKIMEDIA COMMONS. (2007). File :Hqmc600.svg. <https://commons.wikimedia.org/wiki/File:Hqmc600.svg>.
- [159] WING SUE, D. (2020). *Microaggressions in Everyday Life : Race, Gender, and Sexual Orientation*. Wiley, 2e édition.
- [160] WOLFSON, N., KRAIGER, K. et FINKELSTEIN, L. (2011). The relationship between diversity climate perceptions and workplace attitudes. *The Psychologist-manager Journal*, 14(3). 10.1080/10887156.2011.546170.
- [161] WONG, C.-Y. (1994). *Introduction to High-Energy Heavy-Ion Collisions*. World Scientific Publishing Compagny.
- [162] XELLA, S. (2013). Physics object reconstruction in the ATLAS experiment. Rapport technique, CERN, Genève. <https://cds.cern.ch/record/1519113>.
- [163] XENON COLLABORATION. (2021). Search for inelastic scattering of WIMP dark matter in XENON1T. *Physical Review D*, 103(6). 10.1103/PhysRevD.103.063028.
- [164] ZHOU, L., PAN, S., WANG, J. *et al.* (2017). Machine learning on big data : Opportunities and challenges. *Neurocomputing*, 237. 10.1016/j.neucom.2017.01.026.
- [165] ZOLOTOV, O., ROMANOVSKAYA, Y. et RZHANNIKOVA, V. (2019). On Definition of BigData. *The European Physical Journal Conferences*, 224(04011). 10.1051/epj-conf/201922404011.